

## REVIEW

## Genome evolution and biodiversity in teleost fish

J-N Volff

BioFuture Research Group, Physiologische Chemie I, Biozentrum, University of Würzburg, am Hubland, D-97074 Würzburg, Germany

Teleost fish, which roughly make up half of the extant vertebrate species, exhibit an amazing level of biodiversity affecting their morphology, ecology and behaviour as well as many other aspects of their biology. This huge variability makes fish extremely attractive for the study of many biological questions, particularly of those related to evolution. New insights gained from different teleost species and sequencing projects have recently revealed several peculiar features of fish genomes that might have played a role in fish evolution and speciation. There is now substantial evidence that a round of tetraploidization/rediploidization has taken place during the early evolution of the ray-finned fish lineage, and that hundreds of duplicate pairs generated by this event have been maintained over hundreds of millions of years of evolution. Differential loss or subfunction partitioning of such gene duplicates might have been involved in the generation

of fish variability. In contrast to mammalian genomes, teleost genomes also contain multiple families of active transposable elements, which might have played a role in speciation by affecting hybrid sterility and viability. Finally, the amazing diversity of sex determination systems and the plasticity of sex chromosomes observed in teleost might have been involved in both pre- and postmating reproductive isolation. Comparison of data generated by current and future genome projects as well as complementary studies in other species will allow one to approach the molecular and evolutionary mechanisms underlying genome diversity in fish, and will certainly significantly contribute to our understanding of gene evolution and function in humans and other vertebrates. *Heredity* (2005) **94**, 280–294. doi:10.1038/sj.hdy.6800635  
Published online 22 December 2004

**Keywords:** teleost fish; genome duplication; sex chromosomes; sex determination; transposable elements; speciation

## Introduction

Fishes are an extremely diverse group of vertebrate aquatic animals usually breathing through gills throughout life and having fins and scales. They include jawless fishes (hagfishes, lampreys), cartilaginous fishes (sharks, rays) and bony fishes (coelacanth, lungfishes and ray-finned fishes) (Nelson, 1994). With more than 23 500 species, ray-finned fishes (actinopterygians) represent more than 95% of all living fish species, and roughly make up half of the extant vertebrate species. More than 99.8% of ray-finned fishes belong to the teleosts. Bichirs and sturgeons are examples of nonteleost ray-finned fishes (Figure 1). According to traditional views, fishes do not form a unique monophyletic group distinct from tetrapods (mammals, reptiles, birds and amphibians) (Nelson, 1994; but see also Rasmussen and Arnason, 1999). Rather, bony fishes are thought to be more related to tetrapods than to jawless and cartilaginous fishes (Figure 1). This means that humans and other land vertebrates probably all share a 360–450 million years old fish ancestor.

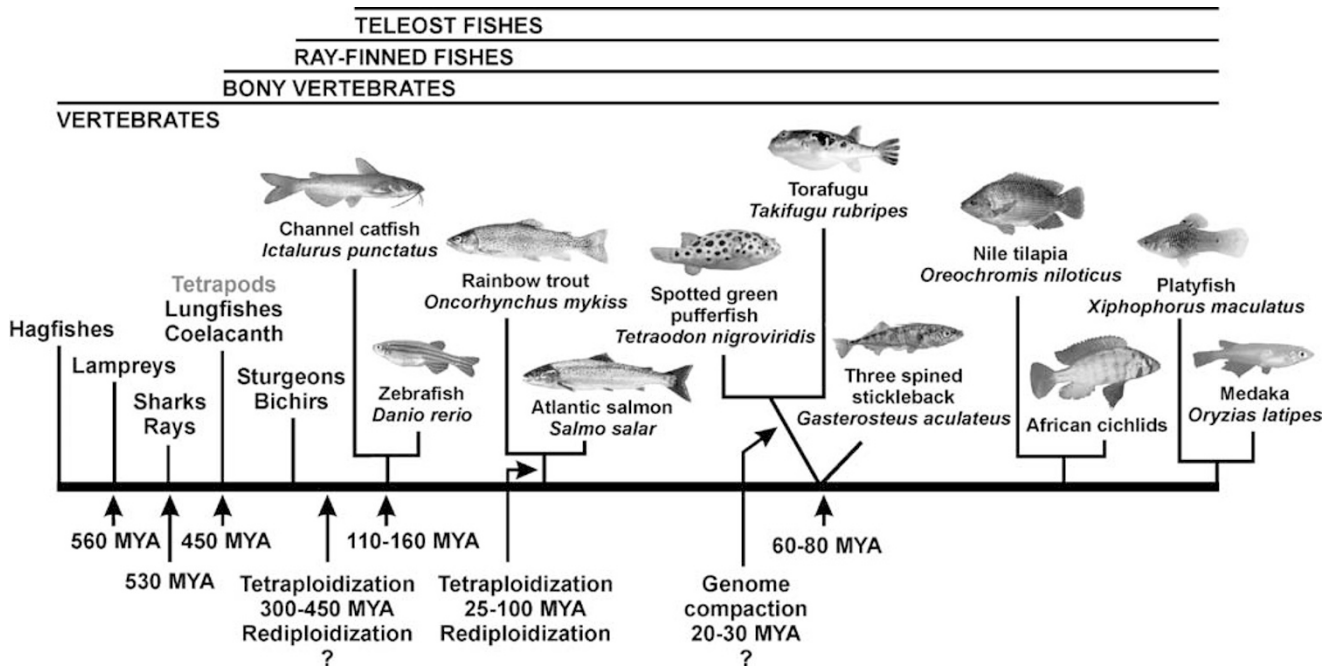
Fishes show a remarkable level of diversity affecting their morphology, ecology, behavior and genomes as well as multiple other facets of their biology (Nelson, 1994). This makes them extremely attractive for the study of many evolutionary questions related to diverse aspects

of biology. Fish biodiversity is important to humans at the economical, ecological and cultural points of view, and its maintenance is an important challenge for the next generations. New insights from several fish species and sequencing projects have shed new light on the organization and evolution of fish genomes and now allow one to approach the evolutionary mechanisms possibly underlying biodiversity in the fish lineage.

## Principal teleost fish models

Several teleost fish species are particularly studied at the genetic and genomic levels; some of them have been, are or will certainly be subjects of whole-genome sequencing projects (Figure 1 and Table 1). For species like the Atlantic salmon *Salmo salar*, the rainbow trout *Oncorhynchus mykiss*, the Nile tilapia *Oreochromis niloticus* and the channel catfish *Ictalurus punctatus*, genetics and genomics programmes have at least partially an economical motivation. Such programmes essentially aim to identify at the molecular level qualitative and quantitative trait loci (QTLs), controlling among others growth, reproduction, pigmentation, environmental tolerance or resistance to disease, which are highly relevant traits for aquaculture.

Other fish species are widely used as 'piscine mice' in developmental biology. This is particularly true for the zebrafish *Danio rerio* and the Japanese medaka *Oryzias latipes*, two well-established complementary models for the study of different aspects of vertebrate organogenesis (Grunwald and Eisen, 2002; Wittbrodt *et al.*, 2002; Furutani-Seiki and Wittbrodt, 2004). In both species, gene function can be studied routinely *in vivo* in the laboratory particularly through transgenesis and 'mor-



**Figure 1** The fish lineage. Origin of fish pictures: Manfred Scharl 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).

pholino' antisense oligonucleotide 'knockdown' technology (the classical knockout technology as used in the mouse is not available at the moment in fish). Large-scale mutagenesis programmes have been performed in the zebrafish and more recently in the medaka (Driever *et al*, 1996; Haffter *et al*, 1996; Furutani-Seiki *et al*, 2004). Some mutants were obtained in one species and not in the other, demonstrating the complementarity of the models.

Another small aquarium fish, the platyfish *Xiphophorus maculatus*, is a traditional model for the study of the development of pigment cells and melanoma (Scharl, 1995; Meierjohann *et al*, 2004), but its use for *in vivo* analysis is restricted by the fact that this fish is viviparous (livebearer). The pufferfishes *Takifugu rubripes* (Torafugu) and *Tetraodon nigroviridis* (spotted green pufferfish) are pure genomics models. They are studied because of the compactness of their genome, characterized by small intronic and intergenic regions (Brenner *et al*, 1993; Roest Crolius *et al*, 2000), but at the moment these animals cannot be crossed routinely in the laboratory and are not usable *in vivo* for functional analysis of gene function. Finally, cichlids from the great lakes of East Africa, sticklebacks and guppies (*Poecilia reticulata*) are important models to study the molecular basis of speciation and the evolutionary mechanisms shaping development and behaviour (Peichel *et al*, 2001; Brooks, 2002; Verheyen *et al*, 2003; Kocher, 2004; Shapiro *et al*, 2004).

The different domains of research described here for each fish species do not exclude other types of investigations. For example, the rainbow trout is not only studied with regard to its economical interest but is also a model for more fundamental research in carcinogenesis, toxicology, comparative immunology, physiology and others (Thorgaard *et al*, 2002). Several topics are investigated in parallel in different species. Sex determi-

nation and sex differentiation are studied in numerous fishes including medaka, zebrafish, salmonids, platyfish, sticklebacks, tilapia and others (Baroiller and D'Cotta, 2001; Devlin and Nagahama, 2002; Volff and Scharl, 2002). Such comparative analyses are of the highest importance to understand the mechanisms driving differential evolution in fish sublineages.

## Fish genome projects

The genome of the pufferfish *Takifugu rubripes* was, after the human genome, the second vertebrate genome to be sequenced (Aparicio *et al*, 2002). Its sequencing through whole-genome shotgun (WGS) strategy allowed the first genome-wide comparison between two vertebrate species. Pufferfish and mammals have approximately the same number of genes, but the Torafugu genome is 8–9 times smaller than the human genome. This is principally due to the fact that nonexonic regions (intronic and intergenic sequences) are generally – but not always – much shorter in the pufferfish than in humans, because of a relative paucity of repetitive sequences. The third assembly of the Torafugu genome is available and consists of approximately 8000 genomic scaffolds covering approximately 95% of the nonrepetitive fraction of the genome. Genome compaction is also observed in the green spotted pufferfish *T. nigroviridis*, the genome of which has been almost completely sequenced too (Jaillon *et al*, 2004). Sequence data have been assembled in approximately 50 000 contigs covering 312 Mbp of the 385 Mbp genome. These contigs have been further linked particularly through fluorescent *in situ* hybridization of genomic clones on Tetraodon chromosomes (Jaillon *et al*, 2004). The genome sequence of both smooth pufferfishes is useful for the identification of coding and regulatory sequences in humans and other vertebrates through

