ANNUAL REVIEWS

Annual Review of Animal Biosciences

The Divergent Genomes of Teleosts

Vydianathan Ravi and Byrappa Venkatesh

Institute of Molecular and Cell Biology, A*STAR (Agency for Science, Technology and Research), Biopolis, Singapore 138673; email: raviv@imcb.a-star.edu.sg, mcbbv@imcb.a-star.edu.sg

Annu. Rev. Anim. Biosci. 2018. 6:47–68

The *Annual Review of Animal Biosciences* is online at animal.annualreviews.org

https://doi.org/10.1146/annurev-animal-030117-014821

Keywords

whole-genome duplication, evolutionary rate, intron turnover, conserved noncoding elements, conserved syntenic blocks, phenotypic diversity

Abstract

Boasting nearly 30,000 species, teleosts account for half of all extant vertebrates and approximately 98% of all ray-finned fish species (Actinopterygii). Teleosts are also the largest and most diverse group of vertebrates, exhibiting an astonishing level of morphological, physiological, and behavioral diversity. Previous studies had indicated that the teleost lineage has experienced an additional whole-genome duplication event. Recent comparative genomic analyses of teleosts and other bony vertebrates using spotted gar (a nonteleost ray-finned fish) and elephant shark (a cartilaginous fish) as outgroups have revealed several divergent features of teleost genomes. These include an accelerated evolutionary rate of protein-coding and nucleotide sequences, a higher rate of intron turnover, loss of many potential *cis*-regulatory elements and shorter conserved syntenic blocks. A combination of these divergent genomic features might have contributed to the evolution of the amazing phenotypic diversity and morphological innovations of teleosts.

INTRODUCTION

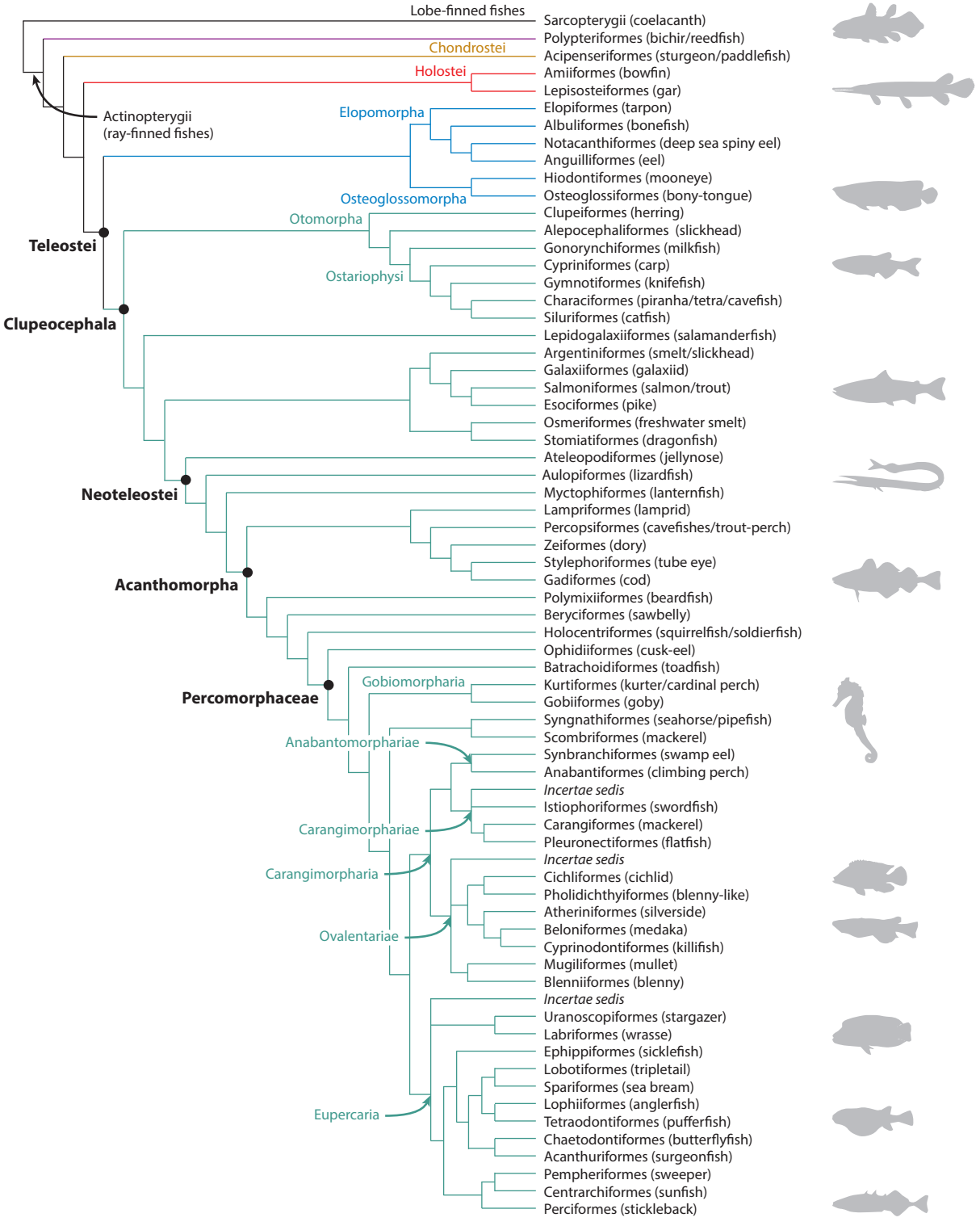
With over 68,000 known species (1; **http://www.iucnredlist.org**), vertebrates are the most dominant and successful group of animals on earth, inhabiting both terrestrial and aquatic habitats. Vertebrates are divided into two broad groups: the jawless vertebrates (Agnatha) and jawed vertebrates (Gnathostomes), with the former group represented by only two extant lineages, lampreys and hagfishes (together known as cyclostomes). Gnathostomes, conversely, comprise the vast majority, accounting for roughly 99% of all vertebrates. There are two main groups within Gnathostomes: cartilaginous fishes (Chondrichthyes), which include sharks, rays, skates, and chimaeras, and bony vertebrates (Osteichthyes), which include ray-finned fishes, lobe-finned fishes, and tetrapods (land animals). Ray-finned fishes (Actinopterygii) are the most dominant group within the vertebrates, and with approximately 30,500 species, they comprise roughly half of all living vertebrates. Two major groups can be recognized in ray-finned fishes: teleosts and basal, nonteleost ray-finned fishes, which comprise four orders—Polypteriformes (e.g., bichir), Acipenseriformes (e.g., sturgeon and paddlefish), Lepisosteiformes (e.g., spotted gar), and Amiiformes (e.g., bowfin).

Teleosts are a monophyletic group comprising approximately 70 orders and approximately 500 families (**Figure 1**) whose phylogenetic relationships are not yet fully resolved (2). At least 27 distinct morphological synapomorphies support the monophyly of teleosts (3). These include a diural caudal skeleton, presence of uroneurals, unpaired basibranchial toothplates, mobile premaxilla, and lower jaws without coronoid and supraangular bones (3, 4). Teleost monophyly is also supported by molecular evidence (2, 5). With nearly 30,000 species, teleosts account for approximately 98% of all ray-finned fishes (6) and nearly half of all extant vertebrates. Teleosts exhibit an astonishing level of diversity in terms of their morphology, habitat, physiology, and behavior. Additionally, they are well known for the phenomenon of adaptive radiation, whereby a single lineage can differentiate into several diverse species over a short evolutionary time span. The African cichlids are a classical example of this phenomenon (7, 8). In three East African lakes (Victoria, Malawi, and Tanganyika), adaptive radiation has generated 250–500 species of cichlids per lake over evolutionary time periods ranging from 15,000 to 12 million years, thus displaying the highest speciation rates among vertebrates (7, 8). The wealth of information in terms of phenotypic diversity in extant as well as fossil taxa and the adaptive radiation make teleosts fascinating subjects that offer unique opportunities to understand the genetic basis of adaptation, speciation, and morphological and physiological diversity.

The basic body plan and developmental programs of teleosts are similar to those of mammals. In addition, teleosts and mammals also possess a similar gene repertoire. Teleosts are therefore well suited for comparative studies of humans and have emerged as excellent genetic models for understanding the molecular mechanisms underlying human development and diseases (9, 10). Small-sized teleosts like zebrafish and medaka have become genetic models of choice because of their size, external development, transparent embryos, ease of breeding and maintenance under laboratory conditions, low maintenance cost, high fecundity, short generation times, and established laboratory procedures for fish manipulation (9, 10).

Figure 1

Phylogenetic relationships of ray-finned fishes. Names of orders, subgroups, and their phylogenetic relationships are mainly based on the study by Betancur et al. (2) (DeepFin.org classification v.3: **https://sites.google.com/site/guilleorti/home/classification**). The sister group relation between Osteoglossomorpha and Elopomorpha is based on Bian et al. (105). Relationships within Eupercaria, Ovalentariae, and Carangimorphariae shown in the figure are poorly supported. Common name(s) of a representative fish is shown within parentheses. Silhouettes of some representative teleosts are shown on the right of the tree.



