Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Genetics & Development

Rapidly evolving fish genomes and teleost diversity

Vydianathan Ravi[a] and Byrappa Venkatesh[a,b]

Teleost fishes are the largest and most diverse group of vertebrates. The diversity of teleosts has been attributed to a whole-genome duplication (WGD) event in the ray-finned fish lineage. Recent comparative genomic studies have revealed that teleost genomes have experienced frequent gene-linkage disruptions compared to other vertebrates, and that protein-coding sequences in teleosts are evolving faster than in mammals, irrespective of their duplication status. A significant number of conserved noncoding elements (CNEs) shared between cartilaginous fishes and tetrapods have diverged beyond recognition in teleost fishes. The divergence of CNEs seems to have been initiated in basal ray-finned fishes before the WGD. The fast evolving singleton and duplicated genes as well as the divergent CNEs might have contributed to the diversity of teleost fishes.

Address
[a] Institute of Molecular and Cell Biology, A*STAR (Agency for Science, Technology and Research), Biopolis, Singapore 138673, Singapore
[b] Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119074, Singapore

Corresponding author: Venkatesh, Byrappa (mcbbv@imcb.a-star.edu.sg)

Current Opinion in Genetics & Development 2008, **18**:544–550

This review comes from a themed issue on
Genomes and evolution
Edited by Sarah Teichmann and Nipam Patel

Available online 16th December 2008

0959-437X/$ – see front matter

DOI 10.1016/j.gde.2008.11.001

Introduction

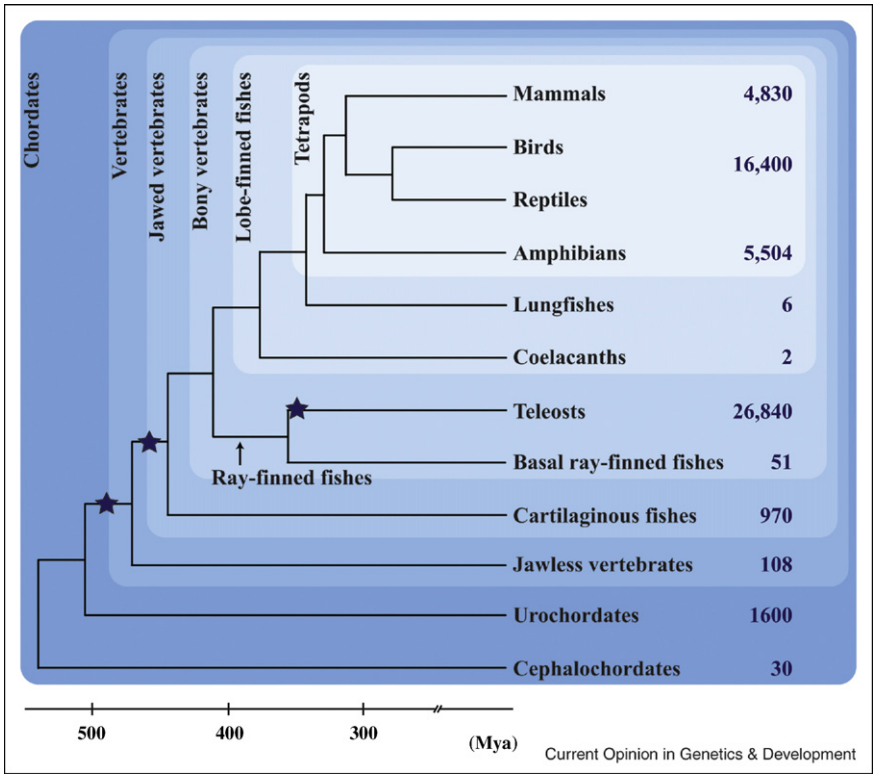
Teleost fishes, with about 27 000 species [1], are the largest and most diverse group of vertebrates. They exhibit remarkable diversity in their morphology, behavior, and adaptations. Teleosts account for more than 99% of ray-finned fishes (Actinopterygians) which diverged from lobe-finned fishes (Sarcopterygians) about 420 million years ago (Mya). The remaining ray-finned fishes, which are basal to teleosts, are represented by only ~50 living species (Figure 1). The sequencing of genes and gene families from teleost fishes in the pregenomic era had unexpectedly revealed the presence of duplicate teleost genes for several human genes. This led to the hypothesis that a whole-genome duplication (WGD) occurred in the ray-finned fish lineage before the diversification of teleost fishes [2–4]. More recently, the sequencing and comparative analysis of whole-genome sequences of teleost fishes such as fugu, *Tetraodon*, and medaka [5,6,7••] have provided compelling evidence for the WGD event in the fish lineage. Timing the gene duplication events using molecular clock and phylogenetic analyses indicated that the duplication occurred around 350 Mya, before the diversification of teleosts [8–11]. Since the basal ray-finned fishes are relatively species-poor, and the WGD predated the teleost radiation, the diversity and complexity of teleost fishes is often attributed to the WGD [10–14]. However, it has been pointed out that when the number and complexities of the extinct ray-finned fish taxa are also considered, there is no correlation between the fish-specific genome duplication and the diversity of teleost fishes [15]. Nonetheless, since major WGD events during vertebrate evolution are preceded by a series of extinction events, it is argued that the WGD could have reduced the probability of extinction through high evolvability, leading ultimately to the increased diversification of the surviving taxa [16]. Besides verifying the fish-specific WGD, recent studies have shown that teleost fishes have experienced a higher rate of chromosomal rearrangements and a faster evolution of protein sequences and conserved noncoding elements (CNEs) compared to cartilaginous fishes and mammals. Here, we review these data and discuss their implications for the evolution and diversity of teleost fishes.