**Table 1** Teleost fish species with genetic and genomic resources

Scientific name	Common name	Genome size <sup>a</sup> (Mpb)	Resources <sup>b</sup>	References	Web resources for analysis <sup>c</sup>
<i>Danio rerio</i>	Zebrafish	1700 (2n = 50)	Genetic map, ESTs	Postlethwait <i>et al</i> (1994) Kelly <i>et al</i> (2000)	<a href="http://zfin.org/">http://zfin.org/</a> and linked sites
			Radiation hybrid map	Hukriede <i>et al</i> (1999) Geisler <i>et al</i> (1999)	<a href="http://zfin.org/">http://zfin.org/</a> and linked sites
			Physical map Genome project		<a href="http://www.sanger.ac.uk/Projects/D_rerio/mapping.shtml">http://www.sanger.ac.uk/Projects/D_rerio/mapping.shtml</a> <a href="http://www.ensembl.org/Danio_rerio/">http://www.ensembl.org/Danio_rerio/</a> <a href="http://www.sanger.ac.uk/Projects/D_rerio/">http://www.sanger.ac.uk/Projects/D_rerio/</a>
<i>Ictalurus punctatus</i>	Channel catfish	1000 (2n = 58)	Genetic map ESTs	Liu <i>et al</i> (2003) He <i>et al</i> (2003)	<a href="http://www.tigr.org/tdb/tgi">http://www.tigr.org/tdb/tgi</a> <a href="http://morag.umsmed.edu/">http://morag.umsmed.edu/</a>
<i>Oncorhynchus mykiss</i>	Rainbow trout	2600 (2n = 58–64)	Genetic map ESTs	Nichols <i>et al</i> (2003) Rexroad <i>et al</i> (2003)	<a href="http://www.tigr.org/tdb/tgi/">http://www.tigr.org/tdb/tgi/</a>
<i>Salmo salar</i>	Atlantic salmon	3000 (2n = 54–60)	Genetic map ESTs	Gilbey <i>et al</i> (2004) Moen <i>et al</i> (2004) Rise <i>et al</i> (2004)	<a href="http://web.uvic.ca/cbr/grasp/">http://web.uvic.ca/cbr/grasp/</a> <a href="http://www.tigr.org/tdb/tgi">http://www.tigr.org/tdb/tgi</a> <a href="http://www.salmongenome.no/cgi-bin/sgp.cgi">http://www.salmongenome.no/cgi-bin/sgp.cgi</a> <a href="http://cegs.stanford.edu/Linkage_Maps.jsp">http://cegs.stanford.edu/Linkage_Maps.jsp</a> <a href="http://cegs.stanford.edu/Physical_map.jsp">http://cegs.stanford.edu/Physical_map.jsp</a> <a href="http://cegs.stanford.edu/stickleback_EST_sequences.jsp">http://cegs.stanford.edu/stickleback_EST_sequences.jsp</a>
<i>Gasterosteus aculeatus</i>	Three-spined stickleback	675 (2n = 42)	Genetic map Physical map ESTs	Peichel <i>et al</i> (2001)	<a href="http://www.thearkdb.org/browser?species=tilapia">http://www.thearkdb.org/browser?species=tilapia</a> <a href="http://hcgs.unh.edu/comp/">http://hcgs.unh.edu/comp/</a> <a href="http://hcgs.unh.edu/fpc/image.php">http://hcgs.unh.edu/fpc/image.php</a>
<i>Oreochromis niloticus</i>	Nile tilapia	1000 (2n = 44)	Genetic map Physical map Genetic map ESTs	Kocher <i>et al</i> (1998) Agresti <i>et al</i> (2000) Albertson <i>et al</i> (2003)	<a href="http://www.thearkdb.org/browser?species=tilapia">http://www.thearkdb.org/browser?species=tilapia</a> <a href="http://hcgs.unh.edu/comp/">http://hcgs.unh.edu/comp/</a> <a href="http://www.tigr.org/tdb/tgi/">http://www.tigr.org/tdb/tgi/</a>
Several species of <i>Takifugu rubripes</i>	East African cichlids Torafugu	400 (2n = 44)	ESTs Genome project	Clark <i>et al</i> (2003) Aparicio <i>et al</i> (2002)	<a href="http://www.fugu-sg.org/index.html">http://www.fugu-sg.org/index.html</a> <a href="http://www.ensembl.org/Fugu_rubripes/">http://www.ensembl.org/Fugu_rubripes/</a>
<i>Tetraodon nigroviridis</i>	Spotted green pufferfish	385 (2n = 42)	Physical map Genome project	Jaillon <i>et al</i> (2004) Jaillon <i>et al</i> (2004)	<a href="http://www.genoscope.cns.fr/externe/tetranew/">http://www.genoscope.cns.fr/externe/tetranew/</a> <a href="http://www.genoscope.cns.fr/externe/tetranew/">http://www.genoscope.cns.fr/externe/tetranew/</a>
<i>Xiphophorus maculatus</i>	Southern platyfish	900 (2n = 48)	Genetic map Gonosome project	Kazianis <i>et al</i> (1996) Froschauer <i>et al</i> (2003)	<a href="http://www.xiphophorus.org/map.htm">http://www.xiphophorus.org/map.htm</a>
<i>Oryzias latipes</i>	Japanese medaka	1000 (2n = 48)	Genetic map Physical map ESTs Genome project	Naruse <i>et al</i> (2004) Khorasani <i>et al</i> (2004) Kimura <i>et al</i> (2004)	<a href="http://gardy.dsp.jst.go.jp/MedGenIn/">http://gardy.dsp.jst.go.jp/MedGenIn/</a> <a href="http://medaka.lab.nig.ac.jp/est_index.html">http://medaka.lab.nig.ac.jp/est_index.html</a> <a href="http://dolphin.lab.nig.ac.jp/medaka/">http://dolphin.lab.nig.ac.jp/medaka/</a> <a href="http://shigen.lab.nig.ac.jp/medaka/genome/top.jsp">http://shigen.lab.nig.ac.jp/medaka/genome/top.jsp</a>

<sup>a</sup>Haploid genome size and chromosome number (also see <http://www.genomesonline.org/> and <http://www.genomesize.com/>).

<sup>b</sup>Linkage maps also exist for other fish species without significant genomic resources including several salmonids, the blind cave fish *Astyanax mexicanus* and the guppy *Poecilia reticulata* (Danzmann and Gharbi (2001) and references therein; Borowsky and Wilkens, 2002; Khoo *et al*, 2003; Woram *et al*, 2004).

<sup>c</sup>Nucleic sequences are generally also accessible using the NCBI blast server (<http://www.ncbi.nih.gov/BLAST/>).

sequence comparison, since sequences with functions should be more conserved than 'useless' sequences. The sequence of *T. nigroviridis* has been used to predict the number of genes in the human genome (Roest Crolius *et al.*, 2000). Importantly, hundreds of putative novel human genes have been discovered by comparing the pufferfish and human genome sequences (Aparicio *et al.*, 2002; Jaillon *et al.*, 2004). In addition, analysis of the *T. nigroviridis* genome revealed the basic structure of the ancestral bony vertebrate genome, which was composed of 12 chromosomes, and allowed to reconstruct many of the chromosome rearrangements which led to the modern human karyotype (Jaillon *et al.*, 2004). Finally, *Takifugu/Tetraodon* comparisons might provide new information about differences between relatively related species, in a manner similar to the comparison between rat and mouse.

The sequencing of the genome of both zebrafish and medaka is close to completion. Sequence reads and preliminary contigs are publicly available, but the final analyses have not been published to date. The sequencing of the zebrafish genome has been initiated in 2001 following two strategies: clone mapping and sequencing from genomic clone libraries and WGS sequencing with subsequent assembly (Table 1). The fourth WGS assembly Zv4 has been released in July 2004. This assembly consists of approximately 21 000 contigs covering 1500 Mbp of the zebrafish genome. The sequencing of the genome of the medaka *Oryzias latipes*, mainly based on a whole shotgun sequence strategy, has been started in 2002. A nine-fold coverage of the genome has been already obtained in May 2004, and a first WGS assembly has been released in July 2004 (about 116 000 sequences covering 840 Mbp; see Naruse *et al.* (2004a) for additional information and useful www resources). The availability and comparison of the zebrafish and medaka genome drafts will allow linking mutant phenotypes to gene functions and shedding a new 'evo-devo' light on the fish and vertebrate lineages.

Finally, for other fishes (Table 1), important expressed sequence tags (ESTs) resources are already available. There is no doubt that some of these species will be subjected to genome sequencing in the near future due to their economical and/or academic relevance, and several proposals way have already been submitted to funding agencies. On the subgenomic level, a project dealing with the physical mapping and sequencing of the sex chromosomes of the platyfish *Xiphophorus maculatus* has been started (Froschauer *et al.*, 2002).

## Fish-specific gene and genome duplications

*Hox* gene clusters and a plethora of other genes have been duplicated in the teleost fish lineage after its divergence from tetrapods (Amores *et al.*, 1998; Wittbrodt *et al.*, 1998; Meyer and Schartl, 1999; Robinson-Rechavi *et al.*, 2001a; Loh *et al.*, 2004; Postlethwait *et al.*, 2004). For some gene pairs (for example *Xmrk/egfrb* in *Xiphophorus*; Volff and Schartl, 2003), the duplication events are rather recent and clearly affected only a restricted chromosomal region in a particular fish sublineage. However, in multiple other cases, the paralogous (duplicated) sequences are much more ancient (Taylor *et al.*, 2001a). Such duplicates have been mapped on different chromosomes within larger duplicated regions (paralogons) in several

divergent teleost fish species (Postlethwait *et al.*, 2000; Woods *et al.*, 2000; Morizot *et al.*, 2001; Taylor *et al.*, 2003; Winkler *et al.*, 2003b; Naruse *et al.*, 2004b). Phylogenomic analysis confirmed the presence in Torafugu, Tetraodon and zebrafish of hundreds of duplicates being co-orthologous to single-copy tetrapod genes (Taylor *et al.*, 2003; Christoffels *et al.*, 2004; Jaillon *et al.*, 2004; Vandepoele *et al.*, 2004). Sequence divergence analysis between paralogs suggests that these genes have been duplicated within the same time window 300–450 million years ago (Taylor *et al.*, 2001a) after divergence of sturgeons from the lineage that led to teleosts (Hoegg *et al.*, 2004). Taken together, these observations have suggested that the ray-finned fish lineage has experienced an event of tetraploidization early during its evolution after its divergence from tetrapods (Figure 1). The large paralogous segments observed in different fish genomes might correspond to remnants of this whole-genome duplication that have been maintained after rediploidization. Two round(s) of tetraploidization/rediploidization (the 1-2-4 rule or 2R hypothesis) might also have occurred earlier during the evolution of the vertebrate lineage before the split between tetrapods and ray-finned fishes (Meyer and Schartl (1999) and references therein), and much more recent independent events of polyploidization have been detected in different fish sublineages including the salmonids (Figure 1; Venkatesh (2003) and references therein). However, the existence of rounds of tetraploidization/rediploidization in the early evolution of vertebrates, and subsequently in ray-finned fishes (the 1-2-4-8 hypothesis, Meyer and Schartl, 1999), is difficult to demonstrate unambiguously because (i) the duplication events are ancient, (ii) the genome has been rediploidized, (iii) in most cases, one of the duplicates has been lost during evolution and (iv) differential evolutionary rates within a pair of paralogs frequently obscure their phylogenetic relationship with orthologs from other species. Therefore, the involvement of successive tetraploidization events in vertebrate evolution is still a matter of intense debate, and an alternative hypothesis involves more regional events of DNA duplication (Robinson-Rechavi *et al.*, 2001b; Seoighe, 2003). Strikingly, some gene families like *hox* or *egfr* (epidermal growth factor receptor) almost perfectly recapitulate the presumed duplication history in fish and other vertebrates. While only one *egfr* gene is present in nonvertebrate animals like the fruit fly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans*, mammals have four *egfr*-related genes, hence supporting the 1-2-4 hypothesis. Teleost fishes have seven *egfr*-related genes, an observation consistent with an additional event of genome duplication followed by the loss of one gene. The much more recent duplication of *egfrb* having generated the *Xmrk* oncogene demonstrates that fish genomes continue to duplicate their genes by more regional events (Volff and Schartl, 2003; Gómez *et al.*, 2004). The same mathematics holds true for the *hox* clusters (Amores *et al.*, 2004).