Genomes of more than 60 teleosts (**Table 1**) have been sequenced to date and have been used for genetic and evolutionary studies via comparative analysis with human and other vertebrate genomes. One of the striking features of the teleost genomes that was revealed by comparison of gene families and whole teleost genomes is that the teleost lineage has experienced an additional whole-genome duplication (WGD), known as teleost-specific genome duplication (TGD) (11–14), on top of the two rounds of WGD events that occurred during the evolution of early vertebrates. The genome of the Japanese pufferfish (*Takifugu rubripes*) was the first teleost genome and the second vertebrate genome sequenced, soon after the completion of the human genome (15). Pufferfish was selected as a model vertebrate genome because of its compact genome, which is only one-eighth the size of the human genome. Comparison of the human and pufferfish genomes confirmed the earlier prediction (16) that they share a similar repertoire of genes. Subsequently, the genomes of genetic models such as medaka, stickleback, and zebrafish were sequenced (17–19). Initially, most of the comparative genomic analyses were carried out between teleost and human genomes without an outgroup vertebrate genome as a reference. Such comparative studies were effective in identifying genetic features conserved between teleosts and human but were unable to infer with confidence whether the unique features observed in one lineage were the result of gain in that lineage or loss in the other. More recently, the genome of a cartilaginous fish, the elephant shark (*Callorhinchus milii*), was generated (20). Cartilaginous fishes are the sister group of bony vertebrates (**Figure 2**) and thus serve as a valuable outgroup for reconstructing the ancestral jawed vertebrate genome, as well as for inferring the shared and derived features of bony vertebrate lineages, such as humans and teleosts. In addition, the genome of a basal nonteleost ray-finned fish, the spotted gar (21), has been sequenced. The spotted gar lineage diverged from the teleost lineage prior to TGD (**Figure 2**) and thus possesses a typical diploid vertebrate genome similar to the human genome. Comparative analyses have shown that the gar genome is organized more like chicken and human genomes than like those of teleosts (21, 22). This makes the spotted gar a useful ray-finned fish outgroup to the teleosts, much closer than the elephant shark. Comparisons of human (and other tetrapods) and teleost genomes together with elephant shark and/or spotted gar genomes have been much more informative in identifying features conserved in human and teleost genomes and features that evolved in the common ancestor of bony vertebrates, and in providing new insights into the genome of the common ancestor of jawed vertebrates. These comparative studies have also been able to highlight several unique features of teleost genomes, indicating that they are highly divergent as compared with genomes of other jawed vertebrates. In this review, we discuss certain distinct features of teleost genomes and examine whether these features have contributed to the phenotypic diversity of teleosts. These distinct features of teleost genomes underscore the need to be cautious when using teleosts as genome or genetic models for comparative studies of human and other vertebrates.

WHOLE-GENOME DUPLICATIONS

The early evolution of vertebrates was accompanied by two rounds of WGD. These ancient WGDs may have spurred the evolution of complex vertebrates from relatively less complex invertebrates. Thereafter WGD is known only in ray-finned fish and amphibian lineages. Ray-finned fishes in particular seem to be more prone to WGD, as tetraploidization has occurred in several lineages of ray-finned fishes. There is now extensive evidence supporting the occurrence of TGD in the common ancestor of teleosts approximately 320 Mya (11–14). In addition to the TGD at the base of the teleosts, there have been recent genome duplication events within some teleost lineages. One of the affected lineages is the salmonids (family Salmonidae), including salmons, trouts, chars, graylings, and whitefishes, which experienced a WGD event, also known as the

Table 1 Ray-finned fish species with published whole-genome sequences

	Common name	Scientific name	Order	Contig N50 (kb)	Scaffold N50 (Mb)	BioProject accession
1	Fugu	*Takifugu rubripes*	Tetraodontiformes	52.9	0.93	PRJNA1434
2	Spotted green pufferfish	*Tetraodon nigroviridis*	Tetraodontiformes	16.0	0.98	PRJNA12350
3	Japanese medaka	*Oryzias latipes*	Beloniformes	9.8	1.41	PRJNA16702
4	Atlantic cod	*Gadus morhua*	Gadiformes	116	1.15	PRJNA299100
5	Stickleback	*Gasterosteus aculeatus*	Perciformes	83.2	10.8	PRJNA151055
6	Platyfish	*Xiphophorus maculatus*	Cyprinodontiformes	22.0	1.1	PRJNA72525
7	Zebrafish	*Danio rerio*	Cypriniformes	25.0	1.55	PRJNA11776
8	Pacific bluefin tuna	*Thunnus orientalis*	Scombriformes	7.6	0.14	PRJDA68701
9	Japanese eel	*Anguilla japonica*	Anguilliformes	NA	0.053	PRJNA158309
10	Tongue sole	*Cynoglossus semilaevis*	Pleuronectiformes	26.5	0.87	PRJNA73987
11	Rainbow trout	*Oncorhynchus mykiss*	Salmoniformes	7.7	0.38	PRJEB4421
12	Electric eel	*Electrophorus electricus*	Gymnotiformes	104.2	NA	PRJNA249073
13	Pufferfish	*Takifugu flavidus*	Tetraodontiformes	2.8	0.30	PRJNA168966
14	Northern pike	*Esox lucius*	Esociformes	16.9	0.7	PRJNA221548
15	Nile tilapia	*Oreochromis niloticus*	Cichliformes	29.3	2.8	PRJNA59571
16	Prinsessan av Burundi	*Neolamprologus brichardi*	Cichliformes	13.2	4.4	PRJNA60365
17	Mbuna	*Metriaclima zebra*	Cichliformes	20.0	3.7	PRJNA60369
18	Nyerere's cichlid	*Pundamilia nyererei*	Cichliformes	22.6	2.5	PRJNA60367
19	Burton's haplo	*Astatotilapia burtoni*	Cichliformes	21.9	1.2	PRJNA60363
20	Common carp	*Cyprinus carpio*	Cypriniformes	68.4	1.0	PRJNA202478
21	Antarctic bullhead notothen	*Notothenia coriiceps*	Perciformes	11.6	0.22	PRJNA66471
22	Mexican tetra (blind cavefish)	*Astyanax mexicanus*	Characiformes	14.7	1.77	PRJNA89115
23	Yellow croaker	*Larimichthys crocea*	Eupercaria	63.1	1.03	PRJNA245366
24	Blue-spotted mudskipper	*Boleophthalmus pectinirostris*	Gobiiformes	20.2	2.31	PRJNA232434
25	Giant-fin mudskipper	*Periophthalmus magnuspinnatus*	Gobiiformes	27.6	0.29	PRJNA232435
26	Blue mudskipper	*Scartelaos histophorus*	Gobiiformes	8.4	0.01	PRJNA232437
27	Giant mudskipper	*Periophthalmodon schlosseri*	Gobiiformes	16.9	0.04	PRJNA232436
28	European seabass	*Dicentrarchus labrax*	Eupercaria	53.0	5.1	PRJEB5099
29	Guppy	*Poecilia reticulata*	Cyprinodontiformes	35.6	5.27	PRJNA238429
30	Grass carp	*Ctenopharyngodon idellus*	Cypriniformes	40.8	6.4	PRJEB5920
31	Fathead minnow	*Pimephales promelas*	Cypriniformes	7.5	0.06	PRJNA227290

(*Continued*)

Table 1 (*Continued*)

	Common name	Scientific name	Order	Contig N50 (kb)	Scaffold N50 (Mb)	BioProject accession
32	African turquoise killifish	*Nothobranchius furzeri*	Cyprinodontiformes	132.5	0.5	PRJEB5837
33	Asian arowana	*Scleropages formosus*	Osteoglossiformes	30.7	5.96	PRJNA290065
34	Cavefish (cyprinid fish)	*Sinocyclocheilus grahami*	Cypriniformes	29.3	1.15	PRJNA274017
35	Cavefish (cyprinid fish)	*Sinocyclocheilus rhinocerous*	Cypriniformes	17.6	0.89	PRJNA274017
36	Cavefish (cyprinid fish)	*Sinocyclocheilus anshuiensis*	Cypriniformes	16.7	1.25	PRJNA274017
37	Monterrey platyfish	*Xiphophorus couchianus*	Cyprinodontiformes	60.0	1.8	PRJNA290781
38	Green swordtail	*Xiphophorus hellerii*	Cyprinodontiformes	30.0	1.6	PRJNA290782
39	Miiuy croaker	*Miichthys miiuy*	Eupercaria	73.3	1.15	PRJNA272995
40	Turbot	*Scophthalmus maximus*	Pleuronectiformes	31.2	4.3	PRJEB11743
41	Spotted gar	*Lepisosteus oculatus*	Lepisosteiformes	68.3	6.9	PRJNA68247
42	Asian seabass	*Lates calcarifer*	Carangimorphariae in sed	1066.0	1.2	PRJNA294489
43	Atlantic herring	*Clupea harengus*	Clupeiformes	21.3	1.84	PRJNA265919
44	Channel catfish	*Ictalurus punctatus*	Siluriformes	77.2	7.73	PRJNA281269
45	Atlantic salmon	*Salmo salar*	Salmoniformes	57.6	2.97	PRJNA72713
46	Mangrove rivulus	*Kryptolebias marmoratus*	Cyprinodontiformes	16.1	0.11	PRJNA290522
47	Ocean sunfish	*Mola mola*	Tetraodontiformes	20.0	9.0	PRJNA305960
48	Amur ide	*Leuciscus waleckii*	Cypriniformes	37.3	0.45	PRJEB12292
49	Marine medaka	*Oryzias melastigma*	Beloniformes	NA	2.57	PRJNA301650
50	Tiger tail seahorse	*Hippocampus comes*	Syngnathiformes	34.7	1.8	PRJNA314292
51	Gulf pipefish	*Syngnathus scovelli*	Syngnathiformes	32.2	0.64	PRJNA355893
52	Japanese flounder	*Paralichthys olivaceus*	Pleuronectiformes	30.5	3.9	PRJNA73673
53	Northern snakehead	*Channa argus*	Anabantiformes	81.4	4.5	PRJNA316547
54	Clearhead icefish	*Protosalanx hyalocranius*	Osmeriformes	17.2	1.16	PRJNA328051
55	Lined seahorse	*Hippocampus erectus*	Syngnathiformes	14.57	1.97	PRJNA347499
56	Blunt snout bream	*Megalobrama amblycephala*	Cypriniformes	49.4	0.84	PRJNA343584
57	European grayling	*Thymallus thymallus*	Salmoniformes	11.2	0.28	PRJEB21333
58	Whale shark	*Rhincodon typus*	Orectolobiformes	5.3	0.005	PRJNA255419
59	Murray cod	*Maccullochella peelii*	Perciformes	52.7	0.11	PRJNA290988
60	Antarctic dragonfish	*Parachaenichthys charcoti*	Perciformes	6.1	0.18	PRJNA330735

Abbreviation: NA, not available.