Higher rate of chromosomal rearrangements in fishes

A WGD is generally followed by an increased rate of interchromosomal rearrangements [17]. A comparison of the recently completed medaka genome with the zebrafish, *Tetraodon*, and human genomes has revealed that eight major interchromosomal rearrangements occurred within a relatively short period of ~50 million years after the WGD in the fish lineage [7••]. Subsequently, while the medaka lineage experienced no major interchromosomal rearrangements, three major rearrangements occurred in the *Tetraodon* lineage. In contrast, the zebrafish lineage has experienced many interchromosomal rearrangements after it diverged from the medaka lineage [7••]. These results suggest that the WGD led to a transient increase in the rate of interchromosomal rearrangements and subsequently the teleost genomes are experiencing differential rates of chromosomal rearrangements.

An analysis of disruption of gene-linkages in vertebrates has suggested that teleost fishes have experienced a higher rate of chromosomal rearrangements compared to other vertebrates. Two studies that used different

Figure 1



Phylogenetic relationships of chordates. The number of extant species in the respective group (based on Ref. [1]) is given on the right. Stars represent the three major WGD events. Two of them occurred at the base of vertebrates, with the first occurring before the divergence of cyclostomes (jawless vertebrates) and the second occurring most likely after the divergence of cyclostomes [35]. The third WGD occurred in the teleost ancestor. Basal ray-finned fishes include the orders Polypteriformes (bichirs), Acipenseriformes (sturgeons and paddlefish), Semionotiformes (gar), and Amiiformes (bowfin). The scale bar represents fossil-based minimum divergence times: jawless vertebrates (Agnatha) and jawed vertebrates (Gnathostomes), 477 Mya [45]; cartilaginous fishes (Chondrichthyes) and bony vertebrates (Osteichthyes), 450 Mya [46]; ray-finned fishes (Actinopterygii) and lobe-finned fishes (Sarcopterygii), 416 Mya [47].

approaches to estimate the rate of gene-linkage disruption have found that the frequency of synteny loss in the terminal branches of teleosts is considerably higher than in the tetrapod lineages [18•,19••]. Recently, a 1.4× coverage sequence representing approximately 75% of the genome of a cartilaginous fish, the elephant shark (*Callorhinchus milii*), was generated based mainly on paired-end sequences of fosmid clones [20••]. A comparison of syntenic genes in the elephant shark (genes that map to the paired ends of a fosmid clone) with their orthologs in human and zebrafish genomes showed that the level of conserved synteny between the elephant shark and zebrafish is less than half of that between the elephant shark and human, and that the conserved clusters of syntenic genes in zebrafish contain fewer genes than in elephant shark and human [20••]. Although the loss of conserved synteny in teleosts can be because of the reciprocal loss of paralogous genes in duplicate chromosomes, comparisons of gene maps have indicated a high frequency of intrachromosomal rearrangements since the divergence of teleost and tetrapod ancestors [21,22]. Moreover, a comparison of gene order across large regions provides support for a higher rate of chromosomal rearrangements in teleost fishes compared to other vertebrates [23,24•]. For example, in contrast to the tight linkage between the human genes for the neurohypophysial hormones, oxytocin and vasopressin, and their orthologs in other tetrapods and coelacanth, the fugu and *Tetraodon* orthologs are separated by several unrelated genes and flanked by nonsyntenic genes [23]. Likewise, most of the genes flanking the duplicated protocadherin clusters in fugu are unrelated to the conserved syntenic genes flanking the unduplicated protocadherin cluster in the elephant shark and human [24•].