At least 500 fish-specific ancient pairs of paralogs are present in the genome of Torafugu and Tetraodon (Christoffels *et al.*, 2004; Jaillon *et al.*, 2004; Vandepoele *et al.*, 2004). Even if this value might be an underestimation, this indicates that most genes are not present as pairs of duplicates but rather as single-copy genes in fish. Therefore, if a whole-genome duplication has taken place during

the evolution of the ray-finned lineage, one copy has been subsequently lost within most pairs of paralogs. This is consistent with the observation that, for the vast majority of gene duplicates, one copy evolves as a pseudogene through degenerative mutations (nonfunctionalization) and/or is eliminated because of its dispensability (Lynch and Conery (2000) and references therein).

Nevertheless, independently of the mechanisms by which they have been generated, the evolutionary scenario behind the persistence of hundreds of functional paralog pairs over hundreds of millions of years of evolution in fish is an extremely interesting question, since the duplication of genetic information is thought to be a major seed of evolution (Ohno, 1970). Rarely, one duplicate might have acquired by chance a mutation conferring a new positively selected beneficial function (neofunctionalization), the other copy fulfilling alone the original function. Alternatively, persistence of gene duplicates could be due to subfunctionalization, that is, through partitioning of ancestral functions between duplicates after complementary degenerative mutations in different regulatory or structural sequences (the duplication-degeneration-complementation model; Force *et al.*, 1999; Lynch and Force, 2000a). In this case, both copies are preserved, since the presence of both is necessary to perform the original function of the ancestral single-copy gene. In teleosts, the evolution of numerous gene duplicates is consistent with the subfunctionalization model (eg Lister *et al.*, 2001; Serluca *et al.*, 2001; Altschmied *et al.*, 2002; McClintock *et al.*, 2002; Cresko *et al.*, 2003; Yu *et al.*, 2003; Amores *et al.*, 2004). For example, mammals and birds have a unique microphthalmia-associated transcription factor gene *mitf*, from which different isoforms are expressed through the use of different promoter sequences and alternative exons. In contrast, fish have two different *mitf* genes present in species as divergent as zebrafish, pufferfish and platyfish. Interestingly, the two *mitf* genes in fish each encode one of the different isoforms that are generated from the single gene in 'higher' vertebrates. Hence, the two *mitf* genes are required together to perform the functions of the unique *mitf* gene in mammals and birds. This partitioning of functions is associated with the degeneration of isoform-specific exons and regulatory sequences (Lister *et al.*, 2001; Altschmied *et al.*, 2002).

It is quite difficult to assess with certainty which mechanisms have been at the origin of the maintenance of paralogs in fish. For example, function partitioning might not be primarily responsible for the persistence of the duplicates but might have arisen subsequently during evolution. In addition, the evolution of several paralog pairs does not completely fit any of the three major simple evolutionary models (non-, neo- and subfunctionalization) since they present divergent functions in teleost and tetrapods (see for instance Winkler *et al.*, 2003b). More information is therefore required about ancestral gene function in the fish lineage *before* the proposed genome duplication for a better understanding of the evolutionary mechanism behind the preservation of hundreds of gene duplicates in teleost fish genomes.

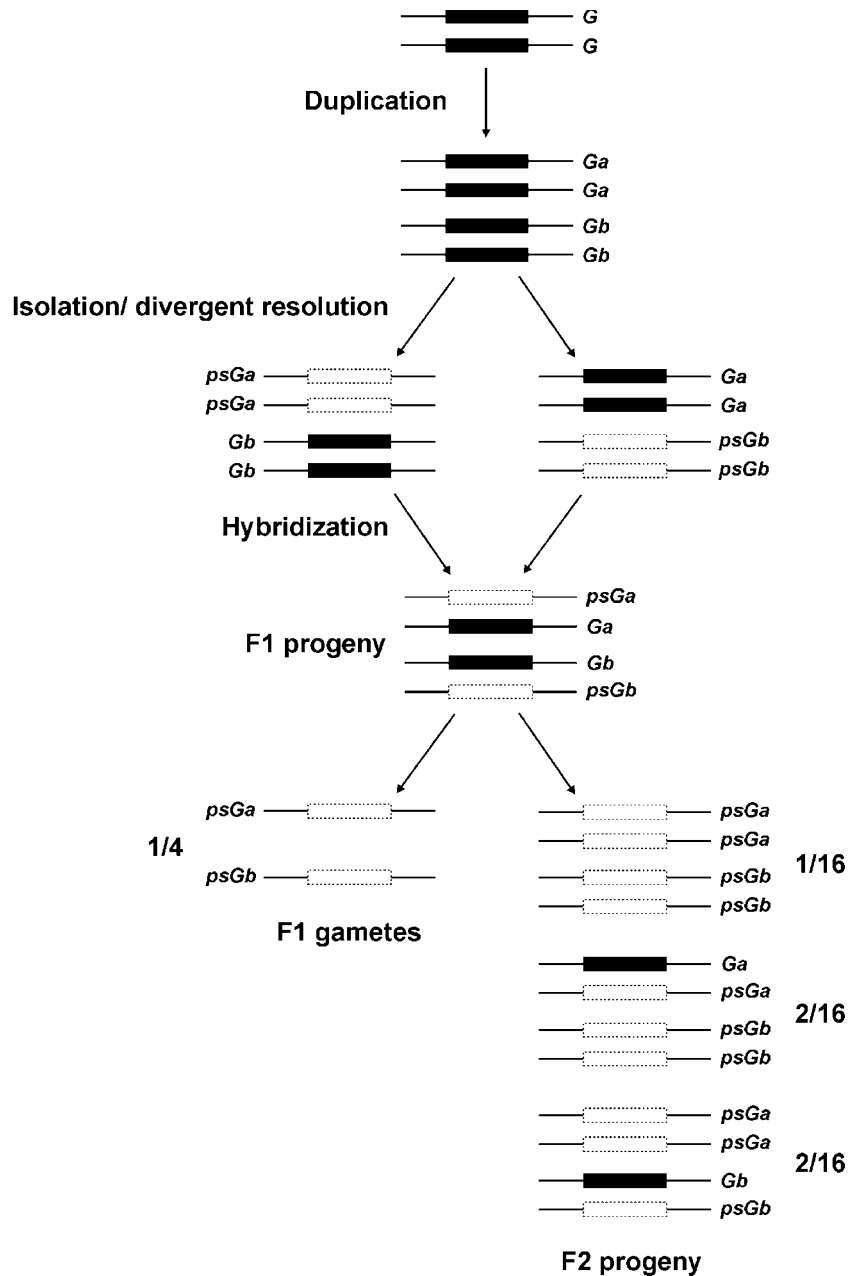
## Gen(om)e duplication and speciation in fish

Genome duplication/rediploidization leads to a massive duplication of genetic information. This might allow the

occurrence of evolutionary novelties necessary for major transitions in evolution and might favour the formation of new species (Ohno, 1970). In addition, divergent resolution of gene duplicates might play an important role in genomic incompatibility between species leading to reduced fertility and/or viability of interspecific hybrids (Werth and Windham, 1991; Lynch and Conery, 2000; Lynch and Force, 2000b; Taylor *et al.*, 2001b; Postlethwait *et al.*, 2004). Imagine the presence, after one round of genome duplication/rediploidization, of two paralogs of the gene *G* (*Ga* and *Gb*), with redundant functions and located on different chromosomes (Figure 2). After geographic isolation between two populations and divergent resolution, *Ga* might be nonfunctionalized through deleterious mutations (pseudogene *psGa*) or simply lost in one population, and *Gb* might do the same in the second population (pseudogene *psGb*). F1 hybrids between both populations would be (*Ga/psGa*; *Gb/psGb*). If *G* is essential for gamete function, 25% of the gametes produced by F1 hybrids will be nonfunctional (no functional copy of the *G* gene) (Lynch and Conery, 2000; Figure 2). If not, crossing between F1 individuals will generate about 6% (1/16) of F2 individuals without any functional *G* gene. In addition, haploinsufficiency (when only one functional allele of *G* is not sufficient to support its normal function) might occur in 25% of the progeny (Figure 2). If divergent resolution occurs for multiple different pairs of duplicates generated by whole-genome duplication, this will result in the passive build-up of reproductive postmating isolation without affecting intraspecific fitness (Lynch and Conery, 2000).

A role of divergent resolution in speciation does not obligatorily involve gene silencing or deletion and can be well conciliated with the subfunctionalization model (Lynch and Conery, 2000). For instance, if the ancestral single-copy gene *G* has two important functions (functions 1 and 2), each involving a specific regulatory sequence, reciprocal divergent partitioning might occur after duplication and geographic isolation (Figure 3). In one population, *Ga* will perform function 1 and *Gb* function 2, while in the second population *Ga* will be responsible for function 2 and *Gb* for function 1. Half of the gametes of the F1 progeny will have genes only for one function. This means that 25% of the gametes will be nonfunctional if one of these functions is important for the gametes, and 50% if both are involved. If not, 12.5% (2/16) of individuals in the F2 progeny will completely lack one function, and as much as 50% might show haploinsufficiency in either function 1 or function 2 (Figure 3). Here again, divergent partitioning for different gene duplicates would result in reproductive isolation and speciation.