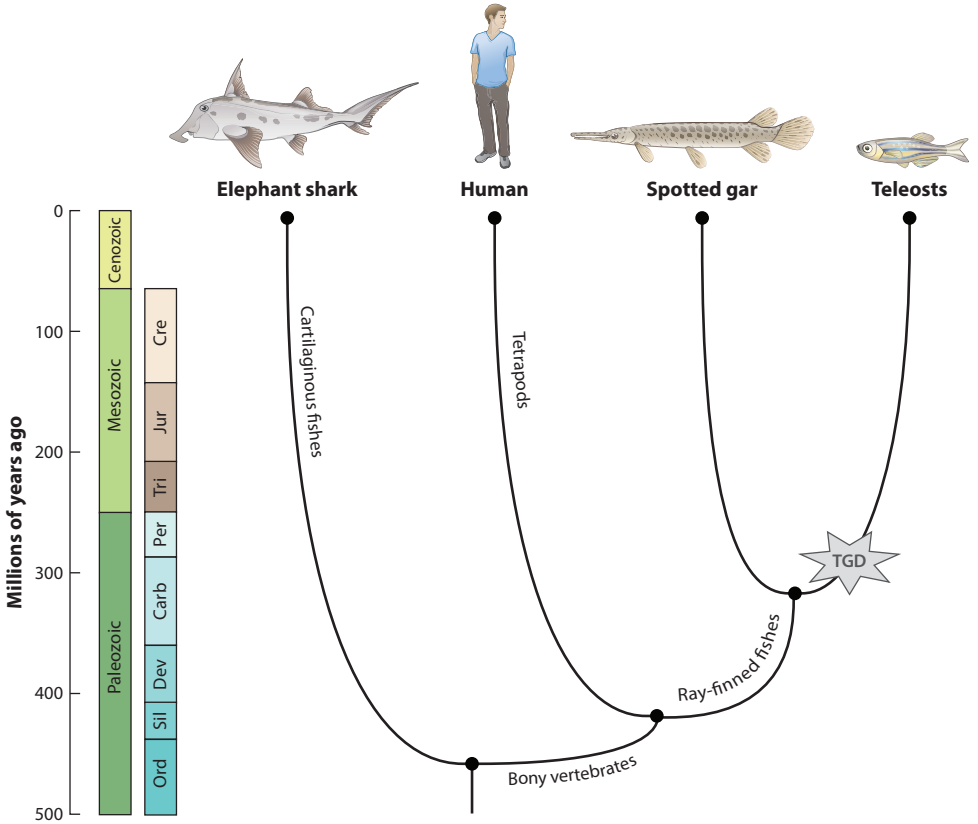


Figure 2

Cartilaginous fishes as an outgroup to bony vertebrates and spotted gar as an outgroup ray-finned fish to teleosts. A simple schematic tree showing the relative phylogenetic positions of teleosts, spotted gar, tetrapods (represented by human), and elephant shark (representing cartilaginous fishes). Elephant shark serves as a valuable reference (outgroup) for evolutionary studies of bony vertebrates, whereas spotted gar is a ray-finned fish outgroup to teleosts. Abbreviations: Carb, Carboniferous; Cre, Cretaceous; Dev, Devonian; Jur, Jurassic; Ord, Ordovician; Per, Permian; Tri, Triassic; Sil, Silurian; TGD, teleost-specific whole-genome duplication event.

salmonid-specific genome duplication, approximately 80 Mya (23, 24). Another affected lineage is that of Cypriniformes, in which several fishes, such as the common carp (*Cyprinus carpio*) and goldfish (*Carassius auratus*) (family Cyprinidae), loaches (Cobitidae), and suckers (Catostomidae), have experienced recent tetraploidization events. In the common carp lineage, several lines of evidence, including cytogenetic analysis and whole-genome sequencing, have confirmed a tetraploidization event approximately 8 Mya (25, 26). Besides teleosts, some basal ray-finned fish lineages have also undergone WGD independently. For example, sturgeons and paddlefish (Acipenseriformes) experienced tetraploidization events independently after their common ancestors diverged approximately 184 Mya (27).

The presence of WGD events at the base of several radiations, such as vertebrates, teleosts, and flowering plants, has led to the notion that WGD is causally related to morphological diversity and evolutionary success. Likewise, the TGD is believed to be responsible for the rapid speciation and adaptive radiation of teleosts. It is proposed that genome duplication events provide additional

genetic raw material on which selection can act and give rise to novel phenotypes. Following genome duplication, there are three common fates for the duplicated genes. The most likely fate is the loss of one of the duplicate loci through accumulation of deleterious mutations, resulting in nonfunctionalization (28). This loss of duplicated gene loci can potentially play a role in speciation via the phenomenon of divergent resolution (i.e., differential gene loss) (29, 30). Divergent resolution (or reciprocal gene loss) can occur when two isolated populations each lose different copies from a duplicated gene set. This process can potentially lead to reproductive isolation of the two populations and thereby result in speciation. Thus, lineages that have experienced genome duplication events, such as teleosts, could diversify via divergent resolution. The other two fates of duplicated genes are subfunctionalization (both copies retained because of partitioning of ancestral gene functions between the duplicates) and neofunctionalization (whereby one of the copies acquires a novel function) (31). The retained duplicate pairs often show a difference in overall expression patterns and expression levels (32). In rainbow trout (*Oncorhynchus mykiss*), following the salmonid-specific genome duplication event ~80 Mya, approximately half of the ancestral genes (48%) have been retained in duplicates (32), and a substantial proportion of them have divergent expression profiles and/or levels. This suggests that these duplicates have been retained owing to sub- or neofunctionalization.

Interestingly, recent studies suggest that there is little or no correlation between TGD and teleost diversity. By quantifying the rates of phenotypic evolution and innovation capacity in terms of size and shape of neopterygian fishes (teleosts plus their immediate ray-finned sister group, Holostei), a recent study showed that teleosts specifically do not show enhanced phenotypic evolution in comparison to holosteans (which include the spotted gar and bowfin) (33). Firstly, the authors noted that in contrast to the extant holosteans, the extinct Mesozoic holosteans showed comparable or even higher rates of size change to the teleosts. In addition, the diversity seen within the teleosts was not spread out uniformly and was in fact due to contributions from a few nested teleost subclades, like percomorphs and otophysans, which comprise more than 80% of the extant teleosts (34). The origin of these teleost subclades (percomorphs ~120 Mya and otophysans ~175 Mya) dates much later than the TGD event and the origin of the teleosts (5). Thus, according to this study, there does not seem to be any correlation between the TGD event and the diversification of teleosts (33). However, Robertson et al. (35) recently proposed that there could be a time lag between WGD events and subsequent evolutionary diversification. Their model, known as lineage-specific ohnologue resolution, predicts that the functional consequence of WGD may not appear explosively but can arise gradually over longer evolutionary periods that may be tens of millions of years after the initial duplication event. Thus, the evolutionary consequence of WGD is much more complex than previously thought and seems to affect lineages differently. Interestingly, lineages like cichlids, which display the highest speciation rates among vertebrates, did not experience a cichlid-specific WGD event (7, 36). In the case of cichlids, an array of genetic and environmental factors has been attributed to their adaptive radiation. In addition, analysis of genomes of five African cichlids revealed that their rapid evolutionary diversification could be due to accumulation of genetic variation within the population as a result of multiple mechanisms, such as accelerated evolution of coding and regulatory sequence, increased frequency of gene duplication, transposable element (TE) insertions near 5′ and 3′ untranslated regions, origin of novel microRNAs, and retention of ancient polymorphisms (36). These changes possibly enabled the cichlids to adapt and diversify in the newly available diverse ecological niches of East African lakes (36). Meier et al. (37) recently showed that hybridization between two divergent ancestral lineages of cichlids provided genetic variation that subsequently recombined and resulted in several cichlid species by exploiting new ecological niches of the East African lakes. In gadiform fishes (order Gadiformes), yet another mechanism has been attributed for the

elevated speciation rate (38). All members of this order lack major histocompatibility complex (MHC) II, which is involved in presenting exogenous antigens originating extracellularly, such as from bacteria. This is apparently compensated by an expansion of the MHC I genes, resulting in a high number of MHC I genes, with several species possessing ~100 copies. Likewise, species within Percomorphaceae also possess a higher number of MHC I genes, with some species having up to 80 copies. Interestingly, diversification rate analyses identified major shifts in the number of species per family in these two groups, thus establishing a positive association between the MHC I gene copy number and the speciation rate (38). Thus, the increase in MHC I diversity could have facilitated rapid speciation of teleosts by enabling them to colonize new environments.

Nevertheless, WGD events do correlate with elevated diversification rates, as seen in vertebrates (>60,000 species); teleosts (~30,000 species); and several other species, such as flowering plants (>350,000 species) (39). WGD events like the TGD likely contribute in some way to their evolutionary success, for example, by helping to overcome environmental changes and escape extinction. This is evident when we compare the species diversity between the extant teleosts and the nonteleost ray-finned fishes. Polyploids seem to be at an advantage when it comes to adapting to rapidly changing environments and major ecological upheavals compared with their diploid relatives (40). There may be a reduced risk of extinction for polyploids because of several factors, including functional redundancy, mutational robustness, and increased rates of evolution and adaptation (41). In fact, during vertebrate evolutionary history, genome duplication events have been preceded by multiple extinct lineages, which have resulted in gaps in the phylogeny of extant taxa prior to the duplication event (42). Therefore, WGDs seem to provide a competitive edge over their diploid progenitors and are especially advantageous when there are newly available niches. However, the direct involvement of WGD in the morphological diversity of teleosts remains unclear.