Protein sequences are evolving faster in fishes

The most common fate of duplicated genes is that while one of the daughter genes continues to be under selective pressure and retains the ancestral function, the other gene diverges and becomes nonfunctional through the accumulation of deleterious mutations. Less frequently,

both daughter genes are retained owing to neofunctionalization of one of the daughter genes (through accumulation of favorable mutations), subfunctionalization of the daughter genes (whereby both accumulate complementary mutations such that the functions of the ancestral gene are now partitioned between the two genes) or a combination of both (subneofunctionalization) [25].

Following WGD in the teleost ancestor, 15–25% of duplicated genes have been retained in pairs [6,22,26••]. Analyses of the evolutionary rates of protein-coding sequences have indeed revealed an asymmetric evolutionary rate between the duplicated gene pairs [10,26••,27,28•,29], which is suggestive of neofunctionalization and/or subfunctionalization. However, what is striking is that even the genes retained in single copy are evolving faster compared to their orthologs in mammals. This feature of teleost genes was first revealed when the substitution rates of a set of 19 genes in teleost fishes and mammals were analyzed. Most of the genes were found to be evolving at a higher rate in teleost fishes than in mammals, irrespective of whether they were duplicated or singletons [30]. Subsequently, pairwise comparisons of a genome-wide set of 5800 orthologous genes in *Tetraodon*, fugu, human, and mouse indicated that the neutral mutation rate between *Tetraodon* and fugu is higher than in human and mouse [6]. It was also observed that the protein sequences between the two pufferfishes are more divergent than between the two mammals even though the two pufferfishes diverged more recently (~32 Mya) than human and mouse (~61 Mya) [6]. Furthermore, comparisons with their orthologs in the urochordate *Ciona intestinalis* showed that the average frequency of nonsynonymous mutations between *Ciona* and *Tetraodon* is higher than between *Ciona* and human [6], providing further support that protein sequences in teleosts are evolving faster than in mammals. A recent investigation of substitution rates in a strict set of singleton and duplicate genes in *Tetraodon* and fugu has provided clear evidence that the singleton genes and at least one copy of the duplicate genes in pufferfishes are evolving significantly faster than their human and mouse orthologs [26••]. Although the asymmetric evolution of the duplicated genes could be attributed to the relaxed selective pressure, it is unclear whether the accelerated substitution rate of singleton genes is also related to WGD. Comparative analysis of substitution rates of singleton and duplicated genes in teleosts and their orthologs in basal ray-finned fishes that escaped WGD and mammals should clarify this.

CNEs have diverged faster in fishes

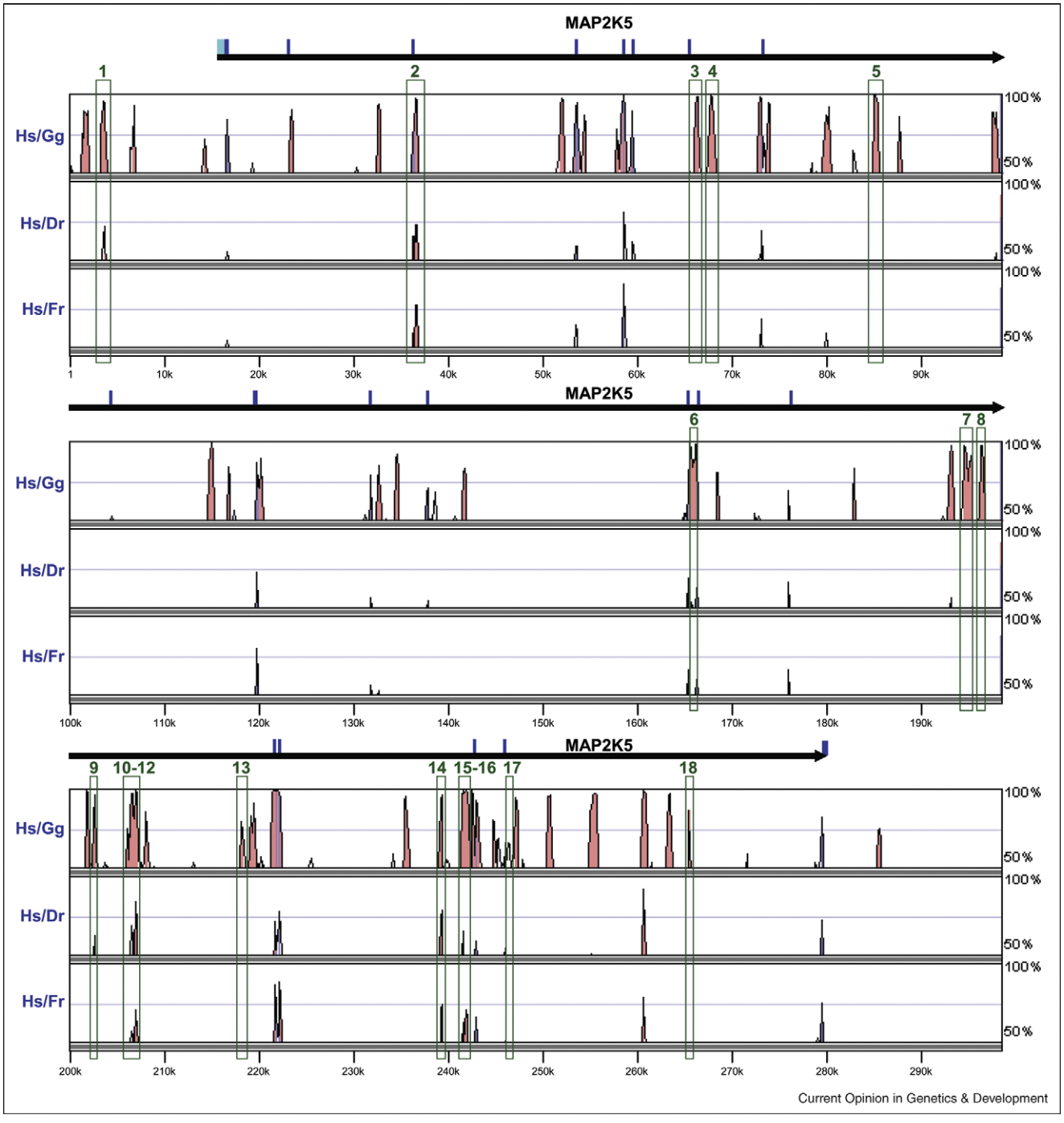
Extreme evolutionary conservation is a useful signature for identifying putative *cis*-regulatory elements in the human genome. Indeed, *in vivo* assay of noncoding elements conserved between human and distant vertebrates has indicated that many of them function as *cis*-regulatory elements directing tissue-specific expression of target genes [31–33]. Attempts to trace the evolutionary history of such elements in the human genome have indicated that they were recruited at different stages during the vertebrate evolution, with substantial numbers recruited in the bony vertebrate ancestor and the tetrapod ancestor [34••]. In fact, some conserved *cis*-regulatory elements can be traced even to cephalochordates (amphioxus [35••,36]), which are the most basal group of chordates.

Although several thousand CNEs have been identified between human and teleost fishes [31–33,34••], comparisons with more ancient vertebrates such as cartilaginous fishes have revealed that a significant number of ancient CNEs have been lost or diverged beyond recognition in teleost fishes. A large number of CNEs shared between the single HoxA cluster in human and horn shark (*Heterodontus francisci*) could not be identified in the duplicated HoxA clusters in zebrafish [37]. Interestingly, the pattern of loss is inconsistent with the duplication–degeneration–complementation model [38] which predicts that the duplicated genes retain complementary subsets of the ancestral *cis*-regulatory elements. Thus, the loss of these ancient CNEs in the zebrafish could be the result of adaptive modifications [37].