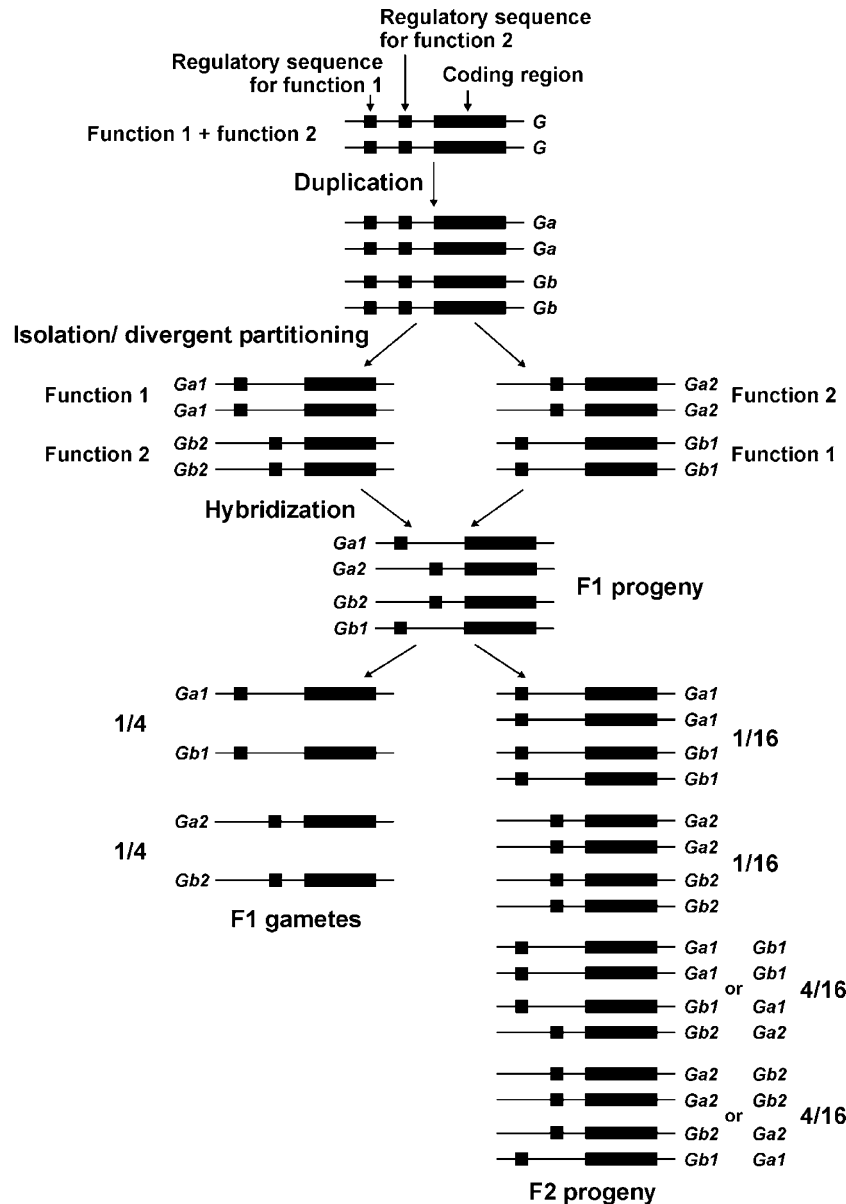
Is there any evidence for divergent resolution in the fish lineage? Several instances of inactivation or loss of a gene duplicate in one species but not in the other have been described in fish, for example, in duplicated *hox* gene clusters (Amores *et al.*, 2004). Demonstrating reciprocal differential loss of duplicates in two divergent fish lineages might be more difficult. If *Ga* is lost in species 1 and *Gb* in species 2, *Gb* of species 1 and *Ga* of species 2 may just look like true orthologs, and additional phylogenetic analyses and mapping experiments will be necessary to demonstrate their paralogy. Putative examples of reciprocal loss have been already



**Figure 2** Divergent resolution of gene duplicates. Black boxes show functional genes, white boxes pseudogenes (ps).

detected through comparative analysis of the genome of medaka and zebrafish (Naruse *et al*, 2004b). Two related helix-loop-helix transcription factor genes called *hey1* and *hey2* are present in the genome of fish and mammals. The *hey1* gene has been duplicated during the early evolution of the ray-fin fish lineage (Winkler *et al*, 2003a). Both *hey1a* and *hey1b* have been maintained in the pufferfishes *T. nigroviridis* and *T. rubripes*, but *hey1b* has been apparently lost in the zebrafish. On the other hand, *hey2* could not be detected in *T. nigroviridis*, and is extremely divergent in *T. rubripes* compared to the well-conserved *hey2* gene of the zebrafish (Winkler *et al*, 2003a). This situation might correspond to a form of reciprocal divergent resolution, with *hey1b* possibly compensating for the loss or extreme divergence of *hey2* in pufferfishes.

Cases of divergent partitioning of subfunctions have also been described in teleost fish. In the rare examples like *mitfa/mitfb* or *sox9a/sox9b*, for which gene duplicate expression and function have been examined in divergent fish species, the major partitioning of ancestral gene functions appears to be ancient (Lister *et al*, 2001; Altschmied *et al*, 2002; Cresko *et al*, 2003). This is manifested by paralog-specific subfunctions conserved in divergent fishes. However, species-specific differences in expression indicative of lineage-specific partitioning have also been found for *sox9a* and *sox9b* between zebrafish and stickleback, suggesting that this phenomenon might indeed be involved in teleost diversification (Cresko *et al*, 2003). Clearly, traditional and functional comparative genomics between pufferfishes, zebrafish, medaka and others will shed new light on the role of



**Figure 3** Divergent subfunction partitioning between gene duplicates.

divergent resolution and partitioning in teleost radiation and subsequent events of speciation.

Gene duplication might also be directly involved in fish diversification by creating 'speciation genes' reducing hybrid fitness. One possible example for that is the *Xmrk* gene (Xiphophorus melanoma receptor tyrosine kinase), corresponding to the dominant *Tu* (Tumour) locus inducing the formation of melanoma in certain interspecific hybrids of the genus *Xiphophorus* (Gordon, 1927; Kosswig, 1928; Anders and Anders, 1978; Schartl, 1995; Meierjohann *et al.*, 2004). After crossing between the *Xmrk*-containing platyfish *Xiphophorus maculatus* and the *Xmrk*-free swordtail *X. helleri*, the F1 hybrid progeny develop noninvasive, superficially spreading nonmalignant pigment cell lesions. Backcrossing of the F1 with the swordtail parent results in the formation of highly invasive melanoma in 25% of the progeny. This corresponds clearly to a reduction of fitness in hybrids, since fishes with melanoma will generally die more or less

rapidly depending on the allele of *Tu* they have inherited. In contrast, the platyfish parent only very rarely develops *Xmrk*-mediated melanoma. *Xmrk* has been recently formed by duplication of the epidermal growth factor receptor gene co-ortholog *egfrb*. Subsequently, mutations in the promoter region have drastically modified its pattern of expression, and mutations conferring ligand-independent constitutive activation have arisen in its extracellular domain (Meierjohann *et al.* (2004) and references therein). According to the classical genetic model, the oncogenic potential of *Xmrk* is repressed by an unlinked tumour suppressor locus called *R* (Anders and Anders, 1978; Schartl, 1995; Meierjohann *et al.*, 2004). The allele of *R* present in the swordtail is unable to repress *Xmrk* (or *R* is simply absent from the swordtail), and the progressive elimination of the platyfish *R* allele in hybrids through crossing leads to the derepression of *Xmrk* and to the formation of melanoma. This situation might be consistent with the

Dobzhansky–Muller model of hybrid incompatibility (Dobzhansky, 1970; Orr and Presgraves, 2000; Wu and Ting, 2004). In this model, two populations are derived from an ancestral population with an *AA/BB* genotype (*A* and *B* are two different genes). *A* evolves into *a* in one population (*aa/BB* genotype) and *B* into *b* in the other population (*AA/bb*). When separated in their own populations, the *a* and *b* alleles do not alter fitness. In contrast, when brought together, incompatibility can occur, resulting in a reduction of fitness in *Aa/Bb* hybrids through partial or full sterility or nonviability. In the Xiphophorus model, *a* (*Xmrk*) would have been created by duplication of *A* (*egfrb*); *B* and *b* would correspond to the different alleles of *R* in the platyfish and the swordtail, respectively. Importantly, the occurrence and significance of hybridization between Xiphophorus species under natural conditions remain to be demonstrated.

## Diversity of transposable elements (TEs) in fish

TEs are sequences able to integrate into new sites within genomes. They are classified into two major classes according to their structure and mechanism of transposition (Curcio and Derbyshire, 2003). Sequences requiring for transposition an RNA intermediate reverse-transcribed into complementary DNA (retrotransposition) are called retroelements. They include reverse transcriptase retrotransposons (LTR or non-LTR retrotransposons depending on the presence of flanking long terminal repeats), retroviruses (reverse transcriptase LTR elements with an envelope gene) and various categories of nonautonomous retroelements like the short interspersed nuclear elements (SINEs). Other elements transposing without reverse transcription are called DNA transposable elements.

Mobile elements can disrupt genes. In addition, ectopic homologous recombination between nonallelic copies of TEs can lead to the formation of deletions, duplications, inversions and translocations, and transposition itself can induce various types of rearrangements at the target site (for a review, see Kazazian, 2004). TEs can also be recruited as exons disrupting an open reading frame, or modify the level and specificity of expression of neighbouring resident genes. The contribution of TEs to mutant phenotypes and genetic diseases can vary considerably between different organisms (Kazazian, 1999).

There is no doubt that transposable elements are drivers of genome evolution (Brosius, 2003; Deininger *et al.*, 2003; Kazazian, 2004). They have been involved in chromosome rearrangements during the evolution of a wide variety of organisms, and retrotransposition has generated at least half of the human and mouse genomes. Particularly, retrotransposition has generated intronless copies of cellular genes (retrogenes), some of them, for example, forming a family of Y-chromosomal genes expressed exclusively in the testis and implicated in male fertility in human (Lahn *et al.*, 2002). TE-derived sequences have been frequently recruited during evolution as regulatory and coding sequences for the host genes (Nekrutenko and Li, 2001; Jordan *et al.*, 2003; Van de Lagemaat *et al.*, 2003). Finally, some TEs, like the

telomeric retrotransposons of *Drosophila*, are apparently directly beneficial to their host, and some mobile sequences have even been domesticated to fulfil new cellular functions (Pardue and DeBaryshe, 2003; Brandt *et al.*, in press).

Almost all types of eukaryotic TEs have been described in teleost fish genomes (Aparicio *et al.*, 2002). Some of these elements are capable of natural insertional mutagenesis (Izsvak *et al.*, 1996; Koga *et al.*, 1996). In order to understand the evolution of TEs in the vertebrate lineage, comparisons between teleost fish and mammalian genomes have been performed particularly for reverse transcriptase retroelements (Voff *et al.*, 2003a). Interestingly, numerous retrotransposons present in vertebrates but absent from mammalian genomes have been identified in the genome of different teleost fish species. As many as nine clades (ancient phylogenetic groups of TEs, the origin of which can be traced back prior to vertebrates) of Ty3/Gypsy-like LTR retrotransposons are found in fish (Poulter and Butler, 1998; Voff *et al.*, 2001b, 2003a), while none of them (with the exception of some domesticated sequences, Brandt *et al.*, in press) are present in the genome of mouse and human. Other major groups of reverse transcriptase retrotransposons present in fish but with no functional equivalent in mammals include Ty1/Copia LTR retrotransposons (Voff *et al.*, 2003a), tyrosine recombinase-encoding retrotransposons (Goodwin and Poulter, 2001, 2004), BEL-like LTR retrotransposons (Frame *et al.*, 2001), Uri endonuclease-encoding *Penelope*-like elements (Lyozin *et al.*, 2001; Voff *et al.*, 2001a) and non-LTR retrotransposons with restriction enzyme-like endonuclease (Voff *et al.*, 2001c; Bouneau *et al.*, 2003). Even for non-LTR retrotransposons with apurinic-apyrimidinic endonuclease, which are extremely well represented in mouse and human genomes (Deininger *et al.*, 2003; Kazazian, 2004), more clades are found in fish genomes (five) than in mammals (three) (Poulter *et al.*, 1999; Voff *et al.*, 1999, 2003a). Taken together, as many as 16–23 clades of reverse transcriptase retroelements have been detected in different fish species, while only six clades are present in mouse and human genomes. A similar situation is also observed for some major families of DNA transposable elements (for example, see Poulter *et al.*, 2003). Hence, the diversity of TEs is much higher in teleost fish than in mammalian genomes, and this phenomenon is also observed inside a particular clade of TE (Furano *et al.*, 2004). Strikingly, even the compact genomes of smooth pufferfishes display a higher diversity of TEs than mammalian genomes, despite their low content of repetitive sequence. Evidence for frequent and recent activity has been provided for numerous families of fish TEs, but their copy number is generally much lower in zebrafish and pufferfish than in mammals. Hence, mobile sequences apparently undergo a higher turnover in teleost fish genomes.