ACCELERATED MOLECULAR EVOLUTIONARY RATE

Previous comparative analysis of a limited set of protein-coding genes from teleosts and mammals had shown that teleost genes are evolving faster than their mammalian counterparts. For example, in a study comparing 19 genes from teleosts and mammals, a majority of the genes (16 out of 19) had higher evolutionary rates in teleosts than in mammals (43). A similar picture emerged subsequently when genome-wide protein sequences were compared between two pufferfishes (fugu and *Tetraodon*) and two mammals (human and mouse) (13). Analysis of fourfold degenerate (4D) sites, which is a proxy for neutral mutation rate, from a genome-wide set of 5,802 orthologous proteins from fugu, *Tetraodon*, human, and mouse showed that the nucleotide divergence rate between fugu and *Tetraodon* was twice as fast as that between human and mouse (13). Pairwise comparisons of the 5,802 orthologous protein sequences showed that the higher nucleotide divergence in the two pufferfish was true for protein sequences as well. Using *Ciona* (a urochordate) as an outgroup, the authors also noted a higher frequency of nonsynonymous substitutions between *Ciona* and *Tetraodon* as compared with *Ciona* and human. Finally, the ratio of nonsynonymous to synonymous changes was found to be much higher between the two pufferfishes as compared with the mammals (13). Thus, these studies based on a few species suggested that the overall molecular evolutionary rate in teleosts is higher than in mammals.

A much clearer picture was obtained when evolutionary rates were compared between genome-wide data sets of multiple teleosts and nonteleost vertebrates using vertebrate outgroups such as the elephant shark and the spotted gar. Phylogenomic analysis using a genome-scale data set of ~700 one-to-one orthologous protein sequences from elephant shark, two teleosts (zebrafish and stickleback), a lobe-finned fish (coelacanth), and seven tetrapods (human, mouse, cow, opossum,

chicken, green anole lizard, and western clawed frog) revealed that these single-copy genes were evolving faster in the two teleosts than in most other bony-vertebrate species (20). Likewise, the neutral nucleotide mutation rate, as estimated by analyzing the 4D sites of ~700 one-to-one orthologs, was found to be the highest for the teleosts (20). A similar analysis based on 243 one-to-one orthologous proteins from 25 jawed vertebrates, including the spotted gar, also showed a distinctly higher evolutionary rate for the teleosts as compared with mammals and other bony vertebrates (21). These analyses have clearly shown that a higher protein and nucleotide evolutionary rate distinguishes teleosts from other jawed vertebrates.

Various physiological and environmental factors have been proposed to explain the variation in molecular evolutionary rates between different groups of organisms. These include body size, weight-specific metabolic rate, generation time, and DNA repair efficiency (44–46). Previous studies have shown an inverse relationship between body size and substitution rate (45). However, body size in teleosts is highly variable and can range from approximately one centimeter (as in *Paedocypris* sp.) to a few meters (as in the ocean sunfish) (47, 48). Although some teleost fishes used for molecular evolutionary analysis, such as zebrafish, medaka, and *Tetraodon*, possess a relatively small body size, their molecular evolutionary rate is comparable to that of even the largest bony fish, the ocean sunfish, which can reach a maximum length of 2.7 m and weigh 2.3 tonnes. Phylogenomic analysis using a set of 1,690 one-to-one orthologs from ocean sunfish and teleosts such as fugu, *Tetraodon*, stickleback, medaka, tilapia, and zebrafish with spotted gar as an outgroup revealed that the enormous ocean sunfish is evolving at a rate similar to those of smaller teleosts like zebrafish and tilapia (49; V. Ravi & B. Venkatesh, unpublished data). Thus, body size does not seem to explain the variation in the molecular evolutionary rate between teleosts and other jawed vertebrates.

Metabolic rate may have a direct relation to evolutionary rate. A higher metabolic rate is hypothesized to result in an increased mutation rate via mechanisms such as an increase in DNA damage by mutagenic by-products of oxidative respiration and an increase in rates of DNA synthesis and nucleotide replacement (45, 50, 51). However, there exists a vast difference in the metabolic rate of teleosts, particularly those inhabiting marine habitats (52), and as such it is unclear whether the higher molecular evolutionary rate of teleosts is related to a higher metabolic rate. Likewise, there are no data to link the higher nucleotide substitution rate of teleosts and DNA repair fidelity.

Generation time, or age at sexual maturity, is another factor that may explain the evolutionary rate variation seen in various organisms. In general, species with a shorter generation time have faster molecular evolutionary rates owing to the higher replication frequency of their germline DNA, thereby increasing the chances of incorporating incorrect bases. Teleosts such as zebrafish, medaka, and *Tetraodon* have short generation times of two to four months (53), comparable to that of rodents, whereas teleosts such as fugu and ocean sunfish possess longer generation times of three to four years. Nevertheless, all these teleost genomes are evolving faster than genomes of other bony vertebrates. Conversely, the spotted gar, whose generation time is approximately one to two years (22), and the elephant shark, with a generation time of three to four years (54), are evolving markedly slower than the teleosts. Thus, the influence of physiological and environmental factors on the molecular evolutionary rate seems to be a highly complex one, and further work is required to clarify the roles of each of these factors. Note that population-genetic factors, such as the effective population size and the number of gametes produced, also affect the molecular evolutionary rates (55). One feature that distinguishes teleosts from most other jawed vertebrates is the TGD event. However, no study directly links WGD with an accelerated neutral nucleotide substitution rate. Thus, the factor(s) that have contributed to the accelerated neutral nucleotide substitution rate of teleosts remain unclear.

A higher evolutionary rate of protein and nucleotide sequences can potentially contribute to the phenotypic diversity of teleosts. The fast-evolving protein sequences have a higher probability of acquiring beneficial residues, which could result in proteins with more efficient or new functions. For example, genes associated with the lateral line system in teleosts were found to be evolving faster than their counterparts in tetrapods and exhibit positive selection (56). Likewise, higher nucleotide evolutionary rate can lead to a higher turnover of *cis*-regulatory elements, resulting in gain of novel regulatory elements or modification/loss of existing regulatory elements. Indeed, evolutionarily conserved noncoding elements (CNEs), a class of *cis*-regulatory elements, have been evolving faster in teleosts compared with other jawed vertebrates (57) (see the section titled Loss of Conserved Noncoding Elements below). Both gain and loss of *cis*-regulatory elements such as transcriptional enhancers have been shown to contribute to phenotypic innovation and diversity of animals (58–61). Thus, the accelerated molecular evolutionary rate in teleosts might have contributed to their phenotypic diversity.

HIGHER RATE OF INTRON TURNOVER

Intron turnover, or loss and gain of introns, is a relatively rare event in vertebrates. Previous studies comparing gain and loss of introns since the divergence of two mammalian lineages (human and mouse) or two teleost lineages (fugu and *Tetraodon*) have indicated a very low level of intron turnover in these lineages with virtually no gain of introns (62–64). These comparisons were, however, limited to a few vertebrate lineages and did not reflect the changes that occurred after the divergence of various vertebrate lineages from a common ancestor. Such a comparison became possible when the whole-genome sequence of the elephant shark, a cartilaginous fish and an outgroup to the bony vertebrates, became available. Genome-wide comparison of ~40,000 orthologous intron positions in elephant shark and several bony vertebrates (human, mouse, cow, opossum, chicken, anole lizard, *Xenopus tropicalis*, stickleback, and zebrafish) revealed a higher level of intron turnover in teleosts compared with the lobe-finned fish lineage (Sarcopterygii, which includes lobe-finned fishes and tetrapods) (**Figure 3**) (20). In the latter group, after an initial gain of 13 introns and loss of 10 introns prior to the diversification of various tetrapods, the intron turnover rate reduced substantially in various tetrapod lineages, with the highest turnover being 11 gains and 13 losses in the anole lizard lineage after it diverged from the common ancestor of tetrapods approximately 330 Mya (**Figure 3**). In contrast, after a gain of 37 introns and a loss of 54 introns in the common ancestor of teleosts, the zebrafish lineage experienced 49 gains and 53 losses, whereas the stickleback lineage experienced 603 gains and 126 losses after the two lineages split approximately 160 Mya (**Figure 3**) (20). These comparisons clearly show that intron turnover is not uncommon in vertebrates and that the teleost lineages in particular have experienced a substantially higher rate of intron gain and loss as compared with other jawed vertebrate lineages.

Several mechanisms/models have been proposed to explain the loss and gain of introns. The main models proposed include the reverse transcription (RT) model, the genomic DNA deletion model, and the nonhomologous end joining (NHEJ) of double-stranded breaks (DSBs) model. The RT model (recombination of partial or full-length reverse transcribed cDNA product with genomic DNA) generally results in a biased loss of 3′ introns with a tendency to lose adjacent introns (62, 65, 66). Unlike the RT model, in the genomic DNA deletion model, there is loss of individual introns, and this loss is usually imprecise as a result of unequal exchange of alleles (63, 67). The NHEJ repair of DSBs model has also been proposed as a mechanism for intron loss (68). This mechanism of intron loss was suggested in *Arabidopsis* based on microsimilarity at the splice sites of lost introns (69).

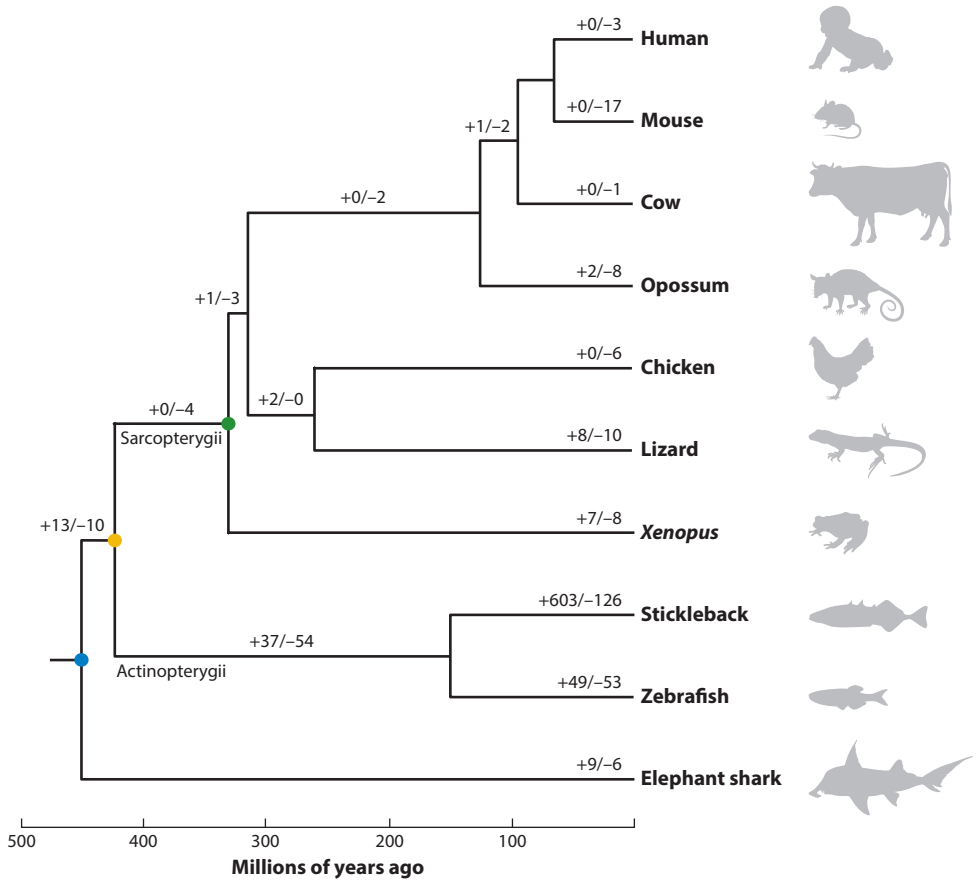


Figure 3

Intron turnover in jawed vertebrates. Phylogenetic tree of representative jawed vertebrates showing the number of introns gained (+) or lost (−) at ancestral nodes as well as terminal branches of the tree [data from Venkatesh et al. (20)]. Sarcopterygii, lobe-finned fishes + tetrapods; Actinopterygii, ray-finned fishes including teleosts. The green dot represents the common ancestor of all tetrapods, the yellow dot represents the common ancestor of bony vertebrates, and the blue dot represents the common ancestor of jawed vertebrates.

At least six mechanisms have been proposed to explain the gain of introns. These include intron transposition, transposon insertion, tandem genomic duplication, NHEJ-mediated intron gain, group II intron insertion, and intron transfer (70). According to the intron transposition model, an intron can be reverse-spliced into the same mRNA at a different position or into another mRNA. Subsequently, RT of this mRNA and recombination of the resultant cDNA molecule with genomic DNA can create a new intron (71–73). TEs that contain cryptic splice sites may get inserted into the genome, thus creating novel introns (67, 69). A recent study revealed large-scale gain of introns in two distantly related algal species via short DNA transposons known as introner elements, each of which possess a splice site toward one end (74). Tandem genomic repeats possessing cryptic splice sites can also result in new introns (75, 76). Genes with highly similar regions, such as in paralogs, may lead to transfer of introns between the two via gene

conversion (72, 77). NHEJ-mediated DSB repair has been implicated in the gain of many introns in *Aspergillus* and *Daphnia* (77, 78).

The reasons behind the higher intron turnover rate in teleosts are not known. Because teleosts have experienced a higher rate of intron gain as well as intron loss, a mechanism(s) that can explain both of the processes seems to be mediating the higher turnover of introns in teleosts. One such mechanism depends on the high level of TE activity in the genome. It is interesting that teleost genomes possess the highest diversity of TEs within the vertebrates (79). In contrast, lobe-finned fishes and tetrapods (sarcopterygians) have lost several TE superfamilies. Unlike in most other vertebrate lineages, including the basal ray-finned fish, the spotted gar, DNA transposons seem to be the predominant form of TE within most teleost genomes (79). In addition, teleost genomes seem to possess more of the recent TEs and less of the ancient TEs, indicating that the TEs have been active in the recent past in teleost genomes (79). Because TEs such as DNA transposons are known to be an important source of novel introns, this might be one of the factors contributing to the higher intron gain in teleost genomes. TEs can also serve as substrates for homologous recombination, resulting in DNA rearrangements such as deletions, duplications, inversions, and translocations (80, 81). This in turn can result in intron loss. Another mechanism that can lead to gain or loss of introns is the NHEJ-mediated DSB repair. However, there is no evidence to suggest that fish genomes experience a higher level of DSB than other vertebrates. Interestingly, a dramatic acceleration in the intron turnover rate of duplicated genes compared with orthologous genes has been shown in two species of malaria parasite (*Plasmodium falciparum* and *Plasmodium yoelii yoelii*) (82). Teleost genomes are known to have retained approximately 20% of duplicate genes post-TGD. These duplicate genes might have contributed to the higher rate of intron turnover. Thus, several factors are likely to have resulted in the higher intron turnover in teleosts.

It is not clear whether a higher turnover of introns can contribute to phenotypic diversity of teleosts. Here we have taken liberty to speculate on some possible mechanisms. Gain of introns creates additional genomic templates where new *cis*-regulatory elements can evolve and alter or modulate the expression pattern of associated genes. Similarly, loss of introns can result in the deletion of regulatory elements, which can also lead to an altered expression pattern of their target genes. Moreover, intron turnover has been shown to affect the level of gene expression. Intron-bearing genes generally exhibit higher expression levels as compared with intronless genes in yeast as well as in mammals (83, 84). Thus, intron turnover not only can lead to altered expression patterns but also can potentially lead to an increase or decrease of gene expression levels. Alternative splicing is a major mechanism that can generate novel proteins by generating multiple isoforms of proteins that can exhibit differential expression. Gain of introns can facilitate generation of new alternative transcripts and thereby create novel protein isoforms. The altered expression patterns of genes as well as generation of novel protein isoforms have the potential to give rise to phenotypic differences between species. Thus, the higher turnover of introns in teleosts might have played a role in the phenotypic diversity of teleosts.

LOSS OF CONSERVED NONCODING ELEMENTS

Comparative analyses of vertebrate genomes have indicated that a higher proportion of conserved sequences are located in the noncoding portions of the genome rather than the protein-coding regions (85, 86). Vertebrate genomes contain thousands of noncoding elements that are under evolutionary constraint (85). Several studies have shown that these sequences, known as CNEs, are frequently found near genes involved in the regulation of development and transcription (87–89). Functional assays have revealed that a substantial proportion of CNEs function as *cis*-regulatory

elements, such as enhancers that drive tissue-specific expression of the associated gene during early stages of development (87, 88, 90, 91).

Because human and teleost lineages diverged approximately 400 Mya, it was hypothesized that noncoding sequences conserved between them are likely to be functional elements that are under selective constraint, and human-teleost comparison was proposed as an effective strategy for identifying *cis*-regulatory elements in the human genome (92). Indeed, a CNE identified between human, mouse, and fugu was shown to function as a transcriptional enhancer recapitulating the expression pattern of its neighboring gene (92). Subsequent whole-genome comparison of human with fugu and zebrafish genomes identified several hundred CNEs, and functional assay of a subset of them revealed that many CNEs show enhancer activity in one or more tissues (87, 88, 93). Thus, the teleost genome was considered an ideal model for genome-wide identification of evolutionarily conserved functional noncoding elements in the human genome. However, with the availability of sequences from outgroups like the elephant shark, it was soon realized that teleost genomes were in fact missing a large fraction of CNEs that were present in the common ancestor of jawed vertebrates but still retained in elephant shark and human. Survey sequencing of the elephant shark genome and comparison of its sequences to human, fugu, and zebrafish indicated that elephant shark and human shared almost twice the number of CNEs (4,782) compared with fugu-human (2,107) and zebrafish-human (2,838) comparisons (94). In a more detailed comparative analysis of elephant shark sequences with whole genomes of four tetrapods (human, dog, chicken, and *Xenopus*) and four teleosts (zebrafish, stickleback, medaka, and fugu), elephant shark and tetrapods were found to possess roughly twice the number of CNEs (3,862–6,489 CNEs) compared with those found in elephant shark and teleosts (1,465–1,936 CNEs) (57). Thus, the teleosts have lost a substantial proportion of the ancestral jawed vertebrate CNEs that are still conserved in other jawed vertebrates. The nucleotide substitution rate of CNEs retained in teleosts showed that the teleost CNEs were evolving faster than their counterparts in tetrapods, suggesting that the rapid nucleotide substitution rate of teleosts is largely responsible for the missing (divergence beyond recognition) CNEs in teleosts (57). The recent availability of a high-quality, whole-genome assembly of the elephant shark and in-depth comparative analysis with whole genomes of 11 bony vertebrates (human, mouse, cow, opossum, chicken, lizard, *X. tropicalis*, fugu, stickleback, medaka, zebrafish) revealed that the teleosts have lost eight times more CNEs than the tetrapods (20), confirming previous findings. These whole-genome comparisons also showed that a vast majority of teleost CNEs (~90%) were lost despite their putative target genes still being present in the genomes, indicating that the loss of these CNEs was not due to loss of their target genes in teleosts.

It is becoming increasingly recognized that regulatory evolution is a major factor driving morphological variations and phenotypic innovations in animals (61, 95, 96). In addition to gain of enhancers, loss of enhancers has also been shown to contribute to phenotypic differences between different populations of the same species or between closely related species (59, 61, 97, 98). For instance, deletion of a hindlimb enhancer of the *Pitx1* gene has been implicated in the loss of pelvic spines in freshwater populations of the threespine stickleback (59). Similarly, loss of the sensory vibrissae and penile spine enhancer of the androgen receptor gene is associated with the absence of sensory vibrissae and penile spines in the human lineage (58). The above-mentioned comparisons of teleost and other jawed vertebrate genomes have shown that teleosts have lost thousands of potential ancient *cis*-regulatory elements that are still conserved in other jawed vertebrates, like elephant shark and human. Furthermore, CNEs that are retained in teleosts have been evolving rapidly as compared with their homologs in other jawed vertebrates. Both loss and divergent sequences of regulatory elements have the potential to alter the expression patterns of their target genes and to lead to phenotypic differences between species. Given that thousands of potential

cis-regulatory elements have been lost in teleosts and regulatory elements that are retained are accumulating substitutions at a higher rate, it is highly likely that these regulatory changes have played a major role in the phenotypic diversity of teleosts. This hypothesis needs to be verified by a systematic knockout of mouse enhancers that have been retained in other jawed vertebrates but lost in teleosts. Likewise, knockout of zebrafish enhancers that are lost in other teleosts should indicate whether the loss of these enhancers has contributed to phenotypic differences between teleosts.

SHORTER CONSERVED SYNTENIC BLOCKS

Comparisons between the genomes of teleosts, the spotted gar (a nonteleost ray-finned fish), and the elephant shark (a cartilaginous fish) have shown that the teleost lineage has experienced a higher rate of interchromosomal rearrangements (20–22). Interestingly, these rearrangement events seem to have occurred in the common ancestor of teleosts before the TGD event (21). Nevertheless, the higher rearrangement rate in the teleost lineage has resulted in shorter conserved syntenic blocks in teleosts as compared with spotted gar and other jawed vertebrates. In addition to the pre-TGD chromosomal rearrangements, the TGD event has also contributed to the shorter syntenic blocks through differential gene retention/loss without any chromosomal rearrangements. Differential gene retention/loss is the process by which paralogous regions (paralogons) generated by WGD events each lose a different copy of the duplicated gene(s) (see **Figure 4*a*** for a simplified schematic). Assuming the pre-TGD genomic region contained three genes (genes 1–3), following the TGD event, two paralogous regions (a-copy and b-copy) containing duplicate copies of the three genes (1a and 1b, 2a and 2b, and 3a and 3b) are generated. Subsequently, gene 2a is lost from the a-copy, whereas genes 1b and 3b are lost from b-copy. Because of this differential gene loss between the paralogous regions, gene 2b (now a single-copy gene) is now located on a chromosome different from that of genes 1a and 3a (each one is now a single-copy gene). Differential loss of blocks of such duplicate genes can result in disruption of synteny, leading to shorter syntenic blocks between teleosts and nonteleost vertebrates. There are many such disrupted syntenic blocks in teleost genomes. For example, the syntenic block in the *Pax6* locus contains *Wt1*, *Rcn1*, *Pax6*, *Elp4*, *Immp1l*, and *Dnajc24* genes, and this block is totally conserved in elephant shark and tetrapods. However, in zebrafish, which contains duplicate *Pax6* loci owing to the TGD, the *pax6a* locus has retained genes *wt1*, *pax6a*, and *dnajc24*, whereas the *pax6b* locus retains *rcn1*, *pax6b*, *elp4*, and *immp1l* (99) [see another example (*Tmem138-Tmem216* locus) in **Figure 4*b***].

The differential gene loss of duplicate genes in paralogous regions of teleost genomes has actually helped scientists in identifying genomic regulatory blocks (GRBs) and assigning target genes to *cis*-regulatory elements embedded within the GRBs in human and other vertebrate genomes. GRBs are chromosomal segments harboring *cis*-regulatory elements, their target genes, and some unrelated bystander genes (100). The bystander genes are not under the control of the regulatory elements driving the target genes and show unrelated expression patterns. Duplication of GRBs as in teleosts allows the unrelated bystander gene to be lost from one of the loci and thereby enables the identification of the retained target gene and assignment of the regulatory elements to that particular gene (100). This is unlike the case in single-copy GRBs, which are protected from chromosomal breakage. GRBs, in fact, form the basis for large-scale conserved synteny for many chromosomal regions by imposing long-range constraints on their integrity as well as structure, for instance, *Hox* gene clusters and the *Irx* loci (101, 102).

The differential loss of duplicate genes from paralogous regions in teleosts has also posed problems in using teleosts for comparative genomic studies and especially for inferring the evolutionary

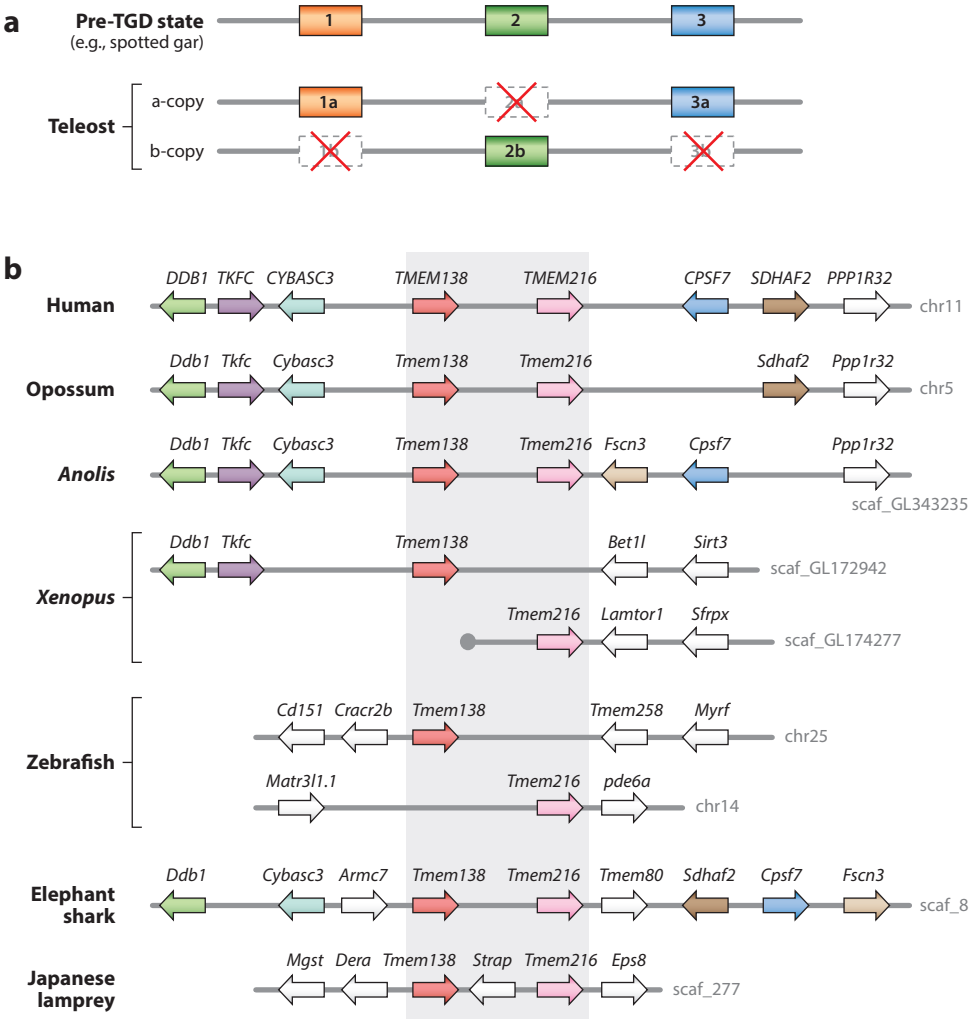


Figure 4

Differential gene loss in duplicate chromosomes of teleosts. (*a*) A simplified schematic showing an example of differential gene loss following the teleost-specific genome duplication (TGD) event. Boxes represent genes, and red crosses represent gene losses. (*b*) The *TMEM138-TMEM216* locus from various vertebrates showing the linkage and orientation of the two genes as well as synteny of flanking genes [data from Lee et al. (103) and Venkatesh et al. (104)]. Note the presence of *Tmem138* and *Tmem216* on different chromosomes and the complete disruption of syntenic blocks in zebrafish compared with elephant shark and tetrapods. In *Xenopus*, the genome assembly comprises scaffolds, and hence, it is not clear whether *Tmem138* and *Tmem216* are present on the same or different chromosomes. Syntenic genes are shaded, whereas nonsyntenic ones are left blank. Grey circles denote scaffold ends.

history of human genes. For instance, in the human genome, two *TMEM* genes, *TMEM138* and *TMEM216*, are closely linked. Mutations in either of them lead to phenotypically indistinguishable ciliopathy, also known as Joubert syndrome (103). In an elegant study, Lee et al. (103) showed that the two genes constituted a functional gene cluster and were tightly coexpressed through a shared regulatory element present in their intergenic region. These two genes are also closely

linked in other mammals and the *Anolis* lizard. However, the two genes were found to be present on different scaffolds in the draft assembly of *X. tropicalis* and on two different chromosomes in zebrafish (**Figure 4*b***) (103). Based on this observation, the authors concluded that the two TMEM genes were brought together into a functional unit at the amphibian-to-reptile evolutionary transition by a chromosomal rearrangement event (103). However, subsequent analysis of this locus in the elephant shark (104) showed that *Tmem138* and *Tmem216* are closely linked in the elephant shark genome similar to in mammals and the *Anole* lizard, indicating that the linkage of the two genes is an ancestral feature and their presence on two different chromosomes in zebrafish is a derived feature. Thus, Lee et al.'s inference based on comparison to zebrafish was incorrect. Analysis of additional teleost species showed that the two *TMEM* genes are present on different chromosomes in other teleosts (medaka, stickleback, fugu, and *Tetraodon*) as well (104). These comparisons show that the two TMEM genes were already linked in the ancestor of all jawed vertebrates, and following the TGD, duplicates of the two TMEM genes were differentially lost in the paralogous chromosomes in teleosts. This example highlights the need to exercise caution when using teleost genomes as a reference for comparative studies. Such comparisons should always include outgroups, such as a cartilaginous fish or a nonteleost ray-finned fish.

SUMMARY

Teleosts are the largest and most diverse group of vertebrates. Recent whole-genome comparisons of teleosts and other vertebrates using outgroups such as the elephant shark and spotted gar have shown that teleost genomes are evolving faster and have accumulated several changes that make them quite divergent as compared with the genomes of other jawed vertebrates. These changes include acceleration in the evolutionary rate of protein-coding sequences and nucleotide sequences, high rate of intron gain and loss in protein-coding genes, loss of thousands of potential *cis*-regulatory elements, and shorter syntenic blocks. Some of these changes likely have contributed to the phenotypic diversity of teleosts. For example, fast-evolving proteins could have led to the evolution of proteins with altered function or altogether novel functions. Likewise, the higher turnover of nucleotide sequences could have contributed to a higher turnover of *cis*-regulatory elements. It is now well recognized that regulatory changes are a major force driving evolution of phenotypic innovations and morphological changes between species. Analysis of CNEs, which are proxies for *cis*-regulatory elements, has provided strong evidence for loss (divergence beyond recognition) as well as divergence in the sequences of regulatory elements in teleosts. These changes have great potential for giving rise to phenotypic changes between different species of teleosts. This hypothesis can be verified by knocking out *cis*-regulatory elements or introducing mutations in *cis*-regulatory elements in mouse or teleost models using the CRISPR/Cas9 system. This gene-editing technology is extremely efficient, such that the phenotype is obvious even in the F0 generation, and thus it can be used even in nonmodel organisms as long as they can be bred in captivity or mature eggs and sperm are available for fertilization. As such this offers a powerful strategy for understanding the genetic basis of phenotypic differences between closely related species of teleosts.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

1. Int. Union Conserv. Nat. 2017. *The IUCN red list of threatened species, version 2017-2*. Cambridge, UK: Int. Union Conserv. Nat., accessed Sep. 14, 2017. **http://www.iucnredlist.org**.
2. Betancur RR, Broughton RE, Wiley EO, Carpenter K, Lopez JA, et al. 2013. The tree of life and a new classification of bony fishes. *PLOS Curr*. 5. **https://doi.org/10.1371/currents.tol.53ba26640df0ccaee75bb165c8c26288**
3. de Pinna MCC. 1996. Teleostean monophyly. In *Interrelationships of Fishes*, ed. MLJ Stiassny, LR Parenti, GD Johnson, pp. 147–62. Cambridge, MA: Academic
4. Patterson C, Rosen D. 1977. Review of ichthyodectiform and other Mesozoic teleost fishes and the theory and practice of classifying fossils. *Bull. Am. Mus. Nat. Hist.* 158:83–172
5. Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, et al. 2012. Resolution of ray-finned fish phylogeny and timing of diversification. *PNAS* 109:13698–703
6. Nelson JS, Grande TC, Wilson MVH. 2016. *Fishes of the World*. Hoboken, NJ: Wiley
7. Kocher TD. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. *Nat. Rev. Genet.* 5:288–98
8. Wagner CE, Harmon LJ, Seehausen O. 2012. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* 487:366–69
9. Dooley K, Zon LI. 2000. Zebrafish: a model system for the study of human disease. *Curr. Opin. Genet. Dev.* 10:252–56
10. Kirchmaier S, Naruse K, Wittbrodt J, Loosli F. 2015. The genomic and genetic toolbox of the teleost medaka (*Oryzias latipes*). *Genetics* 199:905–18
11. Amores A, Force A, Yan YL, Joly L, Amemiya C, et al. 1998. Zebrafish hox clusters and vertebrate genome evolution. *Science* 282:1711–14
12. Christoffels A, Koh EG, Chia JM, Brenner S, Aparicio S, Venkatesh B. 2004. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. *Mol. Biol. Evol.* 21:1146–51
13. Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, et al. 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431:946–57
14. Taylor JS, Braasch I, Frickey T, Meyer A, Van de Peer Y. 2003. Genome duplication, a trait shared by 22,000 species of ray-finned fish. *Genome Res.* 13:382–90
15. Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, et al. 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* 297:1301–10
16. Brenner S, Elgar G, Sandford R, Macrae A, Venkatesh B, Aparicio S. 1993. Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature* 366:265–68
17. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, et al. 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496:498–503
18. Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W, et al. 2007. The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447:714–19
19. Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61
20. Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, et al. 2014. Elephant shark genome provides unique insights into gnathostome evolution. *Nature* 505:174–79
21. Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, et al. 2016. The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nat. Genet.* 48:427–37
22. Amores A, Catchen J, Ferrara A, Fontenot Q, Postlethwait JH. 2011. Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. *Genetics* 188:799–808
23. Lien S, Koop BF, Sandve SR, Miller JR, Kent MP, et al. 2016. The Atlantic salmon genome provides insights into rediploidization. *Nature* 533:200–5
24. 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. Biol. Sci.* 281:20132881

25. Ohno S, Muramoto J, Christian L, Atkin NB. 1967. Diploid-tetraploid relationship among old-world members of the fish family Cyprinidae. *Chromosoma* 23:1–9
26. Xu P, Zhang X, Wang X, Li J, Liu G, et al. 2014. Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nat. Genet.* 46:1212–19
27. Peng Z, Ludwig A, Wang D, Diogo R, Wei Q, He S. 2007. Age and biogeography of major clades in sturgeons and paddlefishes (Pisces: Acipenseriformes). *Mol. Phylogenet. Evol.* 42:854–62
28. Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290:1151–55
29. Semon M, Wolfe KH. 2007. Reciprocal gene loss between *Tetraodon* and zebrafish after whole genome duplication in their ancestor. *Trends Genet.* 23:108–12
30. Taylor JS, Van de Peer Y, Meyer A. 2001. Genome duplication, divergent resolution and speciation. *Trends Genet.* 17:299–301
31. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151:1531–45
32. Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, et al. 2014. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat. Commun.* 5:3657
33. Clarke JT, Lloyd GT, Friedman M. 2016. Little evidence for enhanced phenotypic evolution in early teleosts relative to their living fossil sister group. *PNAS* 113:11531–36
34. 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
35. Robertson F, Gundappa MK, Grammes F, Hvidsten T, Redmond A, et al. 2017. Lineage-specific rediploidization is a mechanism to explain time-lags between genome duplication and evolutionary diversification. *Genome Biol.* 18:111
36. Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, et al. 2014. The genomic substrate for adaptive radiation in African cichlid fish. *Nature* 513:375–81
37. Meier JI, Marques DA, Mwaiko S, Wagner CE, Excoffier L, Seehausen O. 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun.* 8:14363
38. Malmstrom M, Matschiner M, Torresen OK, Star B, Snipen LG, et al. 2016. Evolution of the immune system influences speciation rates in teleost fishes. *Nat. Genet.* 48:1204–10
39. Van de Peer Y, Maere S, Meyer A. 2009. The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* 10:725–32
40. Fawcett JA, Maere S, Van de Peer Y. 2009. Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. *PNAS* 106:5737–42
41. Crow KD, Wagner GP. 2006. 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.* 23:887–92
42. Donoghue PC, Purnell MA. 2005. Genome duplication, extinction and vertebrate evolution. *Trends Ecol. Evol.* 20:312–19
43. Robinson-Rechavi M, Laudet V. 2001. Evolutionary rates of duplicate genes in fish and mammals. *Mol. Biol. Evol.* 18:681–83
44. Britten RJ. 1986. Rates of DNA sequence evolution differ between taxonomic groups. *Science* 231:1393–98
45. Martin AP, Palumbi SR. 1993. Body size, metabolic rate, generation time, and the molecular clock. *PNAS* 90:4087–91
46. Schlotterer C, Amos B, Tautz D. 1991. Conservation of polymorphic simple sequence loci in cetacean species. *Nature* 354:63–65
47. Kottelat M, Britz R, Hui TH, Witte KE. 2006. *Paedocypris*, a new genus of Southeast Asian cyprinid fish with a remarkable sexual dimorphism, comprises the world's smallest vertebrate. *Proc. Biol. Sci.* 273:895–99
48. Pope EC, Hays GC, Thys TM, Doyle TK, Sims DW, et al. 2010. The biology and ecology of the ocean sunfish *Mola*: a review of current knowledge and future research perspectives. *Rev. Fish Biol. Fish.* 20:471–87