Genome-wide comparisons of CNEs in the elephant shark, human, fugu, and zebrafish have indicated that the loss of CNEs in teleosts is a genome-wide phenomenon. As compared to 4782 CNEs (≥70% identity across >100 bp) identified between the human genome and the 1.4× coverage assembly of the elephant shark genome, only 2107 and 2838 CNEs were identified between fugu–human and zebrafish–human genomes, respectively [39••]. Indeed, the divergence of ancient CNEs has been experienced by both singleton (an example shown in Figure 2) and duplicated fugu and zebrafish genes. The more recently completed genomes of medaka and stickleback contain a similar number of ancient CNEs to that in fugu and zebrafish (AP Lee, B Venkatesh, unpublished). Surprisingly, although almost all the 4782 elephant shark and human sequences are conserved in the chicken and dog genomes, only ~35% are recognizable in fugu and zebrafish. This indicates that a significant number of ancient CNEs (~3100) that are shared between cartilaginous fishes and tetrapods have been lost or diverged beyond recognition in teleost fishes.

The absence of a large number of elephant shark–human CNEs in diverse teleosts such as fugu, zebrafish, medaka, and stickleback indicates that the divergence of the CNEs began before the radiation of teleosts. The divergence might have, in fact, started in the stem ray-finned fish lineage as indicated by the pattern of CNEs in the HoxA cluster of bichir, which belongs to the most basal lineage of extant ray-finned fishes [40,41]. Although the

Figure 2



Teleost fishes have lost many ancient CNEs that are conserved in the elephant shark and tetrapods. A VISTA plot of the MLAGAN [48] alignment of human (Hs) *MAP2K5* gene with chicken (Gg), zebrafish (Dr), and fugu (Fr) orthologs showing the CNEs conserved in various vertebrates. The human sequence was used as the base sequence (*x*-axis) for the alignment. The black arrow denotes the *MAP2K5* gene with exons shown as blue boxes. It is a single copy gene in zebrafish, fugu, and other vertebrates. *y-Axis* represents percent sequence identity. CNEs were predicted using the criterion of $\geq$70% identity across $>$100 bp windows. Violet peaks represent conserved exonic sequences and pink peaks represent CNEs. Ancient CNEs conserved between the elephant shark and human (numbered 1–18) are highlighted with green boxes. The IDs of these CNEs are: 1, EH12903; 2, EH12904; 3, EH12906; 4, EH12908; 5, EH12910; 6, EH12912; 7, EH12913; 8, EH12914; 9, EH12916; 10, EH12917; 11, EH12918; 12, EH12919; 13, EH12920; 14, EH12924; 15, EH12925; 16, EH12926; 17, EH12928; 18, EH12930 [39]. All 18 CNEs are conserved in the dog genome (data not shown) while the chicken (Gg) genome has 17 CNEs (CNE 17 is missing). On the other hand, only three of them (CNEs 2, 12, and 14) are conserved in both zebrafish and fugu. One (CNE 1) is conserved in zebrafish but lost (deleted or diverged beyond recognition as a CNE) in fugu, and another one (CNE 16) is conserved in fugu but lost in zebrafish. The remaining 13 (CNEs 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 17, and 18) have been lost in both fugu and zebrafish.

unduplicated HoxA cluster in bichir has lost some of the CNEs shared by the horn shark and human, it has retained some that have been lost in the duplicated HoxA clusters of zebrafish and fugu. In addition, bichir has acquired some new CNEs that are conserved in teleost fishes [42]. Therefore, it appears that the divergence of ancient CNEs began in the stem ray-finned fish lineage and continued during the evolution of teleost fishes.

Conclusions

The teleost fish genomes have experienced a higher rate of gene-linkage disruption and chromosomal rearrangements compared to mammals. The protein-coding sequences in teleost fish genomes are evolving faster than in mammals irrespective of the duplication status of the genes. However, owing to a paucity of protein-coding sequence data from basal ray-finned fishes, it is unclear whether the higher evolutionary rate is unique to teleost fishes or common to all ray-finned fishes. A significant number of ancient CNEs (putative *cis*-regulatory elements) that are shared between cartilaginous fishes and tetrapods have evolved fast and diverged beyond recognition in teleost fishes. The limited data available from basal ray-finned fishes suggest that the divergence of ancient CNEs began in the stem ray-finned fish lineage well before the WGD. Sequencing and comparative analysis of whole-genome sequences from basal ray-finned fishes such as bichir and bowfin that diverged before the WGD should help to verify this hypothesis. The rapid changes in the genome organization, protein-coding sequences, and CNEs in teleost fishes all indicate that their genomes are evolving faster compared to those of mammals and the basal cartilaginous fishes.

The fast evolving protein sequences and the divergent CNEs in teleost fish genomes may have important implications for the morphological diversity and rapid speciation of teleost fishes. The higher neutral mutation rate in teleost protein-coding sequences offers higher chances for selection to act and retain favorable mutations, thereby accelerating the process of neofunctionalization of duplicated and singleton genes. Given that nearly 80% of genes are retained as singletons in teleosts after the WGD, the fast evolving singleton genes might have, in fact, played an important role in the acquisition of evolutionary novelties. Recent studies have indeed shown that positive selection has acted equally on both duplicated and unduplicated genes in teleosts and other vertebrates [43••]. Thus, many of the divergent protein sequences in teleosts may be associated with adaptive evolution. Although the significance of CNEs in vertebrate genomes has not been fully understood, it is hypothesized that most of them, particularly those conserved between distant vertebrates, are likely to be functional elements that are under purifying selection. The functions of only a couple of hundred CNEs have been tested so far, and the results indicate that many of them function as *cis*-regulatory elements directing tissue-specific or developmental stage-specific expression of target genes [31,33]. The divergence or loss of such ancient CNEs in teleosts can potentially alter the regulation of genes resulting in loss or gain of function in a specific cell-type or developmental stage. Such changes in the regulation of genes can result in genetic isolation between populations, leading ultimately to speciation. The altered regulation of genes also has the potential to give rise to marked differences in the morphology, physiology, and behavior between species [44]. Thus, the fast evolving protein-coding sequences and the divergent CNEs might have together played a key role in the diversification and evolutionary success of teleost fishes.

Acknowledgements

Research in the authors' laboratory is supported by the Biomedical Research Council of the A*STAR, Singapore. We would like to thank Eddie Loh, Alison Lee and Patrick Gilligan for their helpful suggestions.

References and recommended reading

Papers of special interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Nelson JS: *Fishes of the World*. edn 4. New York: John Wiley & Sons; 2006.

2. Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang YL *et al.*: **Zebrafish hox clusters and vertebrate genome evolution**. *Science* 1998, **282**:1711-1714.

3. Postlethwait JH, Woods IG, Ngo-Hazelett P, Yan YL, Kelly PD, Chu F, Huang H, Hill-Force A, Talbot WS: **Zebrafish comparative genomics and the origins of vertebrate chromosomes**. *Genome Res* 2000, **10**:1890-1902.

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

5. Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, Dehal P, Christoffels A, Rash S, Hoon S, Smit A *et al.*: **Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes***. *Science* 2002, **297**:1301-1310.

6. Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, Mauceli E, Bouneau L, Fischer C, Ozouf-Costaz C, Bernot A *et al.*: **Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype**. *Nature* 2004, **431**:946-957.

7. •• Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W, Ahsan B, Yamada T, Nagayasu Y, Doi K, Kasai Y *et al.*: **The medaka draft genome and insights into vertebrate genome evolution**. *Nature* 2007, **447**:714-719.
On the basis of comparisons of gene maps of medaka, *Tetraodon*, zebrafish, and human, the karyotype of the ancestral teleost was constructed and timings of major interchromosomal rearrangements during the evolution of teleosts were estimated.

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

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

10. Crow KD, Stadler PF, Lynch VJ, Amemiya C, Wagner GP: **The "fish-specific" Hox cluster duplication is coincident with the origin of teleosts**. *Mol Biol Evol* 2006, **23**:121-136.

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

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

13. Semon M, Wolfe KH: **Reciprocal gene loss between Tetraodon and zebrafish after whole genome duplication in their ancestor**. *Trends Genet* 2007, **23**:108-112.

14. Taylor JS, Van de Peer Y, Meyer A: **Genome duplication, divergent resolution and speciation**. *Trends Genet* 2001, **17**:299-301.

15. Donoghue PC, Purnell MA: **Genome duplication, extinction and vertebrate evolution**. *Trends Ecol Evol* 2005, **20**:312-319.

16. Crow KD, Wagner GP:: **Proceedings of the SMBE Tri-National Young Investigators' Workshop 2005. What is the role of genome duplication in the evolution of complexity and diversity?** *Mol Biol Evol* 2006, **23**:887-892.

17. Otto SP: **The evolutionary consequences of polyploidy**. *Cell* 2007, **131**:452-462.

18. • Semon M, Wolfe KH: **Rearrangement rate following the whole-genome duplication in teleosts**. *Mol Biol Evol* 2007, **24**:860-867.
This study has shown that the terminal teleost branches have experienced a higher rate of disruption of adjacent gene-linkages compared to human, mouse, and chicken lineages.

19. •• Hufton AL, Groth D, Vingron M, Lehrach H, Poustka AJ, Panopoulou G: **Early vertebrate whole genome duplications were predated by a period of intense genome rearrangement**. *Genome Res* 2008, **18**:1582-1591.
The authors use the amphioxus genome as an outgroup to estimate the rate of chromosomal rearrangements during vertebrate evolution. They note that the rate of rearrangement is in fact higher before WGDs rather than after the WGDs. Their analysis also shows that the terminal branches of teleost fishes have a higher rearrangement rate than tetrapod lineages.

20. •• Venkatesh B, Kirkness EF, Loh YH, Halpern AL, Lee AP, Johnson J, Dandona N, Viswanathan LD, Tay A, Venter JC *et al.*: **Survey sequencing and comparative analysis of the elephant shark (*Callorhinchus milii*) genome**. *PLoS Biol* 2007, **5**:e101.
First paper to provide an overview of a cartilaginous fish (the elephant shark) genome.

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

22. Woods IG, Wilson C, Friedlander B, Chang P, Reyes DK, Nix R, Kelly PD, Chu F, Postlethwait JH, Talbot WS: **The zebrafish gene map defines ancestral vertebrate chromosomes**. *Genome Res* 2005, **15**:1307-1314.

23. Gwee PC, Amemiya CT, Brenner S, Venkatesh B: **Sequence and organization of coelacanth neurohypophysial hormone genes: evolutionary history of the vertebrate neurohypophysial hormone gene locus**. *BMC Evol Biol* 2008, **8**:93.

24. • Yu WP, Rajasegaran V, Yew K, Loh WL, Tay BH, Amemiya CT, Brenner S, Venkatesh B: **Elephant shark sequence reveals unique insights into the evolutionary history of vertebrate genes: a comparative analysis of the protocadherin cluster**. *Proc Natl Acad Sci U S A* 2008, **105**:3819-3824.

25. He X, Zhang J: **Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution**. *Genetics* 2005, **169**:1157-1164.

26. •• Brunet FG, Crollius HR, Paris M, Aury JM, Gibert P, Jaillon O, Laudet V, Robinson-Rechavi M: **Gene loss and evolutionary rates following whole-genome duplication in teleost fishes**. *Mol Biol Evol* 2006, **23**:1808-1816.
Using a strict set of singleton and duplicated genes from *Tetraodon* and fugu, the authors provide evidence that singleton genes in pufferfishes are evolving faster than in human and mouse, and that one copy of the duplicated gene in pufferfishes is evolving faster compared to its ortholog in human and mouse.

27. Braasch I, Salzburger W, Meyer A: **Asymmetric evolution in two fish-specifically duplicated receptor tyrosine kinase paralogons involved in teleost coloration**. *Mol Biol Evol* 2006, **23**:1192-1202.

28. • Steinke D, Salzburger W, Braasch I, Meyer A: **Many genes in fish have species-specific asymmetric rates of molecular evolution**. *BMC Genomics* 2006, **7**:20.
This paper provides evidence for asymmetric rate of evolution of duplicated genes in fugu, *Tetraodon*, zebrafish, and medaka.

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

30. Robinson-Rechavi M, Laudet V: **Evolutionary rates of duplicate genes in fish and mammals**. *Mol Biol Evol* 2001, **18**:681-683.

31. Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD *et al.*: **In vivo enhancer analysis of human conserved non-coding sequences**. *Nature* 2006, **444**:499-502.

32. Shin JT, Priest JR, Ovcharenko I, Ronco A, Moore RK, Burns CG, MacRae CA: **Human–zebrafish non-coding conserved elements act in vivo to regulate transcription**. *Nucleic Acids Res* 2005, **33**:5437-5445.

33. Woolfe A, Goodson M, Goode DK, Snell P, McEwen GK, Vavouri T, Smith SF, North P, Callaway H, Kelly K *et al.*: **Highly conserved non-coding sequences are associated with vertebrate development**. *PLoS Biol* 2005, **3**:e7.