The genomic organization of TEs has been extensively analysed, particularly by fluorescent *in situ* hybridization, in the compact genome of the pufferfish *T. nigroviridis* (Dasilva *et al.*, 2002; Bouneau *et al.*, 2003; Fischer *et al.*, 2004). Almost all categories of TEs generally colocalize with other types of repeats (duplicated pseudogenes, minisatellites) in specific heterochromatic regions of the genome. These observations showed that TEs and other repeated elements are generally excluded



from gene-rich regions in *T. nigroviridis*, this underlining the extreme degree of compartmentalization of this compact genome. Hence, the global organization of the genome of the pufferfish is clearly different from that observed in humans, where repeated sequences make up an important fraction of euchromatic DNA, and is more similar to that observed in the fruit fly *D. melanogaster* (Volff *et al.*, 2003a).

## Transposable elements and speciation

Transposable elements might be able to contribute to pre- and postmating reproductive isolation, and therefore might be involved in the formation of new species (Hurst and Schilthuizen, 1998; Hurst and Werren, 2001). TEs are generally active in germ cells, where they might induce insertional mutations and other kinds of rearrangements that could lead to speciation. Fixation of different rearrangements like translocations and inversions in different populations might result in reproductive isolation. In addition, interspecific crossing might activate transposition in hybrids, this leading to hybrid sterility or inviability. Several cases of hybrid dysgenesis involving transposable elements have been described in *Drosophila*. This phenomenon is observed in the progeny of crosses between strains containing multiple functional copies of a particular TE and strains devoid of this active element. For example, the *I* retrotransposon of *D. melanogaster* is repressed and does not transpose in *I* strains containing functional *I* elements, but retrotransposes at very high frequency in the germ line of hybrid females that are produced after crossing *R* females (devoid of active *I* elements) with *I* males (Bucheton *et al.*, 1992). This results in an increased rate of insertions and DNA rearrangements in the germ line of the hybrid females. One important consequence is female sterility, manifested by the nonhatching of most eggs due to early blocking of embryonic development. Hybrid dysgenesis in *Drosophila* can also be mediated by other TEs (Kidwell *et al.*, 1977; Lozovskaya *et al.*, 1990). Derepression of transposition through interspecific hybridization might occur in divergent animal lineages (O'Neill *et al.*, 1998; Labrador *et al.*, 1999).

At the moment, there is no information about a role of TEs in speciation in fish, or about their possible activation in fish hybrids. However, phylogenetic analysis of several retrotransposons from various fish species has revealed the presence of multiple waves of retrotransposition, which might have been associated with speciation events (Volff *et al.*, 2001d). Clearly, the multiple active lineages of TEs present in fish genomes might predispose to rapid speciation. Further studies will be necessary to establish the activity and genomic impact of fish TEs in germ cells and hybrids in order to approach their role in reproductive isolation and species formation.

## Sex chromosome evolution and the diversity of sex determination in fish

Some particular parts of fish genomes are apparently evolving extremely rapidly. This is particularly true for the sex chromosomes, and this phenomenon might be related to the amazing variety of sex determination systems observed in teleosts (for reviews, Baroiller and D'Cotta, 2001; Devlin and Nagahama, 2002; Volff and

Schartl, 2002). All different forms of genetic sex determination have been observed in fish, including both male heterogamety (males are XY and females are XX) and female heterogamety (males are ZZ and females are ZW), autosomal influences and polygenic sex determination. Sex chromosomes can display very variable degrees of molecular differentiation. Sex determination can also be influenced or determined by environmental factors including the temperature and the pH value of the water, or the fish density. Numerous fish species are hermaphrodites, either simultaneous (male and female at the same time), protandrous (first male and then female) or protogynous (first female and then male). In a same fish, different types of sex determination systems can coexist (eg genetic sex determination and influence of temperature in the Nile Tilapia *Oreochromis niloticus*). Different systems of genetic sex determination can be found in the same fish genus (eg *Oreochromis* spp.) and even in the same species (eg *Xiphophorus maculatus*).

The molecular and evolutionary mechanisms driving sex determination and its variability in teleost fish are poorly understood. The *Sry* gene, inducing the male phenotype in human, mouse and other vertebrates, is clearly absent from fish genomes. Importantly, recent studies on the medaka *Oryzias latipes*, a fish with a XX/XY sex determination system, have revealed how rapidly novel master sex-determining genes and sex chromosomes can evolve in teleosts. Using a positional cloning strategy (Matsuda *et al.*, 2002) and a candidate gene approach (Nanda *et al.*, 2002), two different groups have independently identified in this small fish the first master sex-determining gene of a nonmammalian vertebrate. This gene is *dmrt1bY* (aka *DMY*), a Y chromosome-specific duplicate of an autosomal gene called *dmrt1*. *Dmrt1* is a putative transcription factor apparently ubiquitously involved in sex determination/differentiation in vertebrates, and is member of a family of proteins containing a conserved DNA-binding motif called the DM domain (Volff *et al.*, 2003d and cited references). Some DM domain proteins are involved in the induction of sexual dimorphism in divergent invertebrates including flies and nematodes (Zarkower, 2002). DM domain genes other than *dmrt1* have been identified in fish and mammals, and several of them might be involved in gonad development in the mouse (Brunner *et al.*, 2001; Kondo *et al.*, 2002; Kim *et al.*, 2003; Winkler *et al.*, 2004).

Medaka males have two types of *dmrt1* genes, the autosomal *dmrt1* and the Y-specific *dmrt1bY*. *Dmrt1bY* has probably been formed by a large transchromosomal duplication from linkage group 9 onto another autosome, which became the neo-Y-chromosome by this way (Nanda *et al.*, 2002; Volff and Schartl, 2002; Schartl, 2004). Other genes were included in this duplication, but, in contrast to *dmrt1bY*, they all subsequently degenerated (Nanda *et al.*, 2002). Hence, *dmrt1bY* is apparently the only functional gene in the Y-specific part of the sex chromosomes of the medaka, strongly suggesting that it indeed corresponds to the master sex determining gene. Its expression pattern is also consistent with a role in sex determination: *dmrt1bY* is expressed only in male embryos, and expression occurs prior to the morphological differentiation of gonads. In adults, transcripts are found exclusively in the Sertoli cells of the testis. Finally, natural mutations in *dmrt1bY* result in XY sex-reversed

females (Matsuda *et al.*, 2002). However, the existence of spontaneous sex-reversed XX males in the medaka indicates that a full male phenotype can also occasionally develop in the absence of *dmrt1bY* (Nanda *et al.*, 2003).

The high degree of sequence identity between the autosomal *dmrt1* gene and the Y-specific *dmrt1bY* suggests a recent origin for the master sex-determining gene of the medaka (Nanda *et al.*, 2002). This was confirmed by evolutionary analyses, and *dmrt1bY* was detected only in a very restricted number of *Oryzias* species (Kondo *et al.*, 2003; Matsuda *et al.*, 2003; Veith *et al.*, 2003). Hence, this gene is not the universal master sex-determining gene in teleost fish (Voff *et al.*, 2003b), and the gene(s) driving sexual dimorphism remain(s) to be discovered for the vast majority of teleost fish species.

No sex-linked markers and no sex chromosomes have been identified so far in the zebrafish and smooth pufferfishes, explaining why alternative models like salmonids, platyfish, tilapia and sticklebacks are necessary to analyse sex determination and sex chromosome evolution in fish. There is no doubt that the master sex-determining gene of these fishes will be identified by positional cloning. This will shed new light on the molecular mechanisms driving the evolution of sex determination and sex chromosomes in fish. In salmonids (male heterogamety), comparative mapping of sex-linked microsatellite markers has already shown that Arctic charr, brown trout, Atlantic salmon and rainbow trout have evolved different sex chromosomes (Woram *et al.*, 2003). Sex-linked markers have been found in the Nile tilapia *Oreochromis niloticus* (XX/XY) (Lee *et al.*, 2003) and the blue tilapia *O. aureus* (ZW/ZZ) (Lee *et al.*, 2004), and the putative sex chromosomes have been identified by synaptonemal complex analysis (for a review, Griffin *et al.*, 2002). In the threespine stickleback, sequencing of X- and Y-specific bacterial artificial chromosome clones from the sex determination region revealed many sequence differences between X and neo-Y chromosomes (Peichel *et al.*, 2004). In the platyfish *Xiphophorus maculatus*, a species with three sex chromosomes (X, Y and W), megabase-sized bacterial artificial chromosome contigs covering the sex-determining region of the X and Y chromosomes have been constructed and partially sequenced (Voff and Schartl, 2001; Froschauer *et al.*, 2002, unpublished). As observed for the Y chromosome-specific region of both medaka and threespine stickleback (Nanda *et al.*, 2002; Peichel *et al.*, 2004), the sex determination region of the platyfish displays a high level of genomic instability characterized by frequent transpositions, duplications and deletions (Voff *et al.*, 2003c). Genes present in this region frequently undergo mutations and rearrangements. This phenomenon is associated with a high genetic variability of traits like pigmentation, melanoma phenotype or puberty, which are controlled by gene loci closely linked to the master sex-determining gene in the platyfish (Voff and Schartl, 2001). Whether the genomic plasticity of sex-determining regions is directly implicated in the variability of sex determination systems is still an open question.

