

Fugu Genome Analysis Provides Evidence for a Whole-Genome Duplication Early During the Evolution of Ray-Finned Fishes

Alan Christoffels,* Esther G. L. Koh,* Jer-ming Chia,* Sydney Brenner,* Samuel Aparicio,*† and Byrappa Venkatesh*

*Institute of Molecular and Cell Biology, Singapore; and †University of Cambridge, Department of Oncology, Hutchison-MRC Research Centre, Cambridge, UK

With about 24,000 extant species, teleosts are the largest group of vertebrates. They constitute more than 99% of the ray-finned fishes (Actinopterygii) that diverged from the lobe-finned fish lineage (Sarcopterygii) about 450 MYA. Although the role of genome duplication in the evolution of vertebrates is now established, its role in structuring the teleost genomes has been controversial. At least two hypotheses have been proposed: a whole-genome duplication in an ancient ray-finned fish and independent gene duplications in different lineages. These hypotheses are, however, based on small data sets and lack adequate statistical and phylogenetic support. In this study, we have made a systematic comparison of the draft genome sequences of Fugu and humans to identify paralogous chromosomal regions (“paralogons”) in the Fugu that arose in the ray-finned fish lineage (“fish-specific”). We identified duplicate genes in the Fugu by phylogenetic analyses of the Fugu, human, and invertebrate sequences. Our analyses provide evidence for 425 fish-specific duplicate genes in the Fugu and show that at least 6.6% of the genome is represented by fish-specific paralogons. We estimated the ages of Fugu duplicate genes and paralogons using the molecular clock. Remarkably, the ages of duplicate genes and paralogons are clustered, with a peak around 350 MYA. These data strongly suggest a whole-genome duplication event early during the evolution of ray-finned fishes, probably before the origin of teleosts.

Introduction

The teleosts, a monophyletic group of ray-finned fishes, are the largest and the most diverse group of vertebrates. With about 24,000 species, teleosts comprise about half the number of extant vertebrate species and more than 99% of ray-finned fishes (Actinopterygii) that diverged from the lobe-finned fish lineage (Sarcopterygii) about 450 MYA. Although the role of genome duplication(s) in structuring the genomes of vertebrates is now established (McLysaght, Hokamp, and Wolf 2002; Gu, Wang, and Gu 2002; Panopoulou et al. 2003), the role of gene and genome duplications in the evolution of teleosts has been controversial. The discovery of up to seven *Hox* clusters in teleosts such as zebrafish, Fugu, and medaka, compared with four *Hox* clusters in mammals, has led to the hypothesis that a whole-genome duplication occurred in the ray-finned fish lineage (“fish-specific”) after its divergence from the lobe-finned fish lineage (Amores et al. 1998; Aparicio et al. 2002; Naruse et al. 2000). It has also been proposed that the whole-genome duplication in the fish lineage provided the additional genetic material that spurred the radiation of teleosts (Amores et al. 1998; Wittbrodt, Meyer, and Schartl, 1998; Meyer and Schartl 1999; Taylor, Van de Peer, and Meyer 2001; Taylor et al. 2003). The mapping of duplicate copies of zebrafish genes (Gates et al. 1999; Barbazuk et al. 2000; Woods et al. 2000; Postlethwait et al. 2000; Taylor et al. 2003; Winkler et al. 2003) and Fugu genes (Smith et al. 2002; Aparicio et al. 2002; Yu, Brenner, and Venkatesh 2003) to pairs of linkage-groups or duplicate chromosome fragments, and the presence of orthologs for 22 out of 49 pairs of zebrafish duplicate genes in the Fugu (Taylor et al. 2003), appear to