49. Pan H, Yu H, Ravi V, Li C, Lee AP, et al. 2016. The genome of the largest bony fish, ocean sunfish (*Mola mola*), provides insights into its fast growth rate. *GigaScience* 5:36
50. Richter C, Park JW, Ames BN. 1988. Normal oxidative damage to mitochondrial and nuclear DNA is extensive. *PNAS* 85:6465–67
51. Wagner JR, Hu CC, Ames BN. 1992. Endogenous oxidative damage of deoxycytidine in DNA. *PNAS* 89:3380–84
52. Seibel BA, Drazen JC. 2007. The rate of metabolism in marine animals: environmental constraints, ecological demands and energetic opportunities. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362:2061–78
53. Lawrence C, Adatto I, Best J, James A, Maloney K. 2012. Generation time of zebrafish (*Danio rerio*) and medakas (*Oryzias latipes*) housed in the same aquaculture facility. *Lab. Anim.* 41:158–65
54. Sullivan KJ. 1977. Age and growth of the elephant fish *Callorhinchus milii* (Elasmobranchii: Callorhynchidae). *N. Z. J. Mar. Freshw. Res.* 11:745–53
55. Lynch M, Ackerman MS, Gout JF, Long H, Sung W, et al. 2016. Genetic drift, selection and the evolution of the mutation rate. *Nat. Rev. Genet.* 17:704–14
56. Philip S, Machado JP, Maldonado E, Vasconcelos V, O'Brien SJ, et al. 2012. Fish lateral line innovation: insights into the evolutionary genomic dynamics of a unique mechanosensory organ. *Mol. Biol. Evol.* 29:3887–98
57. Lee AP, Kerk SY, Tan YY, Brenner S, Venkatesh B. 2011. Ancient vertebrate conserved noncoding elements have been evolving rapidly in teleost fishes. *Mol. Biol. Evol.* 28:1205–15
58. McLean CY, Reno PL, Pollen AA, Bassan AI, Capellini TD, et al. 2011. Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature* 471:216–19
59. Chan YF, Marks ME, Jones FC, Villarreal G Jr., Shapiro MD, et al. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* 327:302–5
60. Koshikawa S, Giorgianni MW, Vaccaro K, Kassner VA, Yoder JH, et al. 2015. Gain of *cis*-regulatory activities underlies novel domains of wingless gene expression in *Drosophila*. *PNAS* 112:7524–29
61. Wittkopp PJ, Haerum BK, Clark AG. 2008. Regulatory changes underlying expression differences within and between *Drosophila* species. *Nat. Genet.* 40:346–50
62. Coulombe-Huntington J, Majewski J. 2007. Characterization of intron loss events in mammals. *Genome Res.* 17:23–32
63. Loh YH, Brenner S, Venkatesh B. 2008. Investigation of loss and gain of introns in the compact genomes of pufferfishes (*Fugu* and *Tetraodon*). *Mol. Biol. Evol.* 25:526–35
64. Roy SW, Fedorov A, Gilbert W. 2003. Large-scale comparison of intron positions in mammalian genes shows intron loss but no gain. *PNAS* 100:7158–62
65. Ma MY, Che XR, Porceddu A, Niu DK. 2015. Evaluation of the mechanisms of intron loss and gain in the social amoebae *Dictyostelium*. *BMC Evol. Biol.* 15:286
66. Roy SW, Gilbert W. 2005. The pattern of intron loss. *PNAS* 102:713–18
67. Yenerall P, Krupa B, Zhou L. 2011. Mechanisms of intron gain and loss in *Drosophila*. *BMC Evol. Biol.* 11:364
68. Farlow A, Meduri E, Schlotterer C. 2011. DNA double-strand break repair and the evolution of intron density. *Trends Genet.* 27:1–6
69. Fawcett JA, Rouze P, Van de Peer Y. 2012. Higher intron loss rate in *Arabidopsis thaliana* than *A. lyrata* is consistent with stronger selection for a smaller genome. *Mol. Biol. Evol.* 29:849–59
70. Yenerall P, Zhou L. 2012. Identifying the mechanisms of intron gain: progress and trends. *Biol. Direct* 7:29
71. Roy SW, Irimia M. 2009. Mystery of intron gain: new data and new models. *Trends Genet.* 25:67–73
72. Torriani SF, Stukenbrock EH, Brunner PC, McDonald BA, Croll D. 2011. Evidence for extensive recent intron transposition in closely related fungi. *Curr. Biol.* 21:2017–22
73. Denoeud F, Henriet S, Mungpakdee S, Aury JM, Da Silva C, et al. 2010. Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. *Science* 330:1381–85
74. Huff JT, Zilberman D, Roy SW. 2016. Mechanism for DNA transposons to generate introns on genomic scales. *Nature* 538:533–36
75. Gao X, Lynch M. 2009. Ubiquitous internal gene duplication and intron creation in eukaryotes. *PNAS* 106:20818–23

76. Sharpton TJ, Neafsey DE, Galagan JE, Taylor JW. 2008. Mechanisms of intron gain and loss in *Cryptococcus*. *Genome Biol*. 9:R24
77. Zhang LY, Yang YF, Niu DK. 2010. Evaluation of models of the mechanisms underlying intron loss and gain in *Aspergillus* fungi. *J. Mol. Evol.* 71:364–73
78. Li W, Tucker AE, Sung W, Thomas WK, Lynch M. 2009. Extensive, recent intron gains in *Daphnia* populations. *Science* 326:1260–62
79. Chalopin D, Naville M, Plard F, Galiana D, Volff JN. 2015. Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. *Genome Biol. Evol.* 7:567–80
80. Burns KH, Boeke JD. 2012. Human transposon tectonics. *Cell* 149:740–52
81. Kazazian HH Jr. 2004. Mobile elements: drivers of genome evolution. *Science* 303:1626–32
82. Castillo-Davis CI, Bedford TB, Hartl DL. 2004. Accelerated rates of intron gain/loss and protein evolution in duplicate genes in human and mouse malaria parasites. *Mol. Biol. Evol.* 21:1422–27
83. Juneau K, Miranda M, Hillenmeyer ME, Nislow C, Davis RW. 2006. Introns regulate RNA and protein abundance in yeast. *Genetics* 174:511–18
84. Shabalina SA, Ogurtsov AY, Spiridonov AN, Novichkov PS, Spiridonov NA, Koonin EV. 2010. Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. *Mol. Biol. Evol.* 27:1745–49
85. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, et al. 2004. Ultraconserved elements in the human genome. *Science* 304:1321–25
86. Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, et al. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478:476–82
87. Shin JT, Priest JR, Ovcharenko I, Ronco A, Moore RK, et al. 2005. Human-zebrafish non-coding conserved elements act *in vivo* to regulate transcription. *Nucleic Acids Res*. 33:5437–45
88. Woolfe A, Goodson M, Goode DK, Snell P, McEwen GK, et al. 2005. Highly conserved non-coding sequences are associated with vertebrate development. *PLOS Biol*. 3:e7
89. Sandelin A, Bailey P, Bruce S, Engstrom PG, Klos JM, et al. 2004. Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. *BMC Genom*. 5:99
90. Navratilova P, Fredman D, Hawkins TA, Turner K, Lenhard B, Becker TS. 2009. Systematic human/zebrafish comparative identification of *cis*-regulatory activity around vertebrate developmental transcription factor genes. *Dev. Biol.* 327:526–40
91. Visel A, Prabhakar S, Akiyama JA, Shoukry M, Lewis KD, et al. 2008. Ultraconservation identifies a small subset of extremely constrained developmental enhancers. *Nat. Genet.* 40:158–60
92. Aparicio S, Morrison A, Gould A, Gilthorpe J, Chaudhuri C, et al. 1995. Detecting conserved regulatory elements with the model genome of the Japanese puffer fish, *Fugu rubripes*. *PNAS* 92:1684–88
93. Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, et al. 2006. *In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* 444:499–502
94. Venkatesh B, Kirkness EF, Loh YH, Halpern AL, Lee AP, et al. 2006. Ancient noncoding elements conserved in the human genome. *Science* 314:1892
95. Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134:25–36
96. Wray GA. 2007. The evolutionary significance of *cis*-regulatory mutations. *Nat. Rev. Genet.* 8:206–16
97. Indjeian VB, Kingman GA, Jones FC, Guenther CA, Grimwood J, et al. 2016. Evolving new skeletal traits by *cis*-regulatory changes in bone morphogenetic proteins. *Cell* 164:45–56
98. Attanasio C, Nord AS, Zhu Y, Blow MJ, Li Z, et al. 2013. Fine tuning of craniofacial morphology by distant-acting enhancers. *Science* 342:1241006
99. Ravi V, Bhatia S, Gautier P, Loosli F, Tay BH, et al. 2013. Sequencing of *Pax6* loci from the elephant shark reveals a family of *Pax6* genes in vertebrate genomes, forged by ancient duplications and divergences. *PLOS Genet*. 9:e1003177
100. Kikuta H, Laplante M, Navratilova P, Komisarczuk AZ, Engstrom PG, et al. 2007. Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. *Genome Res*. 17:545–55

101. Lee AP, Koh EG, Tay A, Brenner S, Venkatesh B. 2006. Highly conserved syntenic blocks at the vertebrate *Hox* loci and conserved regulatory elements within and outside *Hox* gene clusters. *PNAS* 103:6994–99
102. Maeso I, Irimia M, Tena JJ, Gonzalez-Perez E, Tran D, et al. 2012. An ancient genomic regulatory block conserved across bilaterians and its dismantling in tetrapods by retrogene replacement. *Genome Res.* 22:642–55
103. Lee JH, Silhavy JL, Lee JE, Al-Gazali L, Thomas S, et al. 2012. Evolutionarily assembled *cis*-regulatory module at a human ciliopathy locus. *Science* 335:966–69
104. Venkatesh B, Ravi V, Lee AP, Warren WC, Brenner S. 2013. Basal vertebrates clarify the evolutionary history of ciliopathy-associated genes *Tmem138* and *Tmem216*. *Mol. Biol. Evol.* 30:62–65
105. Bian C, Hu Y, Ravi V, Kuznetsova IS, Shen X, et al. 2016. The Asian arowana (*Scleropages formosus*) genome provides new insights into the evolution of an early lineage of teleosts. *Sci. Rep.* 6:24501

Annual Review of
Animal Biosciences

Volume 6, 2018

Contents

An online log of corrections to *Annual Review of Animal Biosciences* articles may be found at http://www.annualreviews.org/errata/animal