34. •• Stephen S, Pheasant M, Makunin IV, Mattick JS: **Large-scale appearance of ultraconserved elements in tetrapod genomes and slowdown of the molecular clock**. *Mol Biol Evol* 2008, **25**:402-408.
Identified about 13 700 'ultraconserved' elements that are identical across at least 100 bp in three placental mammals. The search for the ultraconserved elements in other tetrapods and teleost fishes showed that 39% were present in the bony vertebrate ancestor, 30% were recruited in the tetrapod ancestor and another 19% were recruited in the amniote ancestor.

35. •• Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK *et al.*: **The amphioxus genome and the evolution of the chordate karyotype**. *Nature* 2008, **453**:1064-1071.
Comparative analysis of the amphioxus genome has provided strong evidence for the two rounds of genome duplications during the origin of vertebrates. Several noncoding elements were found to be conserved between human and amphioxus genomes. Functions of some of these elements were tested and reported in an accompanying paper [36].

36. Holland LZ, Albalat R, Azumi K, Benito-Gutierrez E, Blow MJ, Bronner-Fraser M, Brunet F, Butts T, Candiani S, Dishaw LJ *et al.*: **The amphioxus genome illuminates vertebrate origins and cephalochordate biology**. *Genome Res* 2008, **18**:1100-1111.

37. Chiu CH, Amemiya C, Dewar K, Kim CB, Ruddle FH, Wagner GP: **Molecular evolution of the HoxA cluster in the three major gnathostome lineages**. *Proc Natl Acad Sci U S A* 2002, **99**:5492-5497.

38. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J: **Preservation of duplicate genes by complementary, degenerative mutations**. *Genetics* 1999, **151**:1531-1545.

39. •• Venkatesh B, Kirkness EF, Loh YH, Halpern AL, Lee AP, Johnson J, Dandona N, Viswanathan LD, Tay A, Venter JC *et al.*: **Ancient noncoding elements conserved in the human genome**. *Science* 2006, **314**:1892.
About 4800 conserved noncoding sequences ($\geq$70% identity across $>$100 bp) were identified between the human genome and the genome of a cartilaginous fish, the elephant shark. Only a handful of them have orthologs in invertebrates indicating that the majority of them originated in the vertebrate ancestor. Surprisingly, about 65% of them have been lost or diverged beyond recognition in teleost fishes.

40. Venkatesh B, Erdmann MV, Brenner S: **Molecular synapomorphies resolve evolutionary relationships of extant jawed vertebrates**. *Proc Natl Acad Sci U S A* 2001, **98**:11382-11387.

41. Inoue JG, Miya M, Tsukamoto K, Nishida M: **Basal actinopterygian relationships: a mitogenomic perspective on the phylogeny of the ''ancient fish''**. *Mol Phylogenet Evol* 2003, **26**:110-120.

42. Chiu CH, Dewar K, Wagner GP, Takahashi K, Ruddle F, Ledje C, Bartsch P, Scemama JL, Stellwag E, Fried C *et al.*: **Bichir HoxA cluster sequence reveals surprising trends in ray-finned fish genomic evolution**. *Genome Res* 2004, **14**:11-17.

43. •• Studer R, Duret L, Penel S, Robinson-Rechavi M: **Pervasive positive selection on duplicated and non duplicated vertebrate protein coding genes**. *Genome Res* 2008, **18**:1393-1402.
This is a stringent analysis of the impact of positive selection during vertebrate evolution. The study shows for the first time that positive selection is widespread in vertebrate evolution, and that there is no change in the prevalence of positive selection after the WGD during the origin of vertebrates and in the teleost ancestor.

44. Wray GA: **The evolutionary significance of *cis*-regulatory mutations**. *Nat Rev Genet* 2007, **8**:206-216.

45. Janvier P: **Palaeontology: modern look for ancient lamprey**. *Nature* 2006, **443**:921-924.

46. Sansom IJ, Smith MM, Smith MP: **Scales of thelodont and shark-like fishes from the Ordovician of Colorado**. *Nature* 1996, **379**:628-630.

47. Benton MJ, Donoghue PC: **Paleontological evidence to date the tree of life**. *Mol Biol Evol* 2007, **24**:26-53.

48. Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, Green ED, Sidow A, Batzoglou S: **LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA**. *Genome Res* 2003, **13**:721-731.