support the fish-specific, whole-genome duplication hypothesis. However, phylogenetic analyses of some duplicate genes from zebrafish (Taylor et al. 2001, 2003) and other teleosts (Robinson-Rechavi et al. 2001a, 2001b) suggest that these genes might be the result of independent gene duplications in different lineages, rather than a whole-genome duplication in a common ancestor. Tracing the history of gene duplications in the teleost lineage is also confounded by the fact that ray-finned fishes appear to be more prone to tetraploidization than other vertebrates. Tetraploidization has occurred independently in several lineages of teleosts, as well as in “nonteleost” basal ray-finned fishes (see Venkatesh [2003]). Thus, in the absence of whole-genome sequences, phylogenetic analysis of small data sets may not reliably resolve the evolutionary origin of additional genes, particularly if the tetraploidization event is followed by large-scale secondary gene loss.

The studies of duplicate genes in teleosts so far are based on small data sets before the completion of a teleost genome and, thus, lack adequate statistical and phylogenetic support. Furthermore, because of the small size of the data sets, it has not been possible to estimate the precise ages of fish-specific duplicate genes. Thus, the extent and the timing of gene duplication events in the teleost lineage remain unclear. Recently, the draft genome sequence of a teleost, the Fugu, was completed (Aparicio et al. 2002). At 385 Mb, Fugu has one of the smallest genomes among vertebrates. The compact size of the genome has been attributed to the paucity of repetitive sequences and other nonessential sequences in the genome. In this study, we have made a systematic comparison of the Fugu and human genomes to estimate the extent of fish-specific paralogous chromosomal fragments (“paralogons”) and fish-specific duplicate genes in the Fugu genome. We estimated the ages of the Fugu duplicate genes using the molecular clock. Our results provide strong evidence for a whole-genome duplication early during the evolution of ray-finned

Key words: whole-genome duplication; paralogons, molecular clock, Fugu.

E-mail: mcbbv@imcb.nus.edu.sg.

Mol. Biol. Evol. 21(6):1146–1151. 2004

DOI:10.1093/molbev/msh114

Advance Access publication March 10, 2004

fishes. Interestingly, the whole-genome duplication appears to have occurred before the origin of teleosts, raising doubts about its role in the radiation of teleosts.

Methods

Sequences

A total of 31,059 proteins have been identified in the “draft” sequence of the Fugu (*Fugu rubripes*) genome (version 6.1.1; 12,403 scaffolds over 2 kb long; 332.5 Mb) (Aparicio et al. 2002). Sequences of these proteins were downloaded from the IMCB Fugu Web site (<http://www.fugu-sg.org>). The human proteome comprising 27,628 proteins was downloaded from Ensembl (release 9.30) (<http://www.ensembl.org/>), and 15,832 *Ciona intestinalis* (*Ciona*) proteins were obtained from the Joint Genome Institute ftp site (<http://genome.jgi-psf.org/Ciona4.download.ftp.html>). The 13,054 *Drosophila melanogaster* (fly) protein sequences were downloaded from the Berkeley *Drosophila* Genome Project (<http://www.fruitfly.org>) (April 2002), and the 20,414 *Caenorhabditis elegans* (nematode) peptides (wormpep77) were downloaded from the Sanger Center ftp site (<ftp://ftp.sanger.ac.uk>). We defined tandem duplicates in the Fugu genome as two proteins with a BlastP expectation value (e) of 10^{-15} or less and separated by a maximum of 20 genes. A total of 310 tandem duplicates were replaced with the longest protein. Alternative splice variants in the Fugu and other proteomes (human, *Ciona*, fly, and nematode) were filtered so that the longest variant was retained as the representative protein. Proteins shorter than 100 residues were also excluded. The final curated protein data sets corresponded to 30,749 Fugu, 21,310 human, 15,832 *Ciona*, 12,892 fly, and 20,414 nematode proteins.

Fugu-Human Protein Families

BlastP searches of 30,749 Fugu proteins were carried out against a combined human and Fugu protein data sets using the following parameters: (1) BLOSUM62 matrix with SEG filtering switched on; (2) expectation cutoff score of $1e^{-07}$; and (3) minimum 50% alignment length of the longer sequence covered by the alignment. This combination of parameters were arrived at after extensive exploration of various parameters, including those described by McLysaght, Hokamp, and Wolfe (2002) and Friedman and Hughes (2001). All Blast searches were carried out on a 75-node Compaq Alpha server, and the data were stored in a MySQL database. Of the 30,749 Fugu proteins, 3,257 did not meet the expectation score threshold ($1e^{-07}$). Another 3,983 proteins were excluded because they did not match 50% of the longest sequence. The remaining 23,509 Fugu proteins and their matching human sequences were grouped into families that each contained two to 248 proteins.

Paralogue Detection

All scaffolds were compared with each other to identify scaffolds that shared protein families. Scaffolds that shared more than one duplicate protein pair were identified as paralogous chromosomal segments (“para-

logons”), irrespective of the order and orientation of the genes. A maximum of 20 unrelated genes were allowed on a paralogon. The paralogon detection algorithm outlined above represents an interscaffold comparison and does not take into account any intrascaffold events. We downloaded data for 1,437 human paralogons that were reported to have originated during the early evolution of chordates (McLysaght, Hokamp, and Wolfe 2002).

Statistical Test for Randomness

We identified paralogons on 1,000 shuffled gene maps to test the statistical significance of our results. The number and the size of paralogons in the shuffled data were compared with the actual data using the t -test.

Fish-Specific Duplicate Genes and the Time of Duplications

To obtain outgroups for Fugu and human protein families, we identified *Ciona*, fly, or nematode orthologs for human proteins by a reciprocal Blast. A total of 3,781 *Ciona*-human orthologs, 1,967 fly-human orthologs, and 2,182 nematode-human orthologs were derived from these searches. After adding these invertebrate out-groups into the respective Fugu-human protein families, a total of 995 families that had one out-group sequence (*Ciona*, fly, or nematode), at least one human sequence and more than one Fugu sequence were retrieved.

The 995 families were analyzed for evidence of duplication events in the fish lineage by phylogenetic analysis. Alignments were generated using ClustalW with default parameters (Thompson, Higgins, and Gibson 1994), and tree topologies were generated by the PHYLIP programs, PROTDIST and NEIGHBOR (Felsenstein 1989). The gamma-corrected substitution rates were calculated using the program GAMMA (Gu and Zhang 1997). For 142 families, the program crashed because of some unexplained error in the data file. Neighbor-joining (NJ) trees were drawn for the remaining 853 protein families with 1,000 bootstraps. Because we were only interested in fish-specific duplications, we filtered out 506 families that did not show a duplication topology for the Fugu sequences. NJ trees were reconstructed for the remaining 347 families (425 duplication nodes). Phylogenetic trees were also reconstructed using the maximum-likelihood (ML) method. This method identified 425 fish-specific duplicate nodes similar to the NJ method, except that 12 of the duplicate gene pairs were different from that predicted by the NJ method. Results of only the NJ method are presented because the molecular clock analysis carried out for dating the duplication event (see below) also uses the distance-based NJ algorithm.

The two-cluster test for rate heterogeneity was applied to 347 families to test for deviation from the molecular clock at 5% significance using TPCV, a program within the LinearTree package (Takezaki, Rzhetsky, and Nei 1995). A total of 236 families did not satisfy the molecular clock hypothesis. Estimates of divergence time of these genes were nevertheless calculated to get an idea of their distribution pattern, and are shown as Supplementary

Table 1
Fish-Specific Paralogons in the Fugu Genome

Paralogon Size ^a	Number of Paralogons	Number of Genes ^b	Size (Mb)	Genome Coverage ^c (%)
2	383	1681	16	4.8
3	63	409	4.3	1.3
4	16	129	1.0	0.3
5	1	11	0.06	0.018
6	5	62	0.47	0.14
Total	468	2292	21.83	6.6

^a Number of duplicate genes present on the paralogon.^b Total number of duplicate genes and other genes on the paralogon.^c Calculation based on 332.5 Mb of draft Fugu genome sequence (Aparicio et al. 2002).

Material online at www.mbe.oupjournals.org. An additional 15 families showed negative or zero branch lengths. Linearized trees were drawn for the remaining 96 families. A total of eight linearized trees showed topology inconsistent with fish-specific duplications. The duplication dates of the 95 Fugu genes in the remainder 88 linearized trees were estimated relative to the divergence time of ray-finned fish and lobe-finned fish (450 Myr) (Kumar and Hedges 1998).

Results and Discussion

Fish-Specific Paralogons

The presence of a large number of paralogons is a distinctive feature of a genome that has experienced whole-genome duplication during its evolution. The sizes and the extent of paralogons are constrained by the secondary loss of genes and the decay of synteny. To determine if the Fugu lineage experienced a genome duplication, we estimated the extent of fish-specific paralogons in the Fugu genome. We generated families of related Fugu and human proteins using an objective set of homology criteria (expectation cutoff score of $1e^{-07}$ and greater than 50% identity), and then mapped families of Fugu proteins onto the Fugu scaffolds. We define paralogons as pairs of scaffolds that contained two or more paralogous genes with a maximum of 20 unrelated genes between them. Seventy-eight paralogons that shared more than one pair of paralogous genes with human paralogons (McLysaght, Hokamp, and Wolfe 2002) were eliminated because they could be the result of ancient chordate duplications that occurred before the ray-finned fish and lobe-finned fish split. The remaining 468 paralogons represent fish-specific paralogons that arose in the ray-finned fish lineage (table 1). However, this set may contain paralogons that arose before the ray-finned fish and lobe-finned fish split but degraded in the mammalian lineage because of secondary gene loss or disruption of synteny. Such paralogons should contain duplicate genes that diverged more than 450 MYA.

It is possible that some paralogons could be formed by chance, after independent gene duplications and transposition of duplicated genes to the same region (Smith, Knight, and Hurst 1999). To test the statistical

Table 2
Paralogons in the Real Draft Fugu Genome Sequence and on 1,000 Simulations of the Draft Genome

Paralogon Size ^a	Number of Paralogons	Mean Number of Paralogons	SD	<i>t</i> -test		
				<i>P</i> -value ^b	Z-score	<i>P</i> -value
2	383	114.6	16.2	-549	<0.005	0.0019
3	63	1.9	1.5	-1402	<0.005	0.0019
4	16	0				
5	1	0				
6	5	0				

^a Number of duplicate gene pairs on a paralogon.^b Probability of finding X ("real genome paralogons") or more paralogons in 1,000 simulations; $X \geq S_{m, \text{real-genome}}/1,000$.

significance of the Fugu paralogons, we generated 1,000 randomly shuffled gene maps of Fugu scaffolds and searched for paralogons (table 2). The probability of finding a paralogon with even two pairs of duplicate genes on such simulated gene maps was found to be extremely low ($P \leq 0.0019$), indicating that few, if any, paralogons identified by us are formed by chance.

The fish-specific paralogons in the Fugu contain two to six paralogous genes, with zero to 16 unrelated genes. Indeed, the majority of paralogons (404 of 468) contain four or fewer unrelated genes. Some representative Fugu paralogons are shown in figure 1. (All the paralogons can be viewed at <http://www.fugu-sg.org/docs/publications.html>). Overall, the fish-specific paralogons we identified span 22 Mb and cover 6.6% of the Fugu genome (table 1). This estimate is most likely on the lower side because of the fragmented nature of the Fugu draft genome sequence.

Fish-Specific Duplicate Genes

All the duplicate genes in a genome are unlikely to be represented on paralogons because of the chromosomal rearrangements that occurred after the duplication event. Therefore, as an independent estimate of the extent of gene duplications, we determined the number of fish-specific duplicate genes in Fugu using the phylogenetic approach. Phylogenetic analysis is the best method available for assigning orthology and paralogy to genes from distant lineages and, thus, for identifying lineage-specific duplicate genes. We generated phylogenetic trees of Fugu and human protein families using invertebrate (*Ciona*, fly, or nematode) orthologs as outgroups and screened for topologies that suggest fish-specific duplications (fig. 2). A total of 425 fish-specific duplicate nodes were identified in 347 families. These genes represent a well-defined set of duplicate genes that arose in the ray-finned fish lineage (the descriptions of all the duplicate genes are given in table S1 of Supplementary Material online; the phylogenetic trees can be viewed at <http://www.fugu-sg.org/docs/publications.html>). This set is an underestimate of the duplicate genes in the Fugu because a large number of proteins were eliminated before the reconstruction of phylogenetic trees. Such proteins either lacked invertebrate

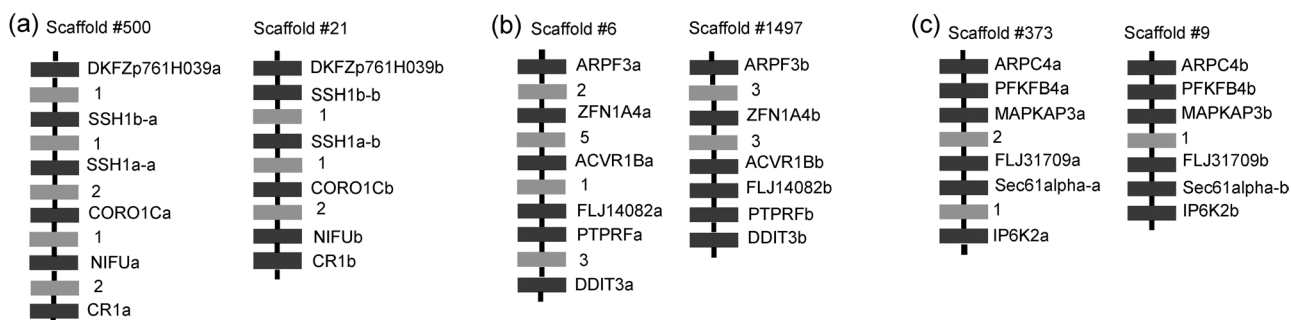


FIG. 1.—Representative fish-specific paralogs (*a–c*) in the Fugu genome. Names of duplicate genes (black boxes) are shown. The number of single genes present in a locus (gray boxes) is indicated. Scaffolds are nonoverlapping fragments of Fugu genomic sequence.

orthologs, or the software failed to calculate the gamma-corrected values (see *Methods*).

Time of Gene Duplications

To estimate the divergence time of Fugu duplicate genes, we reconstructed linearized trees for duplicate genes under the assumption of a molecular clock. The timing for the 95 duplication nodes in 88 protein families that satisfied the molecular clock criterion was estimated relative to the divergence time of lobe-finned fishes and ray-finned fishes (450 Myr) (Kumar and Hedges 1998). Remarkably, two-thirds of the duplications lie between 298 Myr and 393 Myr, with a peak at 345 Myr (fig. 3). These data suggest that large-scale gene duplications occurred in the ray-finned fish lineage around 350 MYA. (The phylogenetic trees for these protein families can be viewed on our Web page at <http://www.fugu-sg.org/docs/publications.html>).

Whole-Genome Duplication

The fish-specific duplicate genes in the Fugu could be the result of large-scale gene duplications that were selected for at a certain time or a single event such as the whole-genome duplication. If the latter is true, then more of the Fugu duplicate genes than expected by chance should be located on paralogs. Our set of 468 Fugu paralogs cover 6.6% of the genome. Therefore, one would expect at least 6.6% of the duplicate genes to be on the paralogs. In fact, 141 out of the 425 duplicate genes (33%) are located on the paralogs (table S2 in Supplementary Material online), indicating that the duplicate genes arose as a result of a single event. This overlap between the duplicate genes

and the paralogs indicate that the paralogs identified by us contain well-defined duplicate genes and extends the set of duplicate genes to include other paralogous genes on the paralogs for which we could not generate phylogenetic trees. Furthermore, mapping the ages of duplicate genes onto the paralogs shows that a large number of paralogs also originated around 350 MYA (figure S1 in Supplementary Material online). Thus, our independent data sets of paralogs and duplicate genes strongly suggest that a large number of duplicate genes were generated in the ray-finned fish lineage around 350 MYA as a result of a single event and not because of independent gene duplications. These results are consistent with a whole-genome duplication event in a ray-finned ancestor that gave rise to the Fugu lineage. Thus, our results support the fish-specific, “whole-genome duplication” hypothesis.

Evolutionary Implications

Based on the presence of orthologous duplicate genes in phylogenetically distant species of teleosts such as zebrafish and Fugu, it has been suggested that an ancient whole-genome duplication provided the additional genetic material that facilitated the radiation of teleosts (Amores et al. 1998; Postlethwait et al. 1998; Meyer and Schartl 1999; Taylor et al. 2003). Our estimation of the ages of duplicate genes in Fugu suggests that the whole-genome duplication occurred in the ray-finned fish lineage around 350 MYA. Interestingly, paleontological evidence suggest that teleosts first appeared around 220 MYA and underwent rapid diversification during the Jurassic and Cretaceous periods (205 to 135 MYA) (Maissey 1996). Thus, the whole-genome duplication appears to have occurred before the origin of teleosts. Alternatively, this could be a case of

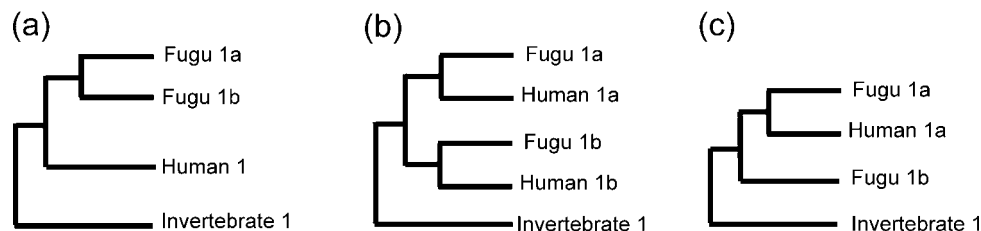


FIG. 2.—Schematic representation of some alternative phylogenetic tree topologies representing gene duplication events in vertebrates. Neighbor-joining trees were generated for closely related Fugu and human protein families using *Ciona*, fruit fly, or nematode as the invertebrate outgroup. (*a*) Fish-specific gene duplication. (*b*) Vertebrate-specific gene duplication. (*c*) Possible vertebrate-specific duplication followed by loss of one copy in the human lineage.

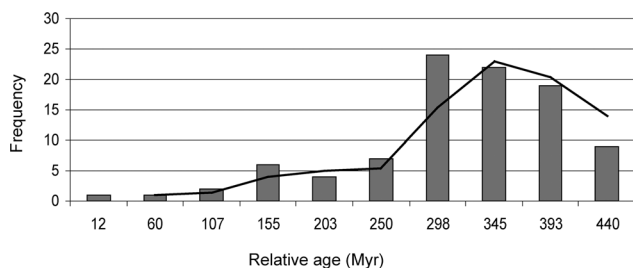


FIG. 3.—Distribution of the estimated ages of fish-specific gene duplications. Ages were estimated relative to the divergence time of lobe-finned fishes and ray-finned fishes (450 Myr). The curve represents the trendline drawn using moving averages.

a dramatic discordance between the molecular data and fossil evidence. If the genome duplication in the ray-finned fish lineage did indeed occur before the origin of teleosts, then genome duplication may not have been the driving force behind the radiation of teleosts. The basal “non-teleost” ray-finned fishes are represented by only four living groups: the polypteriforms (e.g., bichirs), acipenseriforms (sturgeons and paddlefish), semionotiforms (gars), and amiiforms (bowfin). To establish a correlation between the whole-genome duplication and the radiation of teleosts, it would be important to determine the time of the whole-genome duplication in relation to the speciation events of these “nonteleost” ray-finned fishes. Characterization of duplicate genes and paralogons in the basal “nonteleost” ray-finned fishes, as well as in basal teleosts such as osteoglossomorphs (e.g., bonytongues) and elopomorphs (e.g., eels) should help to determine whether the whole-genome duplication event in the ray-finned fish lineage spurred the radiation of teleosts.

Supplementary Material

Table S1 presents descriptions of fish-specific duplicate genes in the *Fugu* genome. Table S2 shows paralogons containing fish-specific *Fugu* duplicate gene pairs identified by phylogenetic analysis. Figure S1 illustrates distribution of the estimated ages of duplicate genes located on the fish-specific paralogons.

Web Site References

<http://www.fugu-sg.org>, IMCB *Fugu* home page.
<http://genome.jgi-psf.org/Ciona4.download.ftp.html>, Joint Genome Institute *Ciona* database.
<http://www.fruitfly.org>, Berkeley *Drosophila* Genome Project.
<ftp://ftp.sanger.ac.uk>, Sanger Center ftp site.
 Figures for the paralogons and phylogenetic trees can be viewed at <http://www.fugu-sg.org/docs/publications.html>.

Acknowledgments

We thank Kenneth Wolfe for providing the details of the human paralogons. This work was supported by the Agency for Science, Technology and Research (A*STAR) of Singapore.

Literature Cited

- Amores, A., A. Force, Y. L. Yan et al. (13 co-authors). 1998. Zebrafish hox clusters and vertebrate genome evolution. *Science* **282**:1711–1714.
- Aparicio, S., J. Chapman, E. Stupka et al. (41 co-authors). 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**:1301–1310.
- Barbazuk, W. B., I. Korf, C. Kadavi, J. Heyen, S. Tate, E. Wun, J. A. Bedell, J. D. McPherson, and S. L. Johnson. 2000. The syntenic relationship of the zebrafish and human genomes. *Genome Res.* **10**:1351–1358.
- Felsenstein, J. 1989. PHYLIP (phylogeny inference package). Version 3.2. *Cladistics* **5**:164–166.
- Friedman, R., and A. L. Hughes. 2001. Pattern and timing of gene duplication in animal genomes. *Genome Res.* **11**:1842–1847.
- Gates, M. A., L. Kim, E. S. Egan, T. Cardozo, H. I. Sirotkin, S. T. Dougan, D. Lashkari, R. Abagyan, A. F. Schier, and W. S. Talbot. 1999. A genetic linkage map for zebrafish: comparative analysis and localization of genes and expressed sequences. *Genome Res.* **9**:334–347.
- Gu, X., Y. Wang, and J. Gu. 2002. Age distribution of human gene families shows significant roles of both large- and small-scale duplications in vertebrate evolution. *Nat. Genet.* **31**:205–209.
- Gu, X., and J. Zhang. 1997. A simple method for estimating the parameter of substitution rate variation among sites. *Mol. Biol. Evol.* **14**:1106–1113.
- Kumar, S., and S. B. Hedges. 1998. A molecular timescale for vertebrate evolution. *Nature* **392**:917–920.
- Maissey, J. G. 1996. *Discovering fossil fishes*. Henry Holt and Company, New York.
- McLysaght, A., K. Hokamp, and K. H. Wolfe. 2002. Extensive genomic duplication during early chordate evolution. *Nat. Genet.* **31**:200–204.
- Meyer, A. and M. Schartl. 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr. Opin. Cell. Biol.* **11**:699–704.
- Naruse, K., S. Fukamachi, H. Mitani et al. (20 co-authors). 2000. A detailed linkage map of medaka, *Oryzias latipes*: comparative genomics and genome evolution. *Genetics* **154**:1773–1784.
- Panopoulou, G., S. Hennig, D. Groth, A. Krause, A. J. Poustka, R. Herwig, M. Vingron, and H. Lehrach. 2003. New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes. *Genome Res.* **13**:1056–1066.
- Postlethwait, J. H., I. G. Woods, P. Ngo-Hazelett, Y. L. Yan, P. D. Kelly, F. Chu, H. Huang, A. Hill-Force, and W. S. Talbot. 2000. Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Res.* **10**:1890–1902.
- Postlethwait, J. H., Y. L. Yan, M. A. Gates et al. (26 co-authors). 1998. Vertebrate genome evolution and the zebrafish gene map. *Nat. Genet.* **18**:345–349.
- Robinson-Rechavi, M., A. S. Carpentier, M. Duffraisse, and V. Laudet. 2001a. How many nuclear hormone receptors are there in the human genome? *Trends Genet.* **17**:554–556.
- Robinson-Rechavi, M., O. Marchand, H. Escriva, and V. Laudet. 2001b. An ancestral whole-genome duplication may not have been responsible for the abundance of duplicated fish genes. *Curr. Biol.* **11**:R458–R459.
- Smith, N. G., R. Knight, and L. D. Hurst. 1999. Vertebrate genome evolution: a slow shuffle or a big bang? *Bioessays* **21**:697–703.

- Smith, S. F., P. Snell, F. Gruetzner, A. J. Bench, T. Haaf, J. A. Metcalfe, A. R. Green, and G. Elgar. 2002. Analyses of the extent of shared synteny and conserved gene orders between the genome of *Fugu rubripes* and human 20q. *Genome Res.* **12**:776–784.
- Takezaki, N., A. Rzhetsky, and M. Nei. 1995. Phylogenetic test of the molecular clock and linearized trees. *Mol. Biol. Evol.* **12**:823–833.
- Taylor, J. S., I. Braasch, T. Frickey, A. Meyer, and Y. Van de Peer. 2003. Genome duplication, a trait shared by 22000 species of ray-finned fish. *Genome Res.* **13**:382–390.
- Taylor, J. S., Y. Van de Peer, I. Braasch, and A. Meyer. 2001. Comparative genomics provides evidence for an ancient genome duplication event in fish. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **356**:1661–1679.
- Taylor, J. S., Y. Van de Peer, and A. Meyer. 2001. Genome duplication, divergent resolution and speciation. *Trends Genet.* **17**:299–301.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 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.
- Venkatesh, B. 2003. Evolution and diversity of fish genomes. *Curr. Opin. Genet. Dev.* **13**:1–5.
- Winkler, C., M. Schafer, J. Duschl, M. Schartl, and J. N. Volff. 2003. Functional divergence of two zebrafish midline growth factors following fish-specific gene duplication. *Genome Res.* **13**:1067–1081.
- Wittbrodt, J., A. Meyer, and M. Schartl. 1998. More genes in fishes? *Bioessays* **20**:511–515.
- Woods, I. G., P. D. Kelly, F. Chu, P. Ngo-Hazelett, Y. Yan, H. Huang, J. H. Postlethwait, and W. S. Talbot. 2000. A comparative map of the zebrafish genome. *Genome Res.* **10**:1903–1914.
- Yu, W. P., S. Brenner, and B. Venkatesh. 2003. Duplication, degeneration and subfunctionalization of the nested synapsin-Timp genes in *Fugu*. *Trends Genet.* **19**:180–183.

Kenneth Wolfe, Associate Editor

Accepted February 25, 2004