## Sex chromosome evolution and speciation

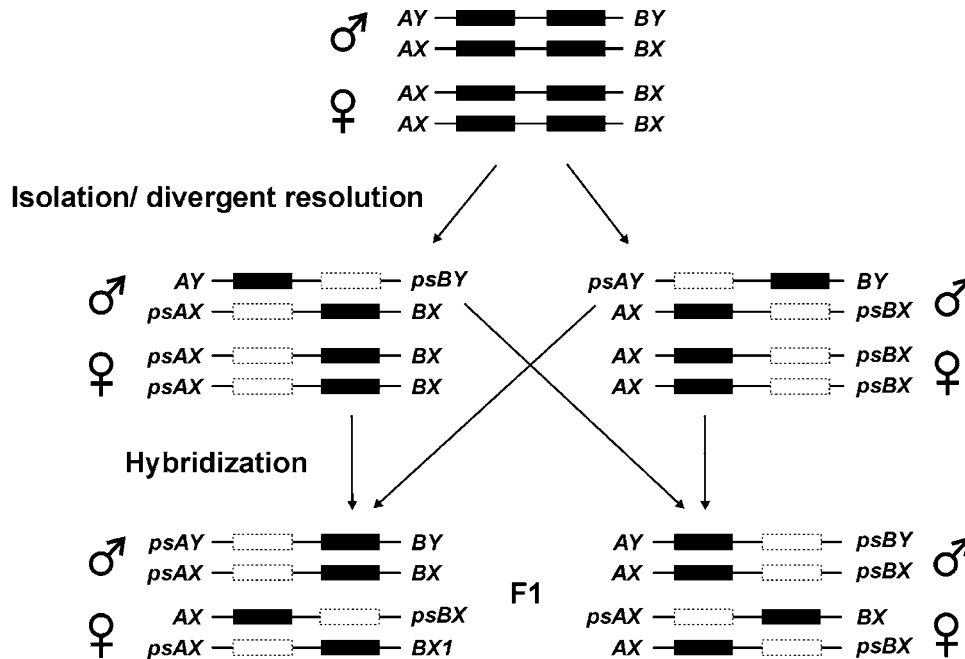
Speciation is intimately associated with the evolution of sex- and reproduction-related traits (eg mating behaviour, fertilization, spermatogenesis, sex determination),

which are frequently controlled by gene loci located on the sex chromosomes. The modification of visual mating cues is probably involved in the establishment of prezygotic barriers between closely related species (prezygotic isolation; Coyne and Orr, 1998). Particularly, colour pattern is a central feature of fish behaviour and evolution, which can serve as mate recognition signals and evolve by sexual selection. African cichlid radiations have been strongly influenced by sexual selection, which principally resulted in the diversification of male colour patterns (Danley and Kocher (2001) and references therein). Numerous examples of sex chromosomal pigmentation loci have been described in different fish species. This is also the case for species of the genus *Xiphophorus* (Voff and Schartl, 2001). In the guppy *Poecilia reticulata*, many of the gene loci controlling the polymorphic male colour patterns involved in mate choice are located in or near the nonrecombining sex-determining region of the Y chromosome (Brooks (2002) and cited references; Lindholm *et al.*, 2004).

Sex-determining regions are apparently very unstable in some fishes, and this phenomenon might account for the high polymorphism generally affecting traits controlled by sex chromosomal loci (Voff and Schartl, 2001). Hence, the rapid divergence of gene loci involved in mate choice within a sex-determining region might speed up prezygotic isolation between two populations. This might affect not only genes involved in pigmentation but also genes playing a role, for example, in sexual maturity, since differences in the time of reproduction might also induce prezygotic isolation (Coyne and Orr (1998) and references therein). Interestingly, a highly polymorphic locus controlling the onset of sexual maturity is closely linked to or located inside the unstable sex-determining region in *Xiphophorus* (Voff and Schartl (2001) and references therein).

In addition, creation of a neo-sex chromosome might disrupt the linkage between the master sex-determining gene and genes involved in mate choice. This might occur, for example, through transposition of the master sex-determining gene from a sex chromosome onto an autosome, as suggested in salmonids (Woram *et al.*, 2003). Another possibility is the creation of a novel master sex-determining gene on an autosome, as observed in the medaka (Nanda *et al.*, 2002). If the mate choice gene, for instance, a Y-linked pigmentation pattern gene, is not controlled directly or indirectly by the master sex-determining itself, the pattern will not be sex-specific anymore, or may even disappear from males. This might initiate the isolation between the population with the ancestral Y chromosome from the population with the neo-Y chromosome. Speciation models based on sexual selection on sex-determining genes associated with colour polymorphisms and incorporating the lability of sex determination in fish have been proposed for the African cichlids (Lande *et al.*, 2001). Divergence of sex determination systems might also lead to hybrid progeny with a reduced fitness due to sex ratio distortion (Voff and Schartl, 2001), and selection for nonbiased sex ratios has been proposed to be involved in sympatric speciation in cichlids (Seehausen *et al.*, 1999).

Differences in sex chromosomes might be a potential source of postzygotic isolation too. For example, if two species have developed different heteromorphic pairs of sex chromosomes, abnormalities in meiosis pairing



**Figure 4** Divergent resolution of Y- and X-chromosomal alleles of genes with male-specific functions. Black boxes show functional genes, white boxes pseudogenes (ps).

might occur in hybrids, with hybrid sterility as a possible consequence. In addition, divergent resolution between X- and Y-chromosomal alleles of genes with male-specific functions (for example, involved in fertility) in different populations might lead to male sterility (or inviability, if one of the functions is essential for the survival of the males) (Lynch and Force, 2000b; Figure 4). If genes *A* and *B* are located both on the proto-sex chromosomes of an ancestral population, *A* and *B* might become Y- and X-specific, respectively, in one population, while they would be X- and Y-specific, respectively, in the other population. All males in the F1 progeny will consequently be sterile, since they completely lack either *A* or *B* (Figure 4). This might correspond to an early stage of reproductive isolation. If this phenomenon only occurs with gene *A*, male sterility will be observed only in one cross, but not in the reciprocal one (Lynch and Force, 2000b). This model is consistent with Haldane's rule, stating that when the F1 hybrid offspring of a cross between a male parent from one line and a female parent from the other line is sterile although otherwise healthy, it will tend to be of the heterogametic sex (Haldane, 1922). Divergent gene loss would be favoured by the high frequency of rearrangement-mediated nonfunctionalization affecting (at least some) sex-determining regions in fish (Volf *et al.*, 2003c). The same model can apply if *A* and *B* correspond to different subfunctions of the same gene, which are then divergently resolved in the two different populations.

## Conclusions

Teleost fish provide an outstanding model to study a multitude of questions related to evolution. This may be linked to the apparent considerable plasticity of their genome, manifested, for example, by a high variability in genome size and chromosome number (Venkatesh, 2003).

Fish genomes also have intrinsic characteristics, which might have been involved in the formation of the amazing diversity of species observed in the teleost lineage. Particularly, there is now substantial evidence that an ancient event of genome duplication (tetraploidization) has provided the evolutionary framework for the diversification of gene functions and for speciation in fish. Completion and comparison of the sequence of the genome of different fish species, as well as subsequent functional genomics approaches will allow the understanding of why hundreds of paralogs have been maintained over hundreds of millions of years of evolution in fish genomes. Such analyses will also provide new information concerning the evolution and evolutionary impact of transposable elements in fish genomes. Clearly, the multiple families of active transposable elements present in teleosts potentially represent powerful evolutionary factors, which might also have played an important role in speciation. Finally, the frequent switching between different sex determination systems and the rapid evolution of sex chromosomes might also be linked to the formation of new species. Many more comparative studies will be necessary to understand why sex determination is so variable in fish in contrast to the situation observed in birds and mammals.

Studies on fish will probably help to better understand the evolution of our own genome and characterize the functions of its genes. Hundreds of new genes and regulatory sequences have already been identified through sequence comparison between pufferfish and human genomes. There is also no doubt that the knowledge gained from analyses in zebrafish and medaka will shed a new light on organogenesis in vertebrates, even if there is already evidence for important differences between human and fish. Sub-functionalized gene paralogs are highly relevant for such

experiments, since they will allow to analyse separately gene functions performed by a unique gene in human (Volff and Schartl, 2003; Postlethwait *et al*, 2004). Finally, analysis of sex determination in fish might allow the discovery of alternative strategies to compensate for the announced destruction of the Y chromosome in humans (Marshall Graves, 2002; Volff *et al*, 2003b). Lessons from current fish models have shown that one species is rarely representative of the complete teleost fish lineage. Multiple comparative analyses will be necessary to understand the evolution of this very diverse group of animals.

## Acknowledgements

I am very grateful to Manfred Schartl, Christoph Winkler and Alexander Froschauer (University of Würzburg) for stimulating discussions and critical reading of the manuscript. Our work is supported by the BioFuture programme of the Bundesministerium für Bildung und Forschung (BMBF) and by the Deutsche Forschungsgemeinschaft (DFG). I also thank the people who kindly provided me with fish pictures.

## References

- Agresti JJ, Seki S, Cnaani A, Poompuang S, Hallerman EM, Umiel N *et al* (2000). Breeding new strains of tilapia: development of an artificial center of origin and linkage map based on AFLP and microsatellite loci. *Aquaculture* **185**: 43–56.
- Albertson RC, Strelman JT, Kocher TD (2003). Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *Proc Natl Acad Sci USA* **100**: 5252–5257.
- Altschmied J, Delfgaauw J, Wilde B, Duschl J, Bouneau L, Volff JN *et al* (2002). Subfunctionalization of duplicate *mitf* genes associated with differential degeneration of alternative exons in fish. *Genetics* **161**: 259–267.
- Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A *et al* (1998). Zebrafish *hox* clusters and vertebrate genome evolution. *Science* **282**: 1711–1714.
- Amores A, Suzuki T, Yan YL, Pomeroy J, Singer A, Amemiya C *et al* (2004). Developmental roles of pufferfish *Hox* clusters and genome evolution in ray-fin fish. *Genome Res* **14**: 1–10.
- Anders A, Anders F (1978). Etiology of cancer as studied in the platyfish-swordtail system. *Biochim Biophys Acta* **516**: 61–95.
- Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, Dehal P *et al* (2002). Whole genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**: 1301–1310.
- Baroiller JF, D'Cotta H (2001). Environment and sex determination in farmed fish. *Comp Biochem Physiol C Toxicol Pharmacol* **130**: 399–409.
- Borowsky R, Wilkens H (2002). Mapping a cave fish genome: polygenic systems and regressive evolution. *J Hered* **93**: 19–21.
- Bouneau L, Fischer C, Ozouf-Costaz C, Froschauer A, Jaillon O, Coutanceau JP *et al* (2003). An active non-LTR retrotransposon with tandem structure in the compact genome of the pufferfish *Tetraodon nigroviridis*. *Genome Res* **13**: 1686–1695.
- Brandt J, Veith AM, Volff JN. A family of neofunctionalized Ty3/gypsy retrotransposon genes in mammalian genomes. *Cytogenet Genome Res* (in press).
- 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**: 265–268.
- Brooks R (2002). Variation in female mate choice within guppy populations: population divergence, multiple ornaments and the maintenance of polymorphism. *Genetica* **116**: 343–358.
- Brosius J (2003). The contribution of RNAs and retroposition to evolutionary novelties. *Genetica* **118**: 99–116.
- Brunner B, Hornung U, Shan Z, Nanda I, Kondo M, Zend-Ajus E *et al* (2001). Genomic organization and expression of the doublesex-related gene cluster in vertebrates and detection of putative regulatory regions for *DMRT1*. *Genomics* **77**: 8–17.
- Bucheton A, Vaury C, Chaboissier MC, Abad P, Pelisson P, Simonelig M (1992). I elements and the *Drosophila* genome. *Genetica* **86**: 175–190.
- 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**: 1146–1151.
- Clark MS, Edwards YJ, Peterson D, Clifton SW, Thompson AJ, Sasaki M *et al* (2003). Fugu ESTs: new resources for transcription analysis and genome annotation. *Genome Res* **13**: 2747–2753.
- Coyne JA, Orr HA (1998). The evolutionary genetics of speciation. *Philos Trans R Soc Lond B Biol Sci* **353**: 287–305.
- Cresko WA, Yan YL, Baltrus DA, Amores A, Singer A, Rodriguez-Mari A *et al* (2003). Genome duplication, subfunction partitioning, and lineage divergence: *Sox9* in stickleback and zebrafish. *Dev Dyn* **228**: 480–489.
- Curcio MJ, Derbyshire KM (2003). The outs and ins of transposition: from *mu* to kangaroo. *Nat Rev Mol Cell Biol* **4**: 865–877.
- Danley PD, Kocher TD (2001). Speciation in rapidly diverging systems: lessons from Lake Malawi. *Mol Ecol* **10**: 1075–1086.
- Danzmann RG, Gharbi K (2001). Gene mapping in fishes: a means to an end. *Genetica* **111**: 3–23.
- Dasilva C, Hadji H, Ozouf-Costaz C, Nicaud S, Jaillon O, Weissenbach J *et al* (2002). Remarkable compartmentalization of transposable elements and pseudogenes in the heterochromatin of the *Tetraodon nigroviridis* genome. *Proc Natl Acad Sci USA* **99**: 13636–13641.
- Deininger PL, Moran JV, Batzer MA, Kazazian Jr HH (2003). Mobile elements and mammalian genome evolution. *Curr Opin Genet Dev* **13**: 651–658.
- Devlin RH, Nagahama Y (2002). Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture* **208**: 191–364.
- Dobzhansky T (1970). *Genetics of the Evolutionary Process*. Columbia University Press: New York.
- Driever W, Solnica-Krezel L, Schier AF, Neuhauss SC, Malicki J, Stemple DL *et al* (1996). A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123**: 37–46.
- Fischer C, Bouneau L, Coutanceau JP, Weissenbach J, Volff JN, Ozouf-Costaz C (2004). Global heterochromatic colocalization of transposable elements with minisatellites in the compact genome of the pufferfish *Tetraodon nigroviridis*. *Gene* **336**: 175–183.
- Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**: 1531–1545.
- Frame IG, Cutfield JF, Poulter RT (2001). New BEL-like LTR-retrotransposons in *Fugu rubripes*, *Caenorhabditis elegans*, and *Drosophila melanogaster*. *Gene* **263**: 219–230.
- Froschauer A, Körting C, Katagiri T, Aoki T, Asakawa S, Shimizu N *et al* (2002). Construction and initial analysis of bacterial artificial chromosome (BAC) contigs from the sex-determining region of the platyfish *Xiphophorus maculatus*. *Gene* **295**: 247–254.
- Furano AV, Duvernell DD, Boissinot S (2004). L1 (LINE-1) retrotransposon diversity differs dramatically between mammals and fish. *Trends Genet* **20**: 9–14.
- Furutani-Seiki M, Sasado T, Morinaga C, Suwa H, Niwa K, Yoda H *et al* (2004). A systematic genome-wide screen for mutations affecting organogenesis in Medaka, *Oryzias latipes*. *Mech Dev* **121**: 647–658.
- Furutani-Seiki M, Wittbrodt J (2004). Medaka and zebrafish, an evolutionary twin study. *Mech Dev* **121**: 629–637.

- Geisler R, Rauch GJ, Baier H, van Bebber F, Bross L, Dekens MP *et al* (1999). A radiation hybrid map of the zebrafish genome. *Nat Genet* **23**: 86–89.
- Gilbey J, Verspoor E, McLay A, Houlihan D (2004). A microsatellite linkage map for Atlantic salmon (*Salmo salar*). *Anim Genet* **35**: 98–105.
- Gómez A, Volf JN, Hornung U, Schartl M, Wellbrock C (2004). Identification of a second *egfr* gene in *Xiphophorus* uncovers an expansion of the epidermal growth factor receptor family in fish. *Mol Biol Evol* **21**: 266–275.
- Goodwin TJ, Poulter RT (2001). The DIRS1 group of retrotransposons. *Mol Biol Evol* **18**: 2067–2082.
- Goodwin TJ, Poulter RT (2004). A new group of tyrosine recombinase-encoding retrotransposons. *Mol Biol Evol* **21**: 746–759.
- Gordon M (1927). The genetics of viviparous top-minnow *Platypoecilus*: the inheritance of two kinds of melanophores. *Genetics* **12**: 253–283.
- Griffin DK, Harvey SC, Campos-Ramos R, Ayling LJ, Bromage NR, Masabanda JS *et al* (2002). Early origins of the X and Y chromosomes: lessons from tilapia. *Cytogenet Genome Res* **99**: 157–163.
- Grunwald DJ, Eisen JS (2002). Headwaters of the zebrafish – emergence of a new model vertebrate. *Nat Rev Genet* **3**: 717–724.
- Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA *et al* (1996). The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **123**: 1–36.
- Haldane JBS (1922). Sex-ratio and unisexual sterility in hybrid animals. *J Genet* **12**: 101–109.
- He C, Chen L, Simmons M, Li P, Kim S, Liu ZJ (2003). Putative SNP discovery in interspecific hybrids of catfish by comparative EST analysis. *Anim Genet* **34**: 445–448.
- Hoegg S, Brinkmann H, Taylor JS, Meyer A (2004). Phylogenetic timing of the fish-specific duplication correlates with the diversification of teleost fish. *J Mol Evol* **59**: 190–203.
- Hukriede NA, Joly L, Tsang M, Miles J, Tellis P, Epstein JA *et al* (1999). Radiation hybrid mapping of the zebrafish genome. *Proc Natl Acad Sci USA* **96**: 9745–9750.
- Hurst D, Schilthuizen M (1998). Selfish elements and speciation. *Heredity* **80**: 2–8.
- Hurst GD, Werren JH (2001). The role of selfish genetic elements in eukaryotic evolution. *Nat Rev Genet* **2**: 597–606.
- Izsvak Z, Ivics Z, Garcia-Estefania D, Fahrenkrug SC, Hackett PB (1996). DANA elements: a family of composite, tRNA-derived short interspersed DNA elements associated with mutational activities in zebrafish (*Danio rerio*). *Proc Natl Acad Sci USA* **93**: 1077–1081.
- Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, Mauceli E *et al* (2004). Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **431**: 946–957.
- Jordan IK, Rogozin IB, Glazko GV, Koonin EV (2003). Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet* **19**: 68–72.
- Kazazian Jr HH (1999). An estimated frequency of endogenous insertional mutations in humans. *Nat Genet* **22**: 130.
- Kazazian Jr HH (2004). Mobile elements: drivers of genome evolution. *Science* **303**: 1626–1632.
- Kazianis S, Morizot DC, McEntire BB, Nairn RS, Borowsky RL (1996). Genetic mapping in *Xiphophorus* hybrid fish: assignment of 43 AP-PCR/RAPD and isozyme markers to multipoint linkage groups. *Genome Res* **6**: 280–289.
- Kelly PD, Chu F, Woods IG, Ngo-Hazelett P, Cardozo T, Huang H *et al* (2000). Genetic linkage mapping of zebrafish genes and ESTs. *Genome Res* **10**: 558–567.
- Khoo G, Lim MH, Suresh H, Gan DK, Lim KF, Chen F *et al* (2003). Genetic linkage maps of the guppy (*Poecilia reticulata*): assignment of RAPD markers to multipoint linkage groups. *Mar Biotechnol* **5**: 279–293.
- Khorasani MZ, Hennig S, Imre G, Asakawa S, Palczewski S, Berger A *et al* (2004). A first generation physical map of the medaka genome in BACs essential for positional cloning and clone-by-clone based genomic sequencing. *Mech Dev* **121**: 903–913.
- Kidwell MG, Kidwell JF, Sved JA (1977). Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility, and male recombination. *Genetics* **86**: 813–833.
- Kim S, Kettlewell JR, Anderson RC, Bardwell VJ, Zarkower D (2003). Sexually dimorphic expression of multiple doublesex-related genes in the embryonic mouse gonad. *Gene Expr Patterns* **3**: 77–82.
- Kimura T, Jindo T, Narita T, Naruse K, Kobayashi D, Shin-I T *et al* (2004). Large-scale isolation of ESTs from medaka embryos and its application to medaka developmental genetics. *Mech Dev* **121**: 915–932.
- Kocher TD (2004). Adaptive evolution and explosive speciation: the cichlid fish model. *Nat Rev Genet* **5**: 288–298.
- Kocher TD, Lee WJ, Sobolewska H, Penman D, McAndrew B (1998). A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). *Genetics* **148**: 1225–1232.
- Koga A, Suzuki M, Inagaki H, Bessho Y, Hori H (1996). Transposable element in fish. *Nature* **383**: 30.
- Kondo M, Froschauer A, Kitano A, Nanda I, Hornung U, Volf JN *et al* (2002). Molecular cloning and characterization of *DMRT* genes from the medaka *Oryzias latipes* and the platyfish *Xiphophorus maculatus*. *Gene* **295**: 213–222.
- Kondo M, Nanda I, Hornung U, Asakawa S, Shimizu N, Mitani H *et al* (2003). Absence of the candidate male sex-determining gene *dmrt1b(Y)* of medaka from other fish species. *Curr Biol* **13**: 416–420.
- Kosswig C (1928). Über Kreuzungen zwischen den Teleostiern *Xiphophorus helleri* und *Platypoecilus maculatus*. *Z Indukt Abstammungs-Vererbungslehre* **47**: 150–158.
- Labrador M, Farre M, Utzet F, Fontdevila A (1999). Interspecific hybridization increases transposition rates of *Osva1*. *Mol Biol Evol* **16**: 931–937.
- Lahn BT, Tang ZL, Zhou J, Barndt RJ, Parvinen M, Allis CD *et al* (2002). Previously uncharacterized histone acetyltransferases implicated in mammalian spermatogenesis. *Proc Natl Acad Sci USA* **99**: 8707–8712.
- Lande R, Seehausen O, van Alphen JJ (2001). Mechanisms of rapid sympatric speciation by sex reversal and sexual selection in cichlid fish. *Genetica* **112–113**: 435–443.
- Lee BY, Hulata G, Kocher TD (2004). Two unlinked loci controlling the sex of blue tilapia (*Oreochromis aureus*). *Heredity* **92**: 543–549.
- Lee BY, Penman DJ, Kocher TD (2003). Identification of a sex-determining region in Nile tilapia (*Oreochromis niloticus*) using bulked segregant analysis. *Anim Genet* **34**: 379–383.
- Lindholm AK, Brooks R, Breden F (2004). Extreme polymorphism in a Y-linked sexually selected trait. *Heredity* **92**: 156–162.
- Lister JA, Close J, Raible DW (2001). Duplicate *mitf* genes in zebrafish: complementary expression and conservation of melanogenic potential. *Dev Biol* **237**: 333–344.
- Liu Z, Karsi A, Li P, Cao D, Dunham R (2003). An AFLP-based genetic linkage map of channel catfish (*Ictalurus punctatus*) constructed by using an interspecific hybrid resource family. *Genetics* **165**: 687–694.
- Loh YH, Christoffels A, Brenner S, Hunziker W, Venkatesh B (2004). Extensive expansion of the Claudin gene family in the teleost fish, *Fugu rubripes*. *Genome Res* **14**: 1248–1257.
- Lozovskaya ER, Scheinker VS, Evgen'ev MB (1990). A hybrid dysgenesis syndrome in *Drosophila virilis*. *Genetics* **126**: 619–623.
- Lynch M, Conery JS (2000). The evolutionary fate and consequences of duplicate genes. *Science* **290**: 1151–1155.
- Lynch M, Force A (2000a). The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**: 459–473.

- Lynch M, Force A (2000b). The origin of interspecific genomic incompatibility via gene duplication. *Am Nat* **156**: 590–605.
- Lyozin GT, Makarova KS, Velikodvorskaia VV, Zelentsova HS, Khechumian RR, Kidwell MG *et al* (2001). The structure and evolution of *Penelope* in the virilis species group of *Drosophila*: an ancient lineage of retroelements. *J Mol Evol* **52**: 445–456.
- Marshall Graves JA (2002). The rise and fall of SRY. *Trends Genet* **18**: 259–264.
- Matsuda M, Nagahama Y, Shinomiya A, Sato T, Matsuda C, Kobayashi T *et al* (2002). *DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature* **417**: 559–563.
- Matsuda M, Sato T, Toyazaki Y, Nagahama Y, Hamaguchi S, Sakaizumi M (2003). *Oryzias latipes* has *DMY*, a gene that is required for male development in the medaka, *O. latipes*. *Zool Sci* **20**: 159–161.
- McClintock JM, Kheirbek MA, Prince VE (2002). Knockdown of duplicated zebrafish *hoxb1* genes reveals distinct roles in hindbrain patterning and a novel mechanism of duplicate gene retention. *Development* **129**: 2339–2354.
- Meierjohann S, Scharl M, Voff JN (2004). Genetic, biochemical and evolutionary facets of *Xmrk*-induced melanoma formation in the fish *Xiphophorus*. *Comp Biochem Physiol C Toxicol Pharmacol* **138**: 281–289.
- Meyer A, Scharl M (1999). Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* **11**: 699–704.
- Moen T, Hoyheim B, Munck H, Gomez-Raya L (2004). A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. *Anim Genet* **35**: 81–92.
- Morizot DC, Nairn RS, Simhambhatla P, Della Coletta L, Trono D, Chovanec L *et al* (2001). *Xiphophorus* genetic linkage map: beginnings of comparative gene mapping in fishes. *Mar Biotechnol* **3**: S153–S161.
- Nanda I, Hornung U, Kondo M, Schmid M, Scharl M (2003). Common spontaneous sex-reversed XX males of the medaka *Oryzias latipes*. *Genetics* **163**: 245–251.
- Nanda I, Kondo M, Hornung U, Asakawa S, Winkler C, Shimizu A *et al* (2002). A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. *Proc Natl Acad Sci USA* **99**: 11778–11783.
- Naruse K, Hori H, Shimizu N, Kohara Y, Takeda H (2004a). Medaka genomics: a bridge between mutant phenotype and gene function. *Mech Dev* **121**: 619–628.
- Naruse K, Tanaka M, Mita K, Shima A, Postlethwait J, Mitani H (2004b). A medaka gene map: the trace of ancestral vertebrate proto-chromosomes revealed by comparative gene mapping. *Genome Res* **14**: 820–828.
- Nekrutenko A, Li WH (2001). Transposable elements are found in a large number of human protein-coding genes. *Trends Genet* **17**: 619–621.
- Nelson JS (1994). *Fishes of the World*, 3rd edn. John Wiley and Sons: New York.
- Nichols KM, Young WP, Danzmann RG, Robison BD, Rexroad C, Noakes M *et al* (2003). A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*). *Anim Genet* **34**: 102–115.
- Ohno S (1970). *Evolution by Gene Duplication*. Springer Verlag: New York.
- O'Neill RJ, O'Neill MJ, Graves JA (1998). Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* **393**: 68–72.
- Orr HA, Presgraves DC (2000). Speciation by postzygotic isolation: forces, genes and molecules. *BioEssays* **22**: 1085–1094.
- Pardue ML, DeBaryshe PG (2003). Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. *Annu Rev Genet* **37**: 485–511.
- Peichel CL, Nereng KS, Ohgi KA, Cole BL, Colosimo PF, Buerkle CA *et al* (2001). The genetic architecture of divergence between threespine stickleback species. *Nature* **414**: 901–905.
- Peichel CL, Ross JA, Matson CK, Dickson M, Grimwood J, Schmutz J *et al* (2004). The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Curr Biol* **14**: 1416–1424.
- Postlethwait J, Amores A, Cresko W, Singer A, Yan YL (2004). Subfunction partitioning, the teleost radiation and the annotation of the human genome. *Trends Genet* **20**: 481–490.
- Postlethwait JH, Johnson SL, Midson CN, Talbot WS, Gates M, Ballinger EW *et al* (1994). A genetic linkage map for the zebrafish. *Science* **264**: 699–703.
- Postlethwait JH, Woods IG, Ngo-Hazelett P, Yan YL, Kelly PD, Chu F *et al* (2000). Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Res* **10**: 1890–1902.
- Poulter R, Butler M (1998). A retrotransposon family from the pufferfish (fugu) *Fugu rubripes*. *Gene* **215**: 241–249.
- Poulter R, Butler M, Ormandy J (1999). A LINE element from the pufferfish (fugu) *Fugu rubripes* which shows similarity to the CR1 family of non-LTR retrotransposons. *Gene* **227**: 169–179.
- Poulter RT, Goodwin TJ, Butler MI (2003). Vertebrate helitrons and other novel Helitrons. *Gene* **313**: 201–212.
- Rasmussen AS, Arnason U (1999). Molecular studies suggest that cartilaginous fishes have a terminal position in the piscine tree. *Proc Natl Acad Sci USA* **96**: 2177–2182.
- Rexroad III CE, Lee Y, Keele JW, Karamycheva S, Brown G, Koop B *et al* (2003). Sequence analysis of a rainbow trout cDNA library and creation of a gene index. *Cytogenet Genome Res* **102**: 347–354.
- Rise ML, von Schalburg KR, Brown GD, Mawer MA, Devlin RH, Kuipers N *et al* (2004). Development and application of a salmonid EST database and cDNA microarray: data mining and interspecific hybridization characteristics. *Genome Res* **14**: 478–490.
- Robinson-Rechavi M, Marchand O, Escriva H, Bardet PL, Zelus D, Hughes S *et al* (2001a). Euteleost fish genomes are characterized by expansion of gene families. *Genome Res* **11**: 781–788.
- Robinson-Rechavi M, Marchand O, Escriva H, Laudet V (2001b). An ancestral whole-genome duplication may not have been responsible for the abundance of duplicated fish genes. *Curr Biol* **11**: R458–R459.
- Roest Crollius H, Jaillon O, Bernot A, Dasilva C, Bouneau L, Fischer C *et al* (2000). Estimate of human gene number provided by genome-wide analysis using *Tetraodon nigroviridis* DNA sequence. *Nat Genet* **25**: 235–238.
- Scharl M (1995). Platyfish and swordtails: a genetic system for the analysis of molecular mechanisms in tumor formation. *Trends Genet* **11**: 185–189.
- Scharl M (2004). A comparative view on sex determination in medaka. *Mech Dev* **121**: 639–645.
- Seehausen O, van Alphen JJM, Lande R (1999). Color polymorphism and sex ratio in a cichlid fish as an incipient stage in sympatric speciation by sexual selection. *Ecol Lett* **2**: 367–378.
- Seoighe C (2003). Turning the clock back on ancient genome duplication. *Curr Opin Genet Dev* **13**: 636–643.
- Serluca FC, Sidow A, Mably JD, Fishman MC (2001). Partitioning of tissue expression accompanies multiple duplications of the Na<sup>+</sup>/K<sup>+</sup> ATPase alpha subunit gene. *Genome Res* **11**: 1625–1631.
- Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, Jonsson B *et al* (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**: 717–723.
- 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**: 382–390.

- Taylor JS, Van de Peer Y, Braasch I, Meyer A (2001a). Comparative genomics provides evidence for an ancient genome duplication event in fish. *Philos Trans R Soc Lond B Biol Sci* **356**: 1661–1679.
- Taylor JS, Van de Peer Y, Meyer A (2001b). Genome duplication, divergent resolution and speciation. *Trends Genet* **17**: 299–301.
- Thorgaard GH, Bailey GS, Williams D, Buhler DR, Kaattari SL, Ristow SS *et al* (2002). Status and opportunities for genomics research with rainbow trout. *Comp Biochem Physiol B Biochem Mol Biol* **133**: 609–646.
- Van de Lagemaat LN, Landry JR, Mager DL, Medstrand P (2003). Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet* **19**: 530–536.
- Vandepoele K, De Vos W, Taylor JS, Meyer A, Van de Peer Y (2004). 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 USA* **101**: 1638–1643.
- Veith AM, Froschauer A, Körtling C, Nanda I, Hanel R, Schmid M *et al* (2003). Cloning of the *dmrt1* gene of *Xiphophorus maculatus*: *dmY/dmrt1Y* is not the master sex-determining gene in the platyfish. *Gene* **317**: 59–66.
- Venkatesh B (2003). Evolution and diversity of fish genomes. *Curr Opin Genet Dev* **13**: 588–592.
- Verheyen E, Salzburger W, Snoeks J, Meyer A (2003). Origin of the superflock of cichlid fishes from Lake Victoria, East Africa. *Science* **300**: 325–329.
- Volff JN, Bouneau L, Ozouf-Costaz C, Fischer C (2003a). Diversity of retrotransposable elements in compact pufferfish genomes. *Trends Genet* **19**: 674–678.
- Volff JN, Hornung U, Scharl M (2001a). Fish retroposons related to the *Penelope* element of *Drosophila virilis* define a new group of retrotransposable elements. *Mol Genet Genomics* **265**: 711–720.
- Volff JN, Kondo M, Scharl M (2003b). Medaka *dmY/dmrt1Y* is not the universal primary sex-determining gene in fish. *Trends Genet* **19**: 196–199.
- Volff JN, Körtling C, Altschmied J, Duschl J, Sweeney K, Wichert K *et al* (2001b). *Jule* from the fish *Xiphophorus* is the first complete vertebrate Ty3/Gypsy retrotransposon from the Mag family. *Mol Biol Evol* **18**: 101–111.
- Volff JN, Körtling C, Froschauer A, Sweeney K, Scharl M (2001c). Non-LTR retrotransposons encoding a restriction enzyme-like endonuclease in vertebrates. *J Mol Evol* **52**: 351–360.
- Volff JN, Körtling C, Froschauer A, Zhou Q, Wilde B, Schultheis C *et al* (2003c). The *Xmrk* oncogene can escape nonfunctionalization in a highly unstable subtelomeric region of the genome of the fish *Xiphophorus*. *Genomics* **82**: 470–479.
- Volff JN, Körtling C, Meyer A, Scharl M (2001d). Evolution and discontinuous distribution of *Rex3* retrotransposons in fish. *Mol Biol Evol* **18**: 427–431.
- Volff JN, Körtling C, Sweeney K, Scharl M (1999). The non-LTR retrotransposon *Rex3* from the fish *Xiphophorus* is widespread among teleosts. *Mol Biol Evol* **16**: 1427–1438.
- Volff JN, Scharl M (2001). Variability of genetic sex determination in poeciliid fishes. *Genetica* **111**: 101–110.
- Volff JN, Scharl M (2002). Sex determination and sex chromosome evolution in the medaka, *Oryzias latipes*, and the platyfish, *Xiphophorus maculatus*. *Cytogenet Genome Res* **99**: 170–177.
- Volff JN, Scharl M (2003). Evolution of signal transduction by gene and genome duplication in fish. *J Struct Funct Genom* **3**: 139–150.
- Volff JN, Zarkower D, Bardwell VJ, Scