Fish Genomes, Comparative Genomics and Vertebrate Evolution

A. Froschauer[1], I. Braasch[2] and J.-N. Volff [2,*]

[1]*TU Dresden, Institut für Zoologie, Zellescher Weg 20b, 01217 Dresden, Germany and* [2]*University of Würzburg, Physiologische Chemie I, Am Hubland, 97074 Würzburg, Germany*

Abstract: Genome sequences from the pufferfishes *Takifugu rubripes* (Fugu) and *Tetraodon nigroviridis*, the zebrafish *Danio rerio* and the medaka *Oryzias latipes* together with genomic data from various other fish species have opened an important era of comparative genomics shedding a new light on the structure and evolution of vertebrate genomes. For instance, comparative analysis of fish genomes has revealed that the ancestral bony vertebrate genome was composed of 12 chromosomes, has confirmed the occurrence of at least one event of genome duplication in the early history of vertebrates and has allowed the identification of conserved regulatory and coding sequences in the human genome. Importantly, major differences have been observed between teleost fish and mammalian genomes. There is now convincing evidence that all teleosts are derived from a common tetraploid fish ancestor. This tetraploidization event arose about 320-350 million years ago in the ray-finned fish lineage, followed by rediploidization and retention of hundreds of duplicate pairs. Divergent evolution of the resulting duplicates has been proposed to be involved in the species richness observed in teleost fishes. Fish genomes also contain many more families of transposable elements than mammals and birds. Finally, while the mammalian and bird lineages possess major sex determination systems with sex chromosomes conserved in very divergent species, fishes have very frequently switched between sex determination mechanisms and repeatedly created novel sex chromosomes during evolution. Hence, teleost fishes display a high level of genomic plasticity, which might be related to the astonishing biodiversity observed in these animals.

Key Words: Fish; Teleost; Vertebrate; Evolution; Genome; Duplication; Sex chromosome; Retroelement.

WHY TO SEQUENCE A FISH GENOME

There are many good reasons to sequence a fish genome. For example, fish genomes can help to understand numerous biological processes in human and their dysfunction in disease. This is particularly true in the domain of developmental biology. Even if teleost fishes and humans are separated by approximately 450 million years of evolution, they both belong to the vertebrate lineage, and should therefore have important developmental pathways in common. Accordingly, some aquarium fish species like the zebrafish *Danio rerio* and the medaka *Oryzias latipes* are established as important complementary models for the study of vertebrate development. The sequence of their genomes should highly facilitate the characterization of the plethora of mutants generated during different large-scale mutagenesis programs, and will certainly shed a new light on development and disease in fishes and humans [1-4]. The zebrafish is increasingly used to study the function of genes involved in human diseases [5-8]. Fishes of the genus *Xiphophorus* are traditional models for cancer research [9].

The pufferfishes *Takifugu rubripes* (Fugu) and *Tetraodon nigroviridis* are pure models of comparative genomics. Their compact genomes principally provide an evolutionary counterpart to the human genome allowing the identification of conserved coding and regulatory sequences [10-12]. Unfortunately, both pufferfishes are extremely difficult to bread in the laboratory, this prohibiting their use for functional analyses.

Other teleost fishes like the East African cichlids and the three-spined stickleback *Gasterosteus aculeatus* are important models to analyse many fundamental questions in evolution and ecology [13-17]. Analysis of their genomes will certainly help to better understand the molecular basis of adaptation and speciation in fish. Various teleost species including the medaka, the platyfish, the three-spined stickleback, the Nile tilapia *Oreochromis niloticus* and several salmonids are used to study sex determination, which is extremely variable in fish [18-19]. The sequencing of the sex chromosomes in these fishes will certainly reveal new sex-determining genes and mechanisms able to drive sexual dimorphism in vertebrates [20]. Applications of fish genomics are multiple and include the study of the interactions between organisms and environment and the analysis of the response to natural or anthropogenic modifications of the biotope (environmental genomics [21]).

Importantly, fish is a vital source of food for people, particularly in Africa and Asia where livestock is relatively scarce, and has a substantial social and economic importance in many countries. Consequently, several genome projects have an economical basis. Such genomic programs, involving species like the Atlantic salmon *Salmo salar*, the rainbow trout *Oncorhynchus mykiss*, the Nile tilapia and the channel catfish *Ictalurus punctatus*, essentially aim to identify at the molecular level qualitative and quantitative trait loci (QTLs) controlling among others growth, reproduction, environ-

*Address correspondence to this author at the Physiologische Chemie I, University of Würzburg, Am Hubland, D-97074 Würzburg, Germany; Tel: +49 (0)931 888 4167; Fax: +49 (0)931 888 4150;
E-mail: volff@biozentrum.uni-wuerzburg.de

1389-2029/06 $50.00+.00 **©2006 Bentham Science Publishers Ltd.**

mental tolerance or resistance to disease, which are relevant traits for the aquaculture.

Since ray-finned fishes constitute the sister group of the lineage containing humans and other tetrapods (Fig. **1**), these multiple genomic efforts on different fish species concomitantly provide independently of their respective motivation an important basis for comparative genomics within the vertebrate lineage and allow a better understanding of the evolution of our own genome. Furthermore, comparison of different piscine genomes are important to understand the molecular mechanisms driving biodiversity in teleost fishes, which represent about half of extant vertebrate species and display a high level of diversity principally affecting their morphology, ecology and behaviour [22].

FISH GENOME PROJECTS

Genome projects generate sequence data principally by clone-by-clone sequencing of large genomic inserts (typically BACs, for bacterial artificial chromosomes [23]), which can be mapped physically on chromosomes by fluorescent *in situ* hybridization (FISH), or by analysis of clones generated by whole genome shotgun (WGS). Both strategies are often combined for the analysis of large genomes [24]. From these data, gene sequences can be predicted, orthology to genes from other organisms can be assessed and syntenic regions (regions with conserved gene content) can be identified [25]. Of particular interest is the identification of new genes, gene families and regulatory sequences as well as the reconstruction of the evolutionary events that shaped genomes.

There is a continuously increasing flux of novel sequences from fish genomes into public databases, including data produced by several fish genome projects (for an overview on current fish genetic and genomic resources, see [26]). The National Center for Biotechnology Information currently lists 113 animal genome projects (http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi) comprising 41 mammals, the Western clawed frog *Xenopus tropicalis*, the chicken *Gallus gallus*, the little skate *Leucoraja erinacea* (cartilaginous fish) as well as three teleost fishes, the pufferfishes *Takifugu rubripes* (Fugu) and *Tetraodon nigroviridis* as well as the zebrafish *Danio rerio*. Added to this list should be the medaka *Oryzias latipes*, which is object of a very advanced project (http://biol1.bio.nagoya-u.ac.jp:8000/), as well as the three-spined stickleback *Gasterosteus aculeatus* (http://www.genome.gov/12512292). Genome projects are being or will be initiated very soon for other teleost species too, including catfish, salmonids and cichlids, for which substantial genetic and genomic resources including a plethora of expressed sequence tags (ESTs) are already available [26].

The genome of the pufferfish *T. rubripes* (Fugu) was the second vertebrate genome to be sequenced, after the human but prior to the mouse genome [27]. The Fugu sequence was recently followed by the draft of the genome of the related green spotted pufferfish *Tetraodon nigroviridis* [28]. Both species were chosen as models for comparative genomics because of their extremely compact genomes, which are with about 400 megabases approximately eight times smaller than the human genome. This relative compaction is mainly due to shorter introns (even if exceptional "giant" genes have been observed in the genome of Fugu [27]) and to a lower content of repetitive sequences, but not to a reduced gene content [10]. However, compact pufferfish genomes unexpectedly contain many more families of transposable elements than the larger mammalian and chicken genomes [29]. One major advantage of the *Tetraodon* genome draft is the physical anchoring of more than 60% of the genome assembly on chromosomes by FISH, providing a major resource for comparative genomics [28]. A genetic linkage map with 200 microsatellite markers has been recently established for Fugu, which will serve as a basis for the development of a physical map [30].

Comparison of predicted genes from *T. rubripes* and *T. nigroviridis* with the human genome led to the identification of approximately 1000 and 900 non-annotated human genes, respectively [27-28]. Furthermore, genomic comparisons of pufferfish *vs.* mammals allowed the detection of putative conserved *cis*-regulatory elements in vertebrates [12]. Twenty-three of 25 such highly conserved non-coding sequences found at the proximity of four genes showed significant enhancer activity in one or more tissues in zebrafish embryos [31].

Assemblies of the genomes of the medaka and zebrafish (http://www.ensembl.org/Danio_rerio/) are publicly available, and their analysis should be published in the next future. Despite problems apparently due to an important polymorphism between the different individual zebrafish genomes sequenced, sequence data produced by the zebrafish genome project, which was started in February 2001 and is based at the Sanger Institute, were already extremely useful to the scientific community in association with a dense genetic map [32]. In order to tackle the problem of polymorphism, the main strategy recently adopted by the Sanger Institute is the "clone-by-clone" sequencing of a BAC library generated from a single double haploid fish.

GENOME DUPLICATIONS AND VERTEBRATE EVOLUTION: OUR POLYPLOID ANCESTORS

It has been suggested that two genome duplications have occurred at the origin of vertebrates 500-800 million years ago (The 2R hypothesis; Fig. **1**) [33, 34 and references therein], but the number and timing of these events remain matter of debate [35]. The 2R hypothesis ("one-two-four" rule) is suggested by the fact that some genes or gene clusters are present as single copies in non-chordate invertebrates as well as in *Ciona* (urochordate) and *Amphioxus* (cephalochordate), but at three to four copies in mammals and birds. This situation is perfectly exemplified by the *Hox* gene clusters, which encode transcription factors determining the positional specification of the anterior-posterior axis [36 and references therein]. Other examples supporting two genome duplications include the major histocompatibility complex loci (MHC) as well as the fibroblast growth factor receptor (Fgfr) and the epidermal growth factor receptor (Egfr) gene families [34 and references therein; 37-38] (Fig. **2**). The first genome duplication event is postulated after the divergence of *Ciona* and *Amphioxus* from the vertebrate lineage but before the split between jawless and jawed vertebrates, the second after the divergence of jawless vertebrates but before the split between cartilaginous and bony vertebrates (Fig. **1**).

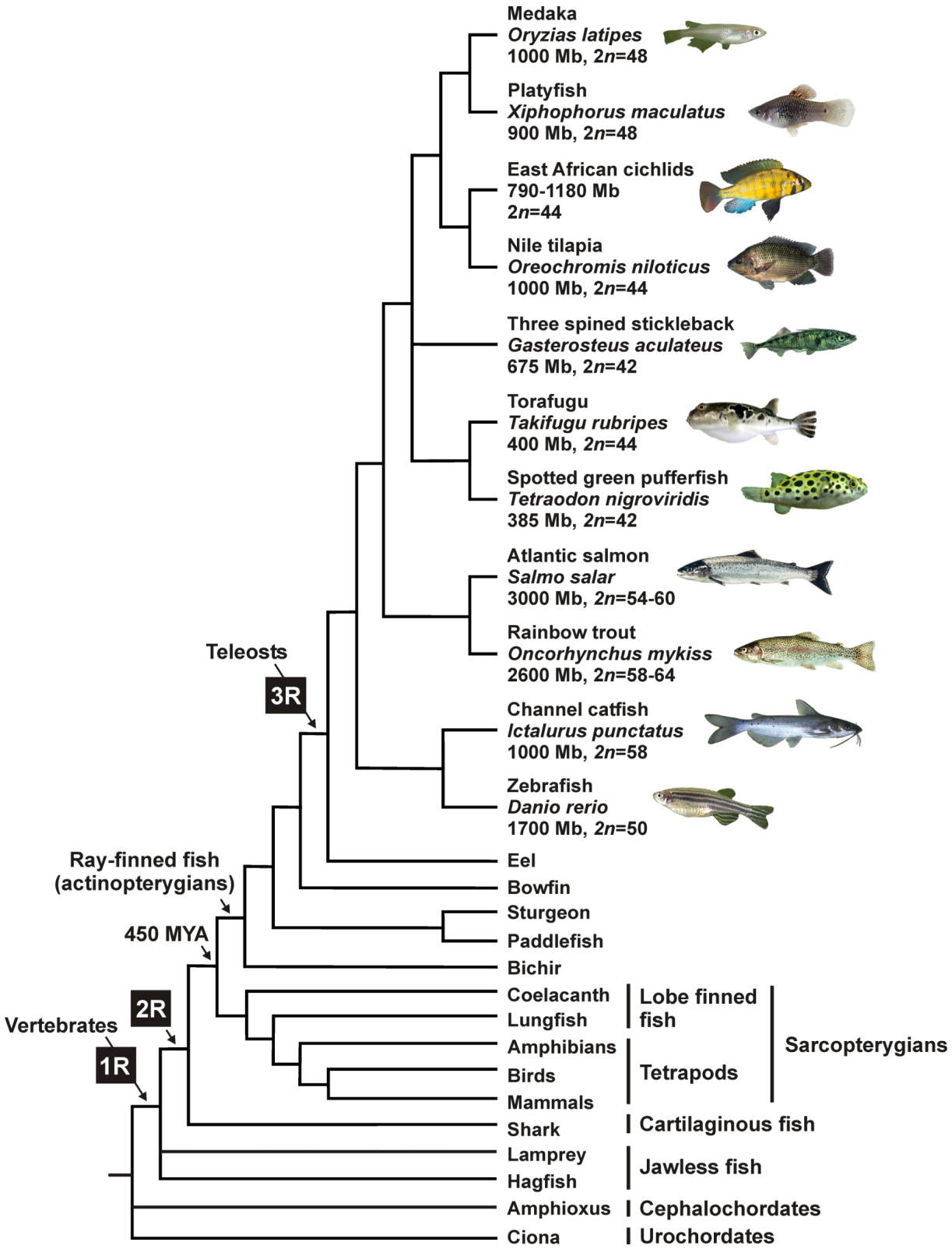

Fig. (1). Schematic view of the chordate lineage with particular emphasis on teleost fish models with substantial genetic and genomic resources. Haploid genome sizes (in megabases, Mb) as well as chromosome numbers are shown. The presumed timing of genome duplication events (1R, 2R, 3R) is shown. MYA, million years ago. Origin of fish pictures: Manfred Schartl and Christoph Winkler (medaka, zebrafish, platyfish and *Tetraodon*), Erwin Schraml (African cichlid), John F. Scarola (rainbow trout and Atlantic salmon), Suzanne L. and Joseph T. Collins (channel catfish), Bernd Ueberschaer (Nile tilapia), Konrad Schmidt (three-spined stickleback) and Greg Elgar (Torafugu).

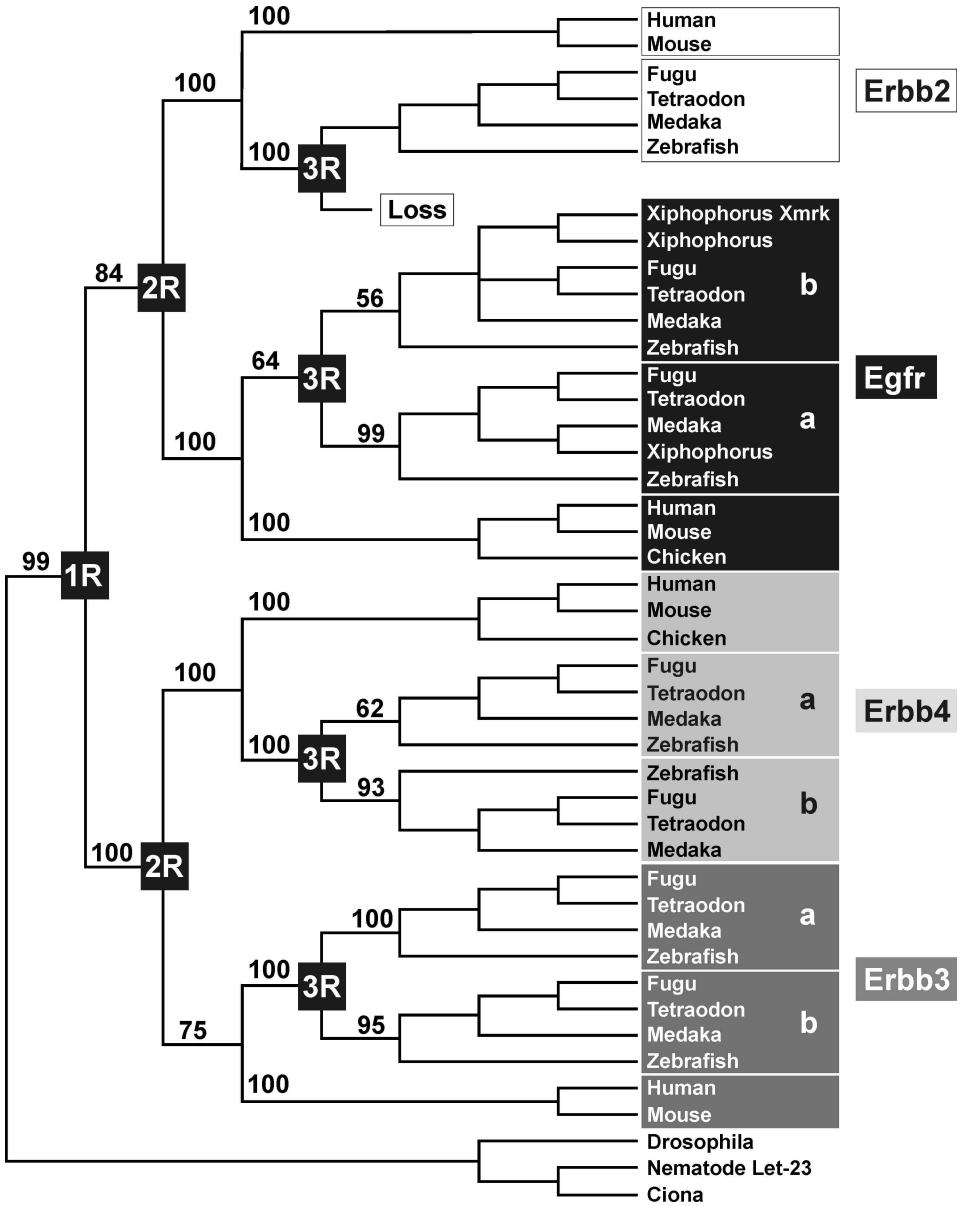

Fig. (2). Molecular phylogeny of epidermal growth factor receptor (Egfr)-related proteins in animals. The consensus tree was obtained from a 690 amino-acid sequence alignment using the neighbour-joining methods (1,000 pseudosamples) [118]; bootstrap values for major nodes are given as percentages. The presumed genome duplication events (1R, 2R, 3R) are shown. "Fish-specific" ancient gene duplicates are indicated by "a" and "b". Accession numbers: Erbb2: *Homo sapiens* NP_001005862, *Mus musculus* NP_034282, *Danio rerio* NM_200119, *Tetraodon nigroviridis* CAG12653; Egfrb: Xmrk *Xiphophorus maculatus* P13388, *Xiphophorus xiphidium* AAD10500, *Danio rerio* XP_700110; Egfra: *Tetraodon nigroviridis* CAG07098, *Xiphophorus xiphidium* AAP55673, *Danio rerio* NP_919405; Egfr: *Homo sapiens* NP_005219, *Mus musculus* NP_997538, *Gallus gallus* P13387; Erbb4: *Homo sapiens* NP_005226, *Mus musculus* XP_136682, chicken NP_001025536; Erbb4a: *Tetraodon nigroviridis* CAF98213; Erbb3 *Homo sapiens* NP_001973, *Mus musculus* NP_034283; Erbb3a *Danio rerio* NM_001005320; Erbb3b: *Danio rerio* NM_001014826; *Takifugu rubripes* AAC34391; *Tetraodon nigroviridis* CAG07406; Egfr from invertebrates: *Drosophila melanogaster* AAR85260, *Caenorhabditis elegans* Let-23 BAA09729, *Ciona intestinalis* BAE06394.

Other sequences are derived from data generated by genome projects of the zebrafish (http://www.ensembl.org/Danio_rerio/), medaka (http://dolphin.lab.nig.ac.jp/medaka/), Tetraodon (http://www.genoscope.cns.fr/externe/tetranew/) and Fugu (http://fugu.biology.qmul.ac.uk/).

Analysis of fish genomes has confirmed that *Hox* clusters and many other genes present as single copies in invertebrates are present at a higher copy number in divergent teleost species as observed in mammals, supporting the occurrence of ancient large-scale duplications in the early evolution of the vertebrate lineage. Interestingly, in a recent study including the complete gene sets of Fugu, *Ciona*, mouse and human, the plotting of the genomic map positions

of paralogous genes that were duplicated prior to the fish-tetrapod split revealed clear patterns of four-way paralogous regions covering a large part of the human genome [39]. This analysis, based in part on fish sequences, provides substantial support for two distinct genome duplication events early in vertebrate evolution.

THE "FISH-SPECIFIC" GENOME DUPLICATION(S)

Strikingly, the zebrafish not only possesses more *Hox* clusters than invertebrates, but also than mammals: seven *hox* gene clusters have been identified in *Danio rerio*, compared to only four in tetrapods. A least seven *Hox* clusters are also present in pufferfishes, medaka, salmonids and cichlids [36, 40-44]. These observations show that large ancient duplications have taken place during early evolution of the ray-finned fish lineage. As *Hox* clusters are generally good indicators of genome duplications, it has been suggested that extra clusters in teleost fishes are remnants of an ancient tetraploidization also called the "fish-specific" genome duplication (FSGD) (the 3R hypothesis [40, 45-48]). This hypothesis was validated by the identification of hundreds of ancient pairs of duplicates co-orthologous to mammalian single copy genes in various teleost genomes [28, 49-51] (Fig. **2**; Table **1**). Most of these paralogues are contemporaneous and have been formed in the ray-finned fish lineage approximately 320-350 million years ago [51-52]. Plotting these ancient "fish-specific" duplicates onto the genomes of zebrafish, *Tetraodon* and medaka revealed the presence of large duplicated regions present on different chromosomes in fishes but corresponding to unique regions in mammals; all teleost chromosomes contain such large paralogous regions [28, 32, 53]. These large duplicated segments are located for example on chromosomes 7/25, 16/19 and 17/20 in zebrafish (Table **1**). Duplicated regions have generally an orthologous counterpart in other teleost species: paralogous segments on chromosomes 7 and 25 of the zebrafish are orthologous to regions on chromosomes 5 and 13 of *Tetraodon*, respectively (Table **1**). Taken together, these observations provide strong evidence for an ancient genome duplication having occurred during the evolution of the ray-finned fish lineage. This event probably took place after the divergence of bowfin, paddlefish, sturgeon and bichir from the stem lineage leading to teleosts [54-55]. Hence, the "fish specific" genome duplication might be coincident with the origin of teleosts. If this tetraploidization arose within a same species (autotetraploidy) or through interspecific hybridization (allotetraploidy) remains to be determined.

More recent tetraploidization (and even octoploidization) events have occurred independently through different mechanisms in several lineages of non-teleost (for example in sturgeons; [56]) and teleost fishes, with various stages of rediploidization [57-58]. The entire family Salmonidae, which includes salmon and trout, descends from an autotetraploid ancestor that underwent genome duplication 25-100 million years ago. Here again, *Hox* clusters are reliable markers for genome duplication: at least 13 *Hox* clusters have been identified in salmonids, compared to 7-8 clusters in diploid teleost species [43]. The common carp *Cyprinus carpio* is a more recent tetraploid; its genome has been duplicated about 12 million years ago possibly by hybridization (allotetraploidy) [59].

Finally, as observed in mammals and other organisms, duplications of more restricted chromosomal regions are also playing an important role in the evolution of gene function and gene families in fish. This is well illustrated by the *claudin* gene family, which comprises 56 members in Fugu compared to 19 genes in human. *Claudin* expansion in teleosts resulted not only from the ancestral "fish-specific" genome duplication but also from multiple tandem duplications [60].

THE EVOLUTIONARY FATE OF DUPLICATED GENES – INSIGHTS FROM FISH

Genetic redundancy probably creates a favourable transient framework for evolutionary innovation, and it has been proposed accordingly that genome duplication events have been associated with major evolutionary transitions [33]. Because of the occurrence of an ancestral tetraploidization in the ray-finned fish lineage, with hundreds of duplicate gene pairs having been conserved in numerous species after rediploidization, teleost fishes represent an excellent model to analyse the consequences of genome duplication on gene function, organismal complexity and organism/species diversity. In addition, the availability of more recent tetraploidized species with variable degrees of rediploidization in the fish lineage might provide important insights into the early steps of this important evolutionary process.

After tetraploidization, each of the gene copies can follow very different evolutionary fates [47 and references therein]. One of the duplicates can be inactivated through mutations into a pseudogene, which will be discarded eventually from the genome. This process called nonfunctionalization apparently removed one copy from about 90% of the duplicate pairs generated by the "fish-specific" genome duplication [28, 32].

On the other hand, more than 2000 pairs of ancient gene duplicates have been maintained in teleost fish genomes [32]. Different mutually non-exclusive major mechanisms have been proposed to explain this persistence [47, 61]. The increase of expression conferred by the presence of two copies might be beneficial for the organism, leading to the maintenance of both duplicates. Duplicated genes can also experience functional divergence. One duplicate might acquire a novel, positively selected function (neofunctionalization). Alternatively, functions of the ancestral single-copy gene might be distributed between the duplicates, making the presence of both copies necessary to fulfil the original functions (subfunction partitioning). If this phenomenon is responsible for the preservation of the duplicates, it is called subfunctionalization (the duplication-degeneration-complementation model; [62, 63]). Reciprocal degenerative mutations in the regulatory sequences of duplicates can lead to quantitative, spatial or temporal subfunction partitioning [47]. Reciprocal mutations in the coding region of the duplicates affecting different functional protein domains can also lead to subfunction partitioning if these domains perform different functions in the original protein [64 and references therein]. Combinations of these models are of course possible: subfunctionalization of duplicated genes has been proposed to correspond to a transition state to neofunctionalization (the sub-neo-functionalization model; [65, 66]).

Table 1. Fifty Examples of Ancient "Fish-Specific" Gene Duplicate Pairs Orthologous to Single-Copy Genes in Mammals

Genes	Zebrafish	Medaka	Fugu	Tetraodon	Others	References
eng1b *eng1a*	Dre1 Dre9	+ +	+ nd	+ nd		[70]
dlx2b *dlx2a*	Dre1 Dre9	nd +	nd +	nd Tni2		[119]
dla(*dll1a*) *dld*(*dll1b*)	Dre1 Dre13	+ +	+ +	+ Tni5		[120]
msxb(*msx3a*) *msxc*(*msx3b*)	Dre1 Dre13	+ +	+ +	Tni17 Tni18		[121, 122]
kitb *kita*	Dre1 Dre20	+ +	+ +	Tni18 Tni1		[123, 124]
mdh1a *mdh1b*	Dre1 +	+ +	+ +	+ Tni17	Sid Sid	[125]
appa *appb*	Dre1 +	+ +	+ +	Tni3 Tni2		[126]
shhb (*twhh*) *shha*	Dre2 Dre7	nd +	nd +	nd Tni6	Cca, Ame	[127, 128]
eng2b *eng2a*	Dre2 Dre7	+ +	nd +	nd Tni6		[70, 129, 130]
fzd8b *fzd8a*	Dre2 Dre24	+ +	+ +	Tni15 +	Oni	[54, 131]
sox9b *sox9a*	Dre3 Dre12	LG8 LG19	+ +	Tni3 +	Gac, Mal Gac, Mal	[82, 83, 132]
hoxB5a *hoxB5b*	Dre3 Dre12	LG8 LG19	LG5 LG1	+ Tni2	Sne Sne	[30, 40, 42, 53]
timp2b *timp2a*	Dre3 Dre12	+ +	+ +	Tni3 Tni2		[73]
sox10b *sox10a*	Dre3 nd	+ +	+ +	+ Tni18		[133-135]
cntn1b *cntn1a*	Dre4 Dre25	+ +	+ +	Tni19 Tni13		[74]
insa *insb*	Dre5 Dre14	+ +	+ +	Tni7 Tni1	Cca, Oni, etc	[136]
notchb *notcha*	Dre5 Dre21	+ +	+ +	Tni12 Tni4		[137]
mitfa *mitfb*	Dre6 Dre13	+ +	+ +	Tni11 Tni9	Xma Xma	[68, 69]
erbb3a *erbb3b*	Dre6 Dre23	+ +	+ +	+ Tni9		[138]
igf1rb *igf1ra*	Dre7 Dre18	+ +	+ +	+ Tni13	Cca, Pol, etc Cca, Pol, etc	[139]
mdka *mdkb*	Dre7 Dre25	nd +	nd +	nd Tni13	Omy Omy	[140]

(Table 1) contd….

Genes	Zebrafish	Medaka	Fugu	Tetraodon	Others	References
isl3(*isl2b*) *isl2*(*isl2a*)	Dre7 Dre25	+ +	+ +	Tni5 Tni13	Ots Ots	[141]
pax6b *pax6a*	Dre7 Dre25	+ +	+ +	Tni5 +		[142]
crabp1b *crabp1a*	Dre7 Dre25	+ +	+ +	Tni5 Tni13		[75]
irx5a *irx5b*	Dre7 Dre25	+ +	+ +	Tni5 +		[143]
hoxD9a *hoxD9b*	Dre9 nd	LG21 LG15	LG1 +	Tni2 Tni17	Omy, Gac	[30, 40, 42, 43, 53]
hoxC6b *hoxC6a*	Dre11 Dre23	nd LG7	nd LG3	nd Tni9	Omy Sne	[30, 40, 42, 43, 53]
snai1a *snai1b*	Dre11 Dre23	+ +	+ +	Tni11 Tni9		[144, 145]
ptenb *ptena*	Dre12 Dre17	+ +	+ +	Tni17 +		[146, 147]
slit1a *slit1b*	Dre13 Dre22	+ +	+ +	Tni17 +		[148]
egfrb *egfra*	Dre14 +	+ +	+ +	+ Tni15	Xph, Xma Xma	[38, 149, 150]
gra *grb*	Dre14 nd	+ +	+ +	Tni1 Tni7	Omy, Abu Omy, Abu	[151, 152]
hoxA9b *hoxA9a*	Dre16 Dre19	LG16 LG11	LG7 LG12	Tni16 Tni21	Sne Sne	[30, 40, 42, 53]
mc5rb *mc5ra*	Dre16 Dre19	nd +	nd +	nd Tni8	 Omy, Cca	[153]
irx1a	Dre16 Dre19	+ +	+ +	+ Tni21		[143]
tpi1b *tpi1a*	Dre16 Dre19	+ +	+ nd	+ nd	Xma Xma	[154]
sox4b *sox4a*	Dre16 Dre19	+ +	+ +	Tni8 Tni21		[134, 155]
crabp2a *crabp2b*	Dre16 Dre19	+ +	+ +	Tni8 Tni21		[156]
flot1b *flot1a*	Dre16 +	+ +	+ +	Tni8 +	Cau	[157]
sox11a *sox11b*	Dre17 Dre20	+ nd	+ +	Tni10 Tni8	Oni Omy	[54, 71, 158]
snap25b *snap25a*	Dre17 Dre20	+ +	+ +	+ Tni14		[159]
bmp2a *bmp2b*	Dre17 Dre20	nd +	nd +	nd Tni14	 Pol	[160]

(Table 1) contd….

Genes	Zebrafish	Medaka	Fugu	Tetraodon	Others	References
runx2a	Dre17	nd	nd	nd		[161]
runx2b	Dre20	+	+	Tni14		
pomca(_)	Dre17	+	+	+	Omy, Oni, etc	[67]
pomcb(_)	Dre20	+	+	Tni14	Omy, etc	
cyp19a1a	Dre18	+	+	Tni5	Oni, Omy	[162, 163]
cyp19a1b	Dre25	+	+	Tni13	Oni, Omy	
spon1a	Dre18	+	+	Tni5		[164]
spon1b	Dre25	+	+	+		
hey1b	Dre19	+	+	Tni8		[165]
hey1a	nd	+	+	Tni21		
syn2b	Dre25	+	+	Tni9		[73]
syn2a	nd	+	+	Tni11		
wt1a	Dre25	+	+	+	Omy	[166]
wt1b	+	+	+	Tni5	Omy	
erbb4a	+	+	+	Tni2		[38]
erbb4b	+	+	+	Tni3		

Abu, *Astatotilapia burtoni*; Ame, *Astyanax mexicanus* (Mexican tetra); Cau, *Carassius auratus* (goldfish); Cca, *Cyprinus carpio* (common carp); Dre, chromosome of the zebrafish *Danio rerio*; Gac, *Gasterosteus aculeatus* (three-spined stickleback); LG, linkage group; Mal, *Monopterus albus* (swamp eel); nd, not detected; Omy, *Oncorhynchus mykiss* (rainbow trout); Oni, *Oreochromis niloticus* (Nile tilapia), Ots, *Oncorhynchus tshawytscha* (Chinook salmon); Pol, *Paralichthys olivaceus* (halibut); Sid, *Sphyraena idiastes* (pelican barracuda); Sne, *Sphoeroides nephelus* (pufferfish), Tni, chromosome of the pufferfish *Tetraodon nigroviridis*; Xph, *Xiphophorus xiphidium*; Xma, *Xiphophorus maculatus* (southern platyfish); +, present but not mapped.

Sequences and gene localization were obtained directly from ZFIN (http://zfin.org) or NCBI (http://www.ncbi.nlm.nih.gov). Sequences not characterized so far were identified through BLAST analysis of the genome of the zebrafish (http://www.ensembl.org/Danio_rerio), medaka (http://dolphin.lab.nig.ac.jp/medaka), Fugu (http://fugu.biology.qmul.ac.uk/blast/) and *Tetraodon* (http://www.ncbi.nlm.nih.gov/BLAST). Phylogenetic relationships were established as described in the legend of figure 2.

Subfunctionalization/ subfunction partitioning has been proposed to explain the persistence of numerous pairs of ancient fish duplicates. For example, the duplicated proopiomelanocortin genes have experienced spatial subfunction partitioning having led to complementary expression in brain and pituitary. In addition, partitioning was also observed at the protein sequence level, with a beta-endorphin segment being apparently functional in the translation product of one of the duplicates but degenerated in the product of the second one [67]. Two isoforms of the microphthalmia-associated transcription factor MITF with different functions, which are produced in mammals using alternative 5' exons and promoters from a single gene, are encoded by two distinct *mitf* genes in teleosts [68-69]. Subfunction partitioning between the *mitf* duplicates is associated with the reciprocal degeneration of isoform-specific alternative first exons and regulatory elements. Other putative examples of subfunction partitioning in teleosts include duplicates encoding Engrailed 2 homeoproteins [47, 70], Sox11 proteins [71], HoxB5 transcription factors [72], Synapsins [73], Cntn1 (Contactin/F3/F11) cell adhesion molecules [74], cellular retinoic acid-binding proteins type I [75], cytoglobins [76] and oestrogen receptors beta [77].

Possible examples of neofunctionalization have been also postulated for some duplicates generated by the “fish-specific” genome duplication [47], and in some cases positive selection was even detected in both duplicates [55]. However, neofunctionalization is generally difficult to demonstrate unambiguously, since functions and expression patterns of duplicates are generally compared with those of the orthologous gene in the mouse or in human. For example, the apparent apparition of a "new" expression pattern in one of the fish paralogues, which is absent for the corresponding gene in mammals, might be indicative of neofunctionalization. However, if this expression pattern was already present in the ancestral fish lineage before duplication, the mechanism shaping the evolution of the duplicates would be subfunction partitioning rather than neofunctionalization.

DIVERGENT PARANOME EVOLUTION IN FISH: A KEY TO BIODIVERSITY

Since the "fish specific" genome duplication predates the origin of teleosts [54, 55], a causal link has been proposed between this event and the huge species diversity observed in this group of fishes [26, 47, 48; but see 78]. With at least 24.000 species, teleost fishes are indeed the most diverse and successful vertebrate taxon [22]. Intuitively, massive duplication of genetic information should dramatically increase the global evolvability of a genome and strongly favour genetic innovation and diversity [55]. The reduction of the paranome itself (i.e. of the set of duplicated genes) might be linked to species formation: non-functionalization of different copies within pairs of paralogues in different populations (divergent resolution) might result in genomic incompatibil-

ity associated with reduced fertility and/or viability of hybrids. This might consequently favour speciation [26, 47, 79-81]. Alternatively, divergent subfunction partitioning of duplicate pairs in different populations would lead to a similar effect [26, 47, 79].

Accordingly, divergent evolution of the "fish-specific" paranome has been already observed in different teleost fish sublineages. It has been estimated that about 50% of genes duplicated in zebrafish are single-copy in pufferfish [32, 49]. However, this value might be an overestimation due to incomplete genome sequence datasets. Divergent resolution has been reported for some *hox* clusters. Zebrafish, medaka and *Tetraodon* all possess seven clusters. However, cluster *hoxDb*, which is present in pufferfish and medaka, is absent from zebrafish. Conversely, cluster *hoxCb*, which has been identified in zebrafish, has been apparently lost in a common ancestor of pufferfish and medaka. Divergent evolution of individual genes within *hox* clusters common to these species has been also observed [36, 42]. Other putative examples of differential loss of ancient duplicate genes in divergent teleost sublineages include genes encoding the Hey1a basic helix-loop-helix transcription factor, the tiggy-winkle hedgehog protein (formally sonic hedgehog b), the Sox10a transcription factor, the Mcr5b melanocortin receptor, the bone morphogenetic protein 2a and the runt-related transcription factor 2a (Table **1**). Finally, comparative analysis of *sox9a* and *sox9b* duplicates in zebrafish, stickleback and medaka has provided evidence for lineage-specific subfunction partitioning [82, 83]. However, for numerous pairs of duplicates, subfunctions have been partitioned early in the teleost lineage before the divergence of medaka and zebrafish.

RECONSTRUCTING THE ANCESTRAL BONY VERTEBRATE GENOME

With the genomes of several teleost fishes and mammals in hand, it was possible to reconstruct the karyotype of their last common ancestor that lived approximately 450 million years ago. Different studies based on high-quality genetic maps of the zebrafish [32, 84] and medaka [53] as well as on the genome sequence draft and physical map of the pufferfish *Tetraodon nigroviridis* [28] congruently concluded that the haploid genome of the ancestral bony vertebrate was constituted by 12 chromosomes. Subsequently, the haploid set of 20-30 chromosomes observed in mammals was achieved through multiple chromosome fissions, while the "fish-specific" genome duplication generated a set of chromosomes from which the 21-25 chromosomes of modern teleosts are derived. Consistently, the nodal value for the haploid number of chromosomes in teleosts is 24. Relatively few interchromosomal rearrangements have occurred in the teleost lineage. Local gene order was principally scrambled by intrachromosomal rearrangements. In contrast, the human genome has been extensively shuffled by interchromosomal rearrangements after its split from ray-finned fishes [28, 32].

DIFFERENTIAL EVOLUTION OF TRANSPOSABLE ELEMENTS IN FISH AND MAMMALS

Transposable elements are mobile DNA sequences able to integrate into new genomic sites within genomes [85-86]. They are not only able to disrupt genes and other sequences through insertion, but can also induce the formation of various types of genomic rearrangements including deletions, duplications, inversions and translocations principally through homologous recombination between non-allelic copies of a same element. Transposition itself can also produce different rearrangements at new insertion sites [85]. Transposable elements can modify the expression of flanking genes either through their own regulatory sequences, or by local modification of chromatin structure. Intronic insertions can also be recruited as exons and hereby disrupt the open reading frame of a gene at the mRNA level.

On the other hand, transposable elements have played a very important role in the evolution of genome structure and gene function in vertebrates and other organisms, and have generated at least half of human and mouse genomes [85-86]. A substantial fraction of regulatory and coding sequences in mammals are derived from transposable elements, and important cellular functions are encoded by domesticated retroelements and DNA transposons [87-90].

Transposable elements are separated in two major classes according to their structure and mechanism of transposition [91]. Essentially, distinction is made between retrotransposable elements (retroelements), which transpose through a mechanism involving the reverse transcription of an mRNA molecule (retrotransposition), and DNA transposons, which do not require reverse transcription for transposition.

Retrotransposable elements and DNA transposons are both present in mammalian and teleost fish genomes, with a much higher global copy number in mouse and human than in pufferfish or zebrafish [27-29, 85, 92]. However, many ancient groups of transposable elements present in invertebrates could be identified in teleost genomes but not in mammals and chicken [29]. This was particularly true for a very diverse group of endogenous retroelements called Ty3/Gypsy long terminal repeat (LTR) retrotransposons. These sequences encode among others a Gag structural protein as well as an integrase and a reverse transcriptase. Their coding region is flanked by two LTRs in direct orientation, and they are structurally and phylogenetically related to vertebrate retroviruses. As many as nine ancient phylogenetic groups of Ty3/Gypsy retrotransposons also present in a variety of invertebrate species were found in the genome of various teleost fish species. In contrast, none of them could be detected in mammalian genomes, with the exception of a family of Ty3/Gypsy retrotransposon-derived domesticated neogenes [90]. Other types of reverse transcriptase-encoding retrotransposable elements were identified in fishes but not in mammals, including Ty1/Copia LTR retrotransposons [29], BEL-like LTR retrotransposons [93], tyrosine recombinase-encoding LTR retrotransposons [94], Penelope-like elements [95-96] and non-LTR retrotransposons with restriction enzyme-like endonuclease [97-98]. Even for non-LTR retrotransposons with apurinic-apyrimidinic endonuclease, which constitute with more than 650.000 copies about 20% of the human and mouse genomes, more families were detected in teleost fishes than in mammals [29], but with much lower copy numbers.

Taken together, from a total of 26 ancient families of non-retroviral retrotransposable elements characterized in

both vertebrates and invertebrates, 22 families were identified in the genome of the zebrafish. In contrast, only three are present in human and mouse genomes. Even the very compact genomes of the smooth pufferfishes with their low repeat content show a higher diversity of endogenous retroelements than mammalian genomes: 22 and 15 families were detected in Fugu and *Tetraodon*, respectively. A similar situation is observed for DNA transposons [92, 99]. On the other hand, endogenous retroviruses, which were introduced independently through infection into the different vertebrate lineages, are present in both teleost fishes and tetrapods, as well as non-coding retroelements called SINEs (short interspersed nuclear elements, for example *Alu* sequences in humans). However, the copy number of SINEs is dramatically higher in mammals than in fishes.

Assuming a mode of vertical transmission, these observations indicate that the great majority of transposable element families present in the genome of the last common ancestor of tetrapods and ray-finned fishes have been eliminated before the split of human and mouse. Absence of most major retroelement families from the chicken genome but their presence in *Xenopus* (unpublished results) suggests that this "genomic purge" arose approximately 300-350 million years ago in the tetrapod lineage. The basis of this massive elimination remains unknown. "Higher" vertebrates might have developed new defence mechanisms not present in fishes and amphibians. Alternatively, very successful families of retrotransposons like L1 in mammals or CR2 in chicken might have supplanted other less efficient groups of transposable elements during evolution.

In contrast to the situation observed in the human genome, where the immense majority of mobile sequences have been inactivated through mutations, numerous families of transposable elements have been recently and are probably still active in different teleost fish genomes. However, their copy number is generally much lower in fishes than in mammals, suggesting a higher turnover in transposable elements in fishes [29].

Analysis by fluorescent *in situ* hybridization of the localization of various types of transposable elements in the compact genome of the pufferfish *Tetraodon nigroviridis* showed that these sequences are generally excluded from gene-rich regions. They rather accumulate together with other categories of repeats (duplicated pseudogenes, minisatellites) in particular heterochromatic regions of the genome [98, 100-101]. These observations demonstrated the extreme degree of compartmentalization of the pufferfish compact genome. Such a situation is not observed in humans, where repeated sequences constitute an important fraction of euchromatic DNA.

SEX CHROMOSOMES IN MAMMALS AND FISH: CONSERVATION VS. CREATIVITY

Mammals and birds possess relatively stable sex determination systems, with conserved sex chromosomes in divergent species of the same group [102]. In most mammals, males are XY and females XX (male heterogamety), with the Y-linked *Sry* gene inducing the male phenotype. In birds, males are ZZ and females ZW, with the Z-linked gene *dmrt1* as a candidate for the master sex-determining gene.

In teleost fishes, the picture is completely different. All possible forms of genetic sex determination have been observed in fishes, with variable influence of environmental factors including the temperature and the pH of the water [18-19, 103-104]. In addition, numerous fish species are hermaphrodites. Within a same species, different types of sex determination can coexist (for example genetic sex determination and influence of temperature in the Nile Tilapia). Different systems of genetic sex determination can be found in a same fish genus (e.g. in *Oreochromis spp.*) and even in a same species (e.g. in the platyfish). The molecular and evolutionary mechanisms driving sex determination and its variability in teleost fishes are poorly understood; mammalian Y chromosome and sex-determining gene *Sry* are absent from the fish lineage.

Recent analyses from different fish species suggest that the diversity of sex determination in fishes and the frequent switching between different sex-determining systems might be linked to an astonishing property of teleost genomes to create new sex chromosomes. This is illustrated by the best example studied to date, the Y chromosome of the medaka *Oryzias latipes*. *O. latipes* has a genetic sex determination with male heterogamety. Using different approaches, two groups have independently isolated the Y-linked master sex-determining gene of the medaka [105-106]. This gene, called *dmrt1bY* (aka *DMY*), is the first master sex-determining gene to be identified in a non-mammalian vertebrate species [19, 107-108]. The *dmrt1bY* gene is a Y-specific duplicate of the autosomal gene *dmrt1*, which encodes a putative transcription factor also involved in sex determination and/or differentiation in tetrapods. Consistent with a role as master sex-determining gene, natural mutations in *dmrt1bY* result in XY sex-reversed females [105].

Dmrt1bY has been formed though a large transchromosomal duplication from linkage group 9 onto another autosome, which consequently became the neo-Y-chromosome; all other genes included in this duplication have been subsequently inactivated by mutations [106]. Importantly, *dmrt1bY* and its autosomal counterpart *dmrt1* are very similar, suggesting that the duplication event that led to the formation of the master gene is relatively recent. Accordingly, it was estimated that *dmrt1bY* has been formed approximately 10 million years ago [109]. *Dmrt1bY* was detected in only a very restricted number of *Oryzias* species, and is absent from more divergent fishes [110-112].

Another recent sex chromosome in fish is the Y chromosome of the three-spined stickleback *Gasterosteus aculeatus*, which is probably less than 10 million years old [113]. Finally, different Y-chromosomes have apparently evolved in different salmonid species [114]. All fish sex-determination regions analyzed so far in fishes are very unstable, and this instability might be relevant for the molecular differentiation between different types of sex chromosomes, for the evolution of sex-linked genes and for the frequent switching between sex determination systems observed in fishes [26]. The evolutionary background behind the repeated formation of new chromosomes in fishes remains completely unknown.

CONCLUSION AND PERSPECTIVES

The analysis of teleost fish genomes has considerably improved our knowledge of the structure and evolution of vertebrate genomes, and has allowed approaching the genomic mechanisms associated with major evolutionary transitions during vertebrate evolution. As a basis for comparative genomics, teleost fishes have significantly contributed to a better understanding of the functioning of the human genome. Interestingly, these analyses have also revealed several remarkable features of fish genomes not observed in tetrapods. Fishes are therefore an outstanding model for the investigation of the molecular processes driving diversity and speciation in living organisms.

One major discovery of the era of fish genomics was the ancestral "fish-specific" genome duplication, which is responsible for the presence of hundreds to thousands of ancient duplicated genes in teleost genomes that are present as single copy genes in human and mouse. Taking this major difference between fishes and mammals into account is very important for disciplines using fishes as models for the investigation of biological processes in human. Due to phenomena like divergent resolution or divergent subfunction partitioning, a same gene might have different functions in two different fishes, or a same function might be performed by different paralogues. Hence, divergent fish models, like zebrafish and medaka in the field of developmental biology [3], should be compared before drawing definite conclusions concerning the function of a given gene. Differential evolution of gene pairs involved in embryonic development in medaka and zebrafish might explain why certain mutants have been obtained in one species but not in the other. On the other hand, partitioning of ancestral functions performed by a same gene between two duplicates will allow the separate study of these functions in fishes. This might be of importance for a better characterization of multifunctional genes in humans, particularly in relation with disease [37, 47].

Much more work will be required to understand the evolutionary processes responsible for the persistence of hundreds of ancient duplicated genes in more than 20,000 different fish species. The development of suitable outgroups for comparison of gene expression and function is certainly one of the major challenges for the future. The ideal pendant for the teleost lineage would be a ray-finned fish having diverged from the stem lineage leading to teleosts before the "fish-specific" genome duplication (for example a sturgeon). Alternatively, teleost fishes should be compared not only with mammals but also with cartilaginous fishes (for example sharks) in order to assess the function and expression of the gene studied before the split of tetrapods and ray-finned fishes. Such comparisons have been already performed at the sequence level for some Hox genes [115, 116]. The chimera *Callorhinchus milli* (elephant fish) has been recently proposed as a cartilaginous fish model for comparative genomics because of the relatively small size of its genome [117]. The fact that two genome duplication events have probably occurred between non-vertebrate chordates and bony vertebrates strongly complicates comparisons with *Ciona* or *Amphioxus*, which also lack several important vertebrate-specific developmental features that might be important for comparative analysis.

Finally, many more different teleost sublineages will need to be tested to understand the evolutionary flexibility of the "fish-specific" paranome and the influence of its differential evolution on organismal and species diversity. The causal relationship between genome duplication and increase in phenotypic complexity, morphological innovation and taxon richness has been recently challenged through the consideration of extinct taxa [78]. In addition, other essential characteristics of teleost fish genomes, like the numerous families of active transposable elements or the turnover of sex chromosomes, might also play a role in the species richness and organismal diversity observed in teleosts. Clearly, the next era of comparative functional genomics in fishes will provide further insights into the general mechanisms driving diversity and speciation.

ACKNOWLEDGEMENTS

The Volff lab is supported by grants from the Bundesministerium für Bildung und Forschung (BMBF, Biofuture programme) and from the Deutsche Forschungsgemeinschaft (DFG). We thank Manfred Schartl and Christoph Winkler for stimulating discussions.

REFERENCES

[1] Granato, M. and Nüsslein-Volhard, C.: Fishing for genes controlling development. *Curr Opin Genet Dev*, **1996,** *6*, 461-468.

[2] Shima, A., Himmelbauer, H., Mitani, H., Furutani-Seiki, M., Wittbrodt, J. and Schartl, M.: Fish genomes flying. Symposium on medaka genomics. *EMBO Rep*, **2003,** *4*, 121-125.

[3] Furutani-Seiki, M. and Wittbrodt, J.: Medaka and zebrafish, an evolutionary twin study. *Mech Dev*, **2004,** *121*, 629-637.

[4] Naruse, K., Hori, H., Shimizu, N., Kohara, Y. and Takeda, H.: Medaka genomics: a bridge between mutant phenotype and gene function. *Mech Dev*, **2004,** *121*, 619-628.

[5] Brownlie, A., Donovan, A., Pratt, S.J., Paw, B.H., Oates, A.C., Brugnara, C., Witkowska, H.E., Sassa, S. and Zon, L.I.: Positional cloning of the zebrafish *sauternes* gene: a model for congenital sideroblastic anaemia. *Nat Genet*, **1998,** *20*, 244-250.

[6] Taylor, M.R., Hurley, J.B., Van Epps, H.A. and Brockerhoff, S.E.: A zebrafish model for pyruvate dehydrogenase deficiency: rescue of neurological dysfunction and embryonic lethality using a ketogenic diet. *Proc Natl Acad Sci U S A*, **2004,** *101*, 4584-4589.

[7] Sadler, K.C., Amsterdam, A., Soroka, C., Boyer, J. and Hopkins, N.: A genetic screen in zebrafish identifies the mutants vps18, nf2 and foie gras as models of liver disease. *Development*, **2005,** *132*, 3561-3572.

[8] Winkler, C., Eggert, C., Gradl, D., Meister, G., Giegerich, M., Wedlich, D., Laggerbauer, B. and Fischer, U.: Reduced U snRNP assembly causes motor axon degeneration in an animal model for spinal muscular atrophy. *Genes Dev*, **2005,** *19*, 2320-2330.

[9] Meierjohann, S., Schartl, M. and Volff, J.-N.: Genetic, biochemical and evolutionary facets of *Xmrk*-induced melanoma formation in the fish *Xiphophorus*. *Comp Biochem Physiol C Toxicol Pharmacol*, **2004,** *138*, 281-289.

[10] Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B. and Aparicio, S.: Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome. *Nature*, **1993,** *366*, 265-268.

[11] Roest Crollius, H., Jaillon, O., Bernot, A., *et al.*: Estimate of human gene number provided by genome-wide analysis using *Tetraodon nigroviridis* DNA sequence. *Nat Genet*, **2000,** *25*, 235-238.

[12] Venkatesh, B. and Yap, W.H.: Comparative genomics using fugu: a tool for the identification of conserved vertebrate cis-regulatory elements. *Bioessays*, **2005,** *27*, 100-107.

[13] McKinnon, J.S. and Rundle, H.D.: Speciation in nature: the threespine stickleback model systems. *Trends in Ecology and Evolution*, **2002,** *17*, 480-488.

[14] Kocher, T.D.: Adaptive evolution and explosive speciation: the cichlid fish model. *Nat Rev Genet*, **2004,** *5*, 288-298.

[15] Salzburger, W. and Meyer, A.: The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. *Naturwissenschaften*, **2004,** *91*, 277-290.

[16] Tickle, C. and Cole, N.J.: Morphological diversity: taking the spine out of three-spine stickleback. *Curr Biol*, **2004,** *14*, R422-424.

[17] Peichel, C.L.: Fishing for the secrets of vertebrate evolution in threespine sticklebacks. *Dev Dyn*, **2005,** *234*, 815-823.

[18] Devlin, R.H. and Nagahama, Y.: Sex determination and differentiation in fish: An overview of genetic, physiological, and environmental influences. *Aquaculture*, **2002,** *208*, 191-364.

[19] Volff, J.-N. and Schartl, M.: Sex determination and sex chromosome evolution in the medaka, *Oryzias latipes*, and the platyfish, *Xiphophorus maculatus*. *Cytogenet Genome Res*, **2002,** *99*, 170-177.

[20] Volff, J.-N., Kondo, M. and Schartl, M.: Medaka *dmY/dmrt1Y* is not the universal primary sex-determining gene in fish. *Trends Genet*, **2003,** *19*, 196-199.

[21] Cossins, A.R. and Crawford, D.L.: Fish as models for environmental genomics. *Nat Rev Genet*, **2005,** *6*, 324-333.

[22] Nelson, J.S. (1994). Fishes of the world. New York: John Wiley & Sons.

[23] Miyake, T. and Amemiya, C.T.: BAC libraries and comparative genomics of aquatic chordate species. *Comp Biochem Physiol C Toxicol Pharmacol*, **2004,** *138*, 233-244.

[24] Green, E.D.: Strategies for the systematic sequencing of complex genomes. *Nat Rev Genet*, **2001,** *2*, 573-583.

[25] Thomas, J.W. and Touchman, J.W.: Vertebrate genome sequencing: building a backbone for comparative genomics. *Trends Genet*, **2002,** *18*, 104-108.

[26] Volff, J.-N.: Genome evolution and biodiversity in teleost fish. *Heredity*, **2005,** *94*, 280-294.

[27] Aparicio, S., Chapman, J., Stupka, E., *et al.*: Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science*, **2002,** *297*, 1301-1310.

[28] Jaillon, O., Aury, J.M., Brunet, F., *et al.*: Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature*, **2004,** *431*, 946-957.

[29] Volff, J.-N., Bouneau, L., Ozouf-Costaz, C. and Fischer, C.: Diversity of retrotransposable elements in compact pufferfish genomes. *Trends Genet*, **2003,** *19*, 674-678.

[30] Kai, W., Kikuchi, K., Fujita, M., Suetake, H., Fujiwara, A., Yoshiura, Y., Ototake, M., Venkatesh, B., Miyaki, K. and Suzuki, Y.: A genetic linkage map for the tiger pufferfish, *Takifugu rubripes*. *Genetics*, **2005,** *171*, 227-238.

[31] Woolfe, A., Goodson, M., Goode, D.K., *et al.*: Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol*, **2005,** *3*, e7.

[32] Woods, I.G., Wilson, C., Friedlander, B., Chang, P., Reyes, D.K., Nix, R., Kelly, P.D., Chu, F., Postlethwait, J.H. and Talbot, W.S.: The zebrafish gene map defines ancestral vertebrate chromosomes. *Genome Res*, **2005,** *15*, 1307-1314.

[33] Ohno, S. (1970). Evolution by gene duplication. New York: Springer Verlag.

[34] Panopoulou, G. and Poustka, A.J.: Timing and mechanism of ancient vertebrate genome duplications -- the adventure of a hypothesis. *Trends Genet*, **2005,** *21*, 559-567.

[35] Skrabanek, L. and Wolfe, K.H.: Eukaryote genome duplication - where's the evidence? *Curr Opin Genet Dev*, **1998,** *8*, 694-700.

[36] Hoegg, S. and Meyer, A.: *Hox* clusters as models for vertebrate genome evolution. *Trends Genet*, **2005,** *21*, 421-424.

[37] Volff, J.-N. and Schartl, M.: Evolution of signal transduction by gene and genome duplication in fish. *J Struct Funct Genomics*, **2003,** *3*, 139-150.

[38] Gomez, A., Volff, J.-N., Hornung, U., Schartl, M. and Wellbrock, C.: Identification of a second *egfr* gene in *Xiphophorus* uncovers an expansion of the epidermal growth factor receptor family in fish. *Mol Biol Evol*, **2004,** *21*, 266-275.

[39] Dehal, P. and Boore, J.L.: Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol*, **2005,** *3*, e314.

[40] Amores, A., Force, A., Yan, Y.L., *et al.*: Zebrafish *hox* clusters and vertebrate genome evolution. *Science*, **1998,** *282*, 1711-1714.

[41] Malaga-Trillo, E. and Meyer, A.: Genome duplications and accelerated evolution of *Hox* genes and cluster architecture in teleost fishes. *Amer Zool*, **2001,** *41*, 676-686.

[42] Amores, A., Suzuki, T., Yan, Y.L., Pomeroy, J., Singer, A., Amemiya, C. and Postlethwait, J.H.: Developmental roles of pufferfish *Hox* clusters and genome evolution in ray-fin fish. *Genome Res*, **2004,** *14*, 1-10.

[43] Moghadam, H.K., Ferguson, M.M. and Danzmann, R.G.: Evolution of *Hox* Clusters in Salmonidae: A Comparative Analysis Between Atlantic Salmon (*Salmo salar*) and Rainbow Trout (*Oncorhynchus mykiss*). *J Mol Evol*, **2005,** *61*, 636-649.

[44] Santini, S. and Bernardi, G.: Organization and base composition of tilapia *Hox* genes: implications for the evolution of *Hox* clusters in fish. *Gene*, **2005,** *346*, 51-61.

[45] Wittbrodt, J., Meyer, A. and Schartl, M.: More genes in fish? *Bioessays*, **1998,** *20*, 511-515.

[46] Meyer, A. and Schartl, M.: Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol*, **1999,** *11*, 699-704.

[47] Postlethwait, J., Amores, A., Cresko, W., Singer, A. and Yan, Y.L.: Subfunction partitioning, the teleost radiation and the annotation of the human genome. *Trends Genet*, **2004,** *20*, 481-490.

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

[49] Taylor, J.S., Braasch, I., Frickey, T., Meyer, A. and Van De Peer, Y.: Genome duplication, a trait shared by 22,000 species of ray-finned fish. *Genome Res*, **2003,** *13*, 382-390.

[50] Christoffels, A., Koh, E.G., Chia, J.M., Brenner, S., Aparicio, S. and 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.

[51] Vandepoele, K., De Vos, W., Taylor, J.S., Meyer, A. and 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.

[52] Taylor, J.S., Van de Peer, Y., Braasch, I. and 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.

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

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

[55] Crow, K.D., Stadler, P.F., Lynch, V.J., Amemiya, C. and Wagner, G.P.: The "Fish Specific" *Hox* Cluster Duplication is Coincident with the Origin of Teleosts. *Mol Biol Evol*, **2005**.

[56] Ludwig, A., Belfiore, N.M., Pitra, C., Svirsky, V. and Jenneckens, I.: Genome duplication events and functional reduction of ploidy levels in sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). *Genetics*, **2001,** *158*, 1203-1215.

[57] Venkatesh, B.: Evolution and diversity of fish genomes. *Curr Opin Genet Dev*, **2003,** *13*, 588-592.

[58] Gregory, R.T. and Mable, B.K. (2005). Polyploidy in animals. In: The evolution of the genome, R.T. Gregory, ed. (London: Elsevier Academic Press), pp. 427-517.

[59] David, L., Blum, S., Feldman, M.W., Lavi, U. and Hillel, J.: Recent Duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci. *Mol Biol Evol*, **2003,** *20*, 1425-1434.

[60] Loh, Y.H., Christoffels, A., Brenner, S., Hunziker, W. and Venkatesh, B.: Extensive expansion of the claudin gene family in the teleost fish, *Fugu rubripes*. *Genome Res*, **2004,** *14*, 1248-1257.

[61] Prince, V.E. and Pickett, F.B.: Splitting pairs: the diverging fates of duplicated genes. *Nat Rev Genet*, **2002,** *3*, 827-837.

[62] Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L. and Postlethwait, J.: Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, **1999,** *151*, 1531-1545.

[63] Lynch, M. and Force, A.: The probability of duplicate gene preservation by subfunctionalization. *Genetics*, **2000,** *154*, 459-473.

[64] Hughes, A.L.: Adaptive evolution after gene duplication. *Trends Genet*, **2002,** *18*, 433-434.

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

[66] Rastogi, S. and Liberles, D.A.: Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol Biol*, **2005,** *5*, 28.

[67] de Souza, F.S., Bumaschny, V.F., Low, M.J. and Rubinstein, M.: Subfunctionalization of expression and peptide domains following the ancient duplication of the proopiomelanocortin gene in teleost fishes. *Mol Biol Evol*, **2005**.

[68] Lister, J.A., Close, J. and Raible, D.W.: Duplicate *mitf* genes in zebrafish: complementary expression and conservation of melanogenic potential. *Dev Biol*, **2001,** *237*, 333-344.

[69] Altschmied, J., Delfgaauw, J., Wilde, B., Duschl, J., Bouneau, L., Volff, J.-N. and Schartl, M.: Subfunctionalization of duplicate *mitf* genes associated with differential degeneration of alternative exons in fish. *Genetics*, **2002,** *161*, 259-267.

[70] Ekker, M., Wegner, J., Akimenko, M.A. and Westerfield, M.: Coordinate embryonic expression of three zebrafish *engrailed* genes. *Development*, **1992,** *116*, 1001-1010.

[71] de Martino, S., Yan, Y.L., Jowett, T., Postlethwait, J.H., Varga, Z.M., Ashworth, A. and Austin, C.A.: Expression of *sox11* gene duplicates in zebrafish suggests the reciprocal loss of ancestral gene expression patterns in development. *Dev Dyn*, **2000,** *217*, 279-292.

[72] Bruce, A.E., Oates, A.C., Prince, V.E. and Ho, R.K.: Additional *hox* clusters in the zebrafish: divergent expression patterns belie equivalent activities of duplicate *hoxB5* genes. *Evol Dev*, **2001,** *3*, 127-144.

[73] Yu, W.P., Brenner, S. and Venkatesh, B.: Duplication, degeneration and subfunctionalization of the nested *synapsin-Timp* genes in Fugu. *Trends Genet*, **2003,** *19*, 180-183.

[74] Haenisch, C., Diekmann, H., Klinger, M., Gennarini, G., Kuwada, J.Y. and Stuermer, C.A.: The neuronal growth and regeneration associated *Cntn1* (F3/F11/Contactin) gene is duplicated in fish: expression during development and retinal axon regeneration. *Mol Cell Neurosci*, **2005,** *28*, 361-374.

[75] Liu, R.Z., Sharma, M.K., Sun, Q., Thisse, C., Thisse, B., Denovan-Wright, E.M. and Wright, J.M.: Retention of the duplicated cellular retinoic acid-binding protein 1 genes (*crabp1a* and *crabp1b*) in the zebrafish genome by subfunctionalization of tissue-specific expression. *Febs J*, **2005,** *272*, 3561-3571.

[76] Fuchs, C., Luckhardt, A., Gerlach, F., Burmester, T. and Hankeln, T.: Duplicated cytoglobin genes in teleost fishes. *Biochem Biophys Res Commun*, **2005,** *337*, 216-223.

[77] Hawkins, M.B., Godwin, J., Crews, D. and Thomas, P.: The distributions of the duplicate oestrogen receptors ER-beta a and ER-beta b in the forebrain of the Atlantic croaker (*Micropogonias undulatus*): evidence for subfunctionalization after gene duplication. *Proc Biol Sci*, **2005,** *272*, 633-641.

[78] Donoghue, P.C.J. and Purnell, M.A.: Genome duplication, extinction and vertebrate evolution. *Trends Ecol Evol*, **2005,** *20*, 312-319.

[79] Lynch, M. and Conery, J.S.: The evolutionary fate and consequences of duplicate genes. *Science*, **2000,** *290*, 1151-1155.

[80] Lynch, M. and Force, A.: The origin of interspecific genomic incompability *via* gene duplication. *Am Naturalist*, **2000,** *17*, 299-301.

[81] Taylor, J.S., Van de Peer, Y. and Meyer, A.: Genome duplication, divergent resolution and speciation. *Trends Genet*, **2001,** *17*, 299-301.

[82] Cresko, W.A., Yan, Y.L., Baltrus, D.A., Amores, A., Singer, A., Rodriguez-Mari, A. and Postlethwait, J.H.: Genome duplication, subfunction partitioning, and lineage divergence: *Sox9* in stickleback and zebrafish. *Dev Dyn*, **2003,** *228*, 480-489.

[83] Klüver, N., Kondo, M., Herpin, A., Mitani, H. and Schartl, M.: Divergent expression patterns of *Sox9* duplicates in teleosts indicate a lineage specific subfunctionalization. *Dev Genes Evol*, **2005,** *215*, 297-305.

[84] Postlethwait, J.H., Woods, I.G., Ngo-Hazelett, P., Yan, Y.L., Kelly, P.D., Chu, F., Huang, H., Hill-Force, A. and Talbot, W.S.: Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Res*, **2000,** *10*, 1890-1902.

[85] Kazazian, H.H., Jr.: Mobile elements: drivers of genome evolution. *Science*, **2004,** *303*, 1626-1632.

[86] Volff, J.-N. (2005). Retrotransposable elements and genome evolution. Basel: Karger Verlag.

[87] Nekrutenko, A. and Li, W.H.: Transposable elements are found in a large number of human protein-coding genes. *Trends Genet*, **2001,** *17*, 619-621.

[88] Jordan, I.K., Rogozin, I.B., Glazko, G.V. and Koonin, E.V.: Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet*, **2003,** *19*, 68-72.

[89] van de Lagemaat, L.N., Landry, J.R., Mager, D.L. and Medstrand, P.: Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet*, **2003,** *19*, 530-536.

[90] Brandt, J., Schrauth, S., Veith, A.M., Froschauer, A., Haneke, T., Schultheis, C., Gessler, M., Leimeister, C. and Volff, J.-N.: Transposable elements as a source of genetic innovation: expression and evolution of a family of retrotransposon-derived neogenes in mammals. *Gene*, **2005,** *345*, 101-111.

[91] Curcio, M.J. and Derbyshire, K.M.: The outs and ins of transposition: from mu to kangaroo. *Nat Rev Mol Cell Biol*, **2003,** *4*, 865-877.

[92] Fischer, C., Bouneau, L., Coutanceau, J.P., Weissenbach, J., Ozouf-Costaz, C. and Volff, J.-N.: Diversity and clustered distribution of retrotransposable elements in the compact genome of the pufferfish *Tetraodon nigroviridis. Cytogenet Genome Res*, **2005,** *110*, 522-536.

[93] Frame, I.G., Cutfield, J.F. and Poulter, R.T.: New BEL-like LTR-retrotransposons in *Fugu rubripes*, *Caenorhabditis elegans*, and *Drosophila melanogaster*. *Gene*, **2001,** *263*, 219-230.

[94] Goodwin, T.J. and Poulter, R.T.: A new group of tyrosine recombinase-encoding retrotransposons. *Mol Biol Evol*, **2004,** *21*, 746-759.

[95] Lyozin, G.T., Makarova, K.S., Velikodvorskaja, V.V., Zelentsova, H.S., Khechumian, R.R., Kidwell, M.G., Koonin, E.V. and Evgen'ev, M.B.: The structure and evolution of Penelope in the *virilis* species group of *Drosophila*: an ancient lineage of retroelements. *J Mol Evol*, **2001,** *52*, 445-456.

[96] Volff, J.-N., Hornung, U. and Schartl, M.: Fish retroposons related to the Penelope element of *Drosophila virilis* define a new group of retrotransposable elements. *Mol Genet Genomics*, **2001,** *265*, 711-720.

[97] Volff, J.-N., Körting, C., Froschauer, A., Sweeney, K. and Schartl, M.: Non-LTR retrotransposons encoding a restriction enzyme-like endonuclease in vertebrates. *J Mol Evol*, **2001,** *52*, 351-360.

[98] Bouneau, L., Fischer, C., Ozouf-Costaz, C., Froschauer, A., Jaillon, O., Coutanceau, J.P., Körting, C., Weissenbach, J., Bernot, A. and Volff, J.-N.: An active non-LTR retrotransposon with tandem structure in the compact genome of the pufferfish *Tetraodon nigroviridis*. *Genome Res*, **2003,** *13*, 1686-1695.

[99] Zhou, Q., Froschauer, A., Schultheis, C., Schmidt, C., Bienert, P., Wenning, M., Dettai, A. and Volff, J.-N.: Helitron transposons on the sex chromosomes of the platyfish *Xiphophorus maculatus* and their evolution in animal genomes. *Zebrafish*, **2005,** *(in press)*.

[100] Dasilva, C., Hadji, H., Ozouf-Costaz, C., Nicaud, S., Jaillon, O., Weissenbach, J. and Crollius, H.R.: Remarkable compartmentalization of transposable elements and pseudogenes in the heterochromatin of the *Tetraodon nigroviridis* genome. *Proc Natl Acad Sci U S A*, **2002,** *99*, 13636-13641.

[101] Fischer, C., Bouneau, L., Coutanceau, J.P., Weissenbach, J., Volff, J.-N. and Ozouf-Costaz, C.: Global heterochromatic colocalization of transposable elements with minisatellites in the compact genome of the pufferfish *Tetraodon nigroviridis*. *Gene*, **2004,** *336*, 175-183.

[102] Graves, J.A. and Shetty, S.: Sex from W to Z: Evolution of vertebrate sex chromosomes and sex determining genes. *J Exp Zool*, **2001,** *290*, 449-462.

[103] Baroiller, J.F. and D'Cotta, H.: Environment and sex determination in farmed fish. *Comp Biochem Physiol C Toxicol Pharmacol*, **2001,** *130*, 399-409.

[104] Schartl, M.: Sex chromosome evolution in non-mammalian vertebrates. *Curr Opin Genet Dev*, **2004,** *14*, 634-641.

[105] Matsuda, M., Nagahama, Y., Shinomiya, A., *et al.*: *DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature*, **2002,** *417*, 559-563.

[106] Nanda, I., Kondo, M., Hornung, U., *et al.*: A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. *Proc Natl Acad Sci U S A*, **2002,** *99*, 11778-11783.

[107] Schartl, M.: A comparative view on sex determination in medaka. *Mech Dev*, **2004,** *121*, 639-645.

[108] Matsuda, M.: Sex determination in the teleost medaka, *Oryzias latipes*. *Annu Rev Genet*, **2005,** *39*, 293-307.

[109] Kondo, M., Nanda, I., Hornung, U., Schmid, M. and Schartl, M.: Evolutionary origin of the medaka Y chromosome. *Curr Biol*, **2004,** *14*, 1664-1669.

[110] Kondo, M., Nanda, I., Hornung, U., Asakawa, S., Shimizu, N., Mitani, H., Schmid, M., Shima, A. and Schartl, M.: Absence of the candidate male sex-determining gene *dmrt1b(Y)* of medaka from other fish species. *Curr Biol*, **2003,** *13*, 416-420.

[111] Matsuda, M., Sato, T., Toyazaki, Y., Nagahama, Y., Hamaguchi, S. and Sakaizumi, M.: *Oryzias curvinotus* has *DMY*, a gene that is required for male development in the medaka, *O. latipes*. *Zoolog Sci*, **2003,** *20*, 159-161.

[112] Veith, A.M., Froschauer, A., Körting, C., Nanda, I., Hanel, R., Schmid, M., Schartl, M. and Volff, J.-N.: Cloning of the *dmrt1* gene of *Xiphophorus maculatus*: *dmY/dmrt1Y* is not the master sex-determining gene in the platyfish. *Gene*, **2003,** *317*, 59-66.

[113] Peichel, C.L., Ross, J.A., Matson, C.K., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Mori, S., Schluter, D. and Kingsley, D.M.: The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Curr Biol*, **2004,** *14*, 1416-1424.

[114] Woram, R.A., Gharbi, K., Sakamoto, T., *et al.*: Comparative genome analysis of the primary sex-determining locus in salmonid fishes. *Genome Res*, **2003,** *13*, 272-280.

[115] Chiu, C.H., Amemiya, C., Dewar, K., Kim, C.B., Ruddle, F.H. and Wagner, G.P.: Molecular evolution of the *HoxA* cluster in the three major gnathostome lineages. *Proc Natl Acad Sci U S A*, **2002,** *99*, 5492-5497.

[116] Chiu, C.H., Dewar, K., Wagner, G.P., *et al.*: Bichir *HoxA* cluster sequence reveals surprising trends in ray-finned fish genomic evolution. *Genome Res*, **2004,** *14*, 11-17.

[117] Venkatesh, B., Tay, A., Dandona, N., Patil, J.G. and Brenner, S.: A compact cartilaginous fish model genome. *Curr Biol*, **2005,** *15*, R82-83.

[118] Saitou, N. and Nei, M.: The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*, **1987,** *4*, 406-425.

[119] Stock, D.W., Ellies, D.L., Zhao, Z., Ekker, M., Ruddle, F.H. and Weiss, K.M.: The evolution of the vertebrate *Dlx* gene family. *Proc Natl Acad Sci U S A*, **1996,** *93*, 10858-10863.

[120] Haddon, C., Smithers, L., Schneider-Maunoury, S., Coche, T., Henrique, D. and Lewis, J.: Multiple delta genes and lateral inhibition in zebrafish primary neurogenesis. *Development*, **1998,** *125*, 359-370.

[121] Ekker, M., Akimenko, M.A., Bremiller, R. and Westerfield, M.: Regional expression of three homeobox transcripts in the inner ear of zebrafish embryos. *Neuron*, **1992,** *9*, 27-35.

[122] Akimenko, M.A., Johnson, S.L., Westerfield, M. and Ekker, M.: Differential induction of four *msx* homeobox genes during fin development and regeneration in zebrafish. *Development*, **1995,** *121*, 347-357.

[123] Parichy, D.M., Rawls, J.F., Pratt, S.J., Whitfield, T.T. and Johnson, S.L.: Zebrafish sparse corresponds to an orthologue of *c-kit* and is required for the morphogenesis of a subpopulation of melanocytes, but is not essential for hematopoiesis or primordial germ cell development. *Development*, **1999,** *126*, 3425-3436.

[124] Mellgren, E.M. and Johnson, S.L.: *Kitb*, a second zebrafish ortholog of mouse Kit. *Dev Genes Evol*, **2005,** *215*, 470-477.

[125] Merrit, T.J. and Quattro, M.: Evolution of the vertebrate cytosolic malate dehydrogenase gene family: duplication and divergence in actinopterygian fish. *J Mol Evol*, **2003,** *56*, 265-276.

[126] Musa, A., Lehrach, H. and Russo, V.A.: Distinct expression patterns of two zebrafish homologues of the human *APP* gene during embryonic development. *Dev Genes Evol*, **2001,** *211*, 563-567.

[127] Roelink, H., Augsburger, A., Heemskerk, J., *et al.*: Floor plate and motor neuron induction by *vhh-1*, a vertebrate homolog of hedgehog expressed by the notochord. *Cell*, **1994,** *76*, 761-775.

[128] Ekker, S.C., Ungar, A.R., Greenstein, P., von Kessler, D.P., Porter, J.A., Moon, R.T. and Beachy, P.A.: Patterning activities of vertebrate hedgehog proteins in the developing eye and brain. *Curr Biol*, **1995,** *5*, 944-955.

[129] Ristoratore, F., Carl, M., Deschet, K., Richard-Parpaillon, L., Boujard, D., Wittbrodt, J., Chourrout, D., Bourrat, F. and Joly, J.S.: The midbrain-hindbrain boundary genetic cascade is activated ectopically in the diencephalon in response to the widespread expression of one of its components, the medaka gene *Ol-eng2*. *Development*, **1999,** *126*, 3769-3779.

[130] Holland, P.W. and Williams, N.A.: Conservation of engrailed-like homeobox sequences during vertebrate evolution. *FEBS Lett*, **1990,** *277*, 250-252.

[131] Wang, Y., Macke, J.P., Abella, B.S., Andreasson, K., Worley, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A. and Nathans, J.: A large family of putative transmembrane receptors homologous to the product of the *Drosophila* tissue polarity gene frizzled. *J Biol Chem*, **1996,** *271*, 4468-4476.

[132] Yan, Y.L., Willoughby, J., Liu, D., Crump, J.G., Wilson, C., Miller, C.T., Singer, A., Kimmel, C., Westerfield, M. and Postlethwait, J.H.: A pair of *Sox*: distinct and overlapping functions of zebrafish *sox9* co-orthologs in craniofacial and pectoral fin development. *Development*, **2005,** *132*, 1069-1083.

[133] Dutton, K.A., Pauliny, A., Lopes, S.S., Elworthy, S., Carney, T.J., Rauch, J., Geisler, R., Haffter, P. and Kelsh, R.N.: Zebrafish colourless encodes *sox10* and specifies non-ectomesenchymal neural crest fates. *Development*, **2001,** *128*, 4113-4125.

[134] Koopman, P., Schepers, G., Brenner, S. and Venkatesh, B.: Origin and diversity of the *SOX* transcription factor gene family: genome-wide analysis in *Fugu rubripes*. *Gene*, **2004,** *328*, 177-186.

[135] Lang, M., Miyake, T., Braasch, I., Tinnemore, D., Siegel, N., Salzburger, W., Amemiya, C.T. and Meyer, A.: A BAC library of the East African haplochromine cichlid fish *Astatotilapia burtoni*. *J Exp Zoolog B Mol Dev Evol*, **2005**.

[136] Irwin, D.M.: A second insulin gene in fish genomes. *Gen Comp Endocrinol*, **2004,** *135*, 150-158.

[137] Bierkamp, C. and Campos-Ortega, J.A.: A zebrafish homologue of the *Drosophila* neurogenic gene *Notch* and its pattern of transcription during early embryogenesis. *Mech Dev*, **1993,** *43*, 87-100.

[138] Lyons, D.A., Pogoda, H.M., Voas, M.G., Woods, I.G., Diamond, B., Nix, R., Arana, N., Jacobs, J. and Talbot, W.S.: *Erbb3* and *erbb2* are essential for schwann cell migration and myelination in zebrafish. *Curr Biol*, **2005,** *15*, 513-524.

[139] Maures, T., Chan, S.J., Xu, B., Sun, H., Ding, J. and Duan, C.: Structural, biochemical, and expression analysis of two distinct insulin-like growth factor I receptors and their ligands in zebrafish. *Endocrinology*, **2002,** *143*, 1858-1871.

[140] Winkler, C., Schafer, M., Duschl, J., Schartl, M. and Volff, J.-N.: Functional divergence of two zebrafish midkine growth factors following fish-specific gene duplication. *Genome Res*, **2003,** *13*, 1067-1081.

[141] Gong, Z., Hui, C.C. and Hew, C.L.: Presence of *isl-1*-related LIM domain homeobox genes in teleost and their similar patterns of expression in brain and spinal cord. *J Biol Chem*, **1995,** *270*, 3335-3345.

[142] Nornes, S., Clarkson, M., Mikkola, I., Pedersen, M., Bardsley, A., Martinez, J.P., Krauss, S. and Johansen, T.: Zebrafish contains two *pax6* genes involved in eye development. *Mech Dev*, **1998,** *77*, 185-196.

[143] Dildrop, R. and Ruther, U.: Organization of Iroquois genes in fish. *Dev Genes Evol*, **2004,** *214*, 267-276.

[144] Thisse, C., Thisse, B., Schilling, T.F. and Postlethwait, J.H.: Structure of the zebrafish *snail1* gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development*, **1993,** *119*, 1203-1215.

[145] Thisse, C., Thisse, B. and Postlethwait, J.H.: Expression of *snail2*, a second member of the zebrafish *snail* family, in cephalic mesendoderm and presumptive neural crest of wild-type and spadetail mutant embryos. *Dev Biol*, **1995,** *172*, 86-99.

[146] Yu, W.P., Pallen, C.J., Tay, A., Jirik, F.R., Brenner, S., Tan, Y.H. and Venkatesh, B.: Conserved synteny between the Fugu and human *PTEN* locus and the evolutionary conservation of vertebrate PTEN function. *Oncogene*, **2001,** *20*, 5554-5561.

[147] Croushore, J.A., Blasiole, B., Riddle, R.C., Thisse, C., Thisse, B., Canfield, V.A., Robertson, G.P., Cheng, K.C. and Levenson, R.: *Ptena* and *ptenb* genes play distinct roles in zebrafish embryogenesis. *Dev Dyn*, **2005,** *234*, 911-921.

[148] Hutson, L.D., Jurynec, M.J., Yeo, S.Y., Okamoto, H. and Chien, C.B.: Two divergent *slit1* genes in zebrafish. *Dev Dyn*, **2003,** *228*, 358-369.

[149] Adam, D., Dimitrijevic, N. and Schartl, M.: Tumor suppression in *Xiphophorus* by an accidentally acquired promoter. *Science*, **1993,** *259*, 816-819.

[150] Goishi, K., Lee, P., Davidson, A.J., Nishi, E., Zon, L.I. and Klagsbrun, M.: Inhibition of zebrafish epidermal growth factor re-

ceptor activity results in cardiovascular defects. *Mech Dev*, **2003,** *120*, 811-822.

[151] Bury, N.R., Sturm, A., Le Rouzic, P., Lethimonier, C., Ducouret, B., Guiguen, Y., Robinson-Rechavi, M., Laudet, V., Rafestin-Oblin, M.E. and Prunet, P.: Evidence for two distinct functional glucocorticoid receptors in teleost fish. *J Mol Endocrinol*, **2003,** *31*, 141-156.

[152] Greenwood, A.K., Butler, P.C., White, R.B., DeMarco, U., Pearce, D. and Fernald, R.D.: Multiple corticosteroid receptors in a teleost fish: distinct sequences, expression patterns, and transcriptional activities. *Endocrinology*, **2003,** *144*, 4226-4236.

[153] Ringholm, A., Fredriksson, R., Poliakova, N., Yan, Y.L., Postlethwait, J.H., Larhammar, D. and Schiöth, H.B.: One melanocortin 4 and two melanocortin 5 receptors from zebrafish show remarkable conservation in structure and pharmacology. *J Neurochem*, **2002,** *82*, 6-18.

[154] Merritt, T.J. and Quattro, J.M.: Evidence for a period of directional selection following gene duplication in a neurally expressed locus of triosephosphate isomerase. *Genetics*, **2001,** *159*, 689-697.

[155] Mavropoulos, A., Devos, N., Biemar, F., Zecchin, E., Argenton, F., Edlund, H., Motte, P., Martial, J.A. and Peers, B.: sox4b is a key player of pancreatic alpha cell differentiation in zebrafish. *Dev Biol*, **2005,** *285*, 211-223.

[156] Sharma, M.K., Saxena, V., Liu, R.Z., Thisse, C., Thisse, B., Denovan-Wright, E.M. and Wright, J.M.: Differential expression of the duplicated cellular retinoic acid-binding protein 2 genes (*crabp2a* and *crabp2b*) during zebrafish embryonic development. *Gene Expr Patterns*, **2005,** *5*, 371-379.

[157] Malaga-Trillo, E., Laessing, U., Lang, D.M., Meyer, A. and Stuermer, C.A.: Evolution of duplicated *reggie* genes in zebrafish and goldfish. *J Mol Evol*, **2002,** *54*, 235-245.

[158] Rimini, R., Beltrame, M., Argenton, F., Szymczak, D., Cotelli, F. and Bianchi, M.E.: Expression patterns of zebrafish *sox11A*, *sox11B* and *sox21*. *Mech Dev*, **1999,** *89*, 167-171.

[159] Risinger, C., Salaneck, E., Soderberg, C., Gates, M., Postlethwait, J.H. and Larhammar, D.: Cloning of two loci for synapse protein Snap25 in zebrafish: comparison of paralogous linkage groups suggests loss of one locus in the mammalian lineage. *J Neurosci Res*, **1998,** *54*, 563-573.

[160] Martinez-Barbera, J.P., Toresson, H., Da Rocha, S. and Krauss, S.: Cloning and expression of three members of the zebrafish *Bmp* family: *Bmp2a*, *Bmp2b* and *Bmp4*. *Gene*, **1997,** *198*, 53-59.

[161] Flores, M.V., Tsang, V.W., Hu, W., Kalev-Zylinska, M., Postlethwait, J., Crosier, P., Crosier, K. and Fisher, S.: Duplicate zebrafish *runx2* orthologues are expressed in developing skeletal elements. *Gene Expr Patterns*, **2004,** *4*, 573-581.

[162] Chiang, E.F., Yan, Y.L., Guiguen, Y., Postlethwait, J. and Chung, B.: Two *Cyp19* (P450 Aromatase) Genes on duplicated zebrafish chromosomes are expressed in ovary or brain. *Mol Biol Evol*, **2001,** *18*, 542-550.

[163] Kishida, M. and Callard, G.V.: Distinct cytochrome P450 aromatase isoforms in zebrafish (*Danio rerio*) brain and ovary are differentially programmed and estrogen regulated during early development. *Endocrinology*, **2001,** *142*, 740-750.

[164] Higashijima, S., Nose, A., Eguchi, G., Hotta, Y. and Okamoto, H.: Mindin/F-spondin family: novel ECM proteins expressed in the zebrafish embryonic axis. *Dev Biol*, **1997,** *192*, 211-227.

[165] Winkler, C., Elmasri, H., Klamt, B., Volff, J.-N. and Gessler, M.: Characterization of *hey* bHLH genes in teleost fish. *Dev Genes Evol*, **2003,** *213*, 541-553.

[166] Bollig, F., Mehringer, R., Perner, B., Hartung, C., Schafer, M., Schartl, M., Volff, J.-N., Winkler, C. and Englert, C.: Identification and comparative expression analysis of a second *wt1* gene in zebrafish. *Dev Dyn*, **2005,** published online Nov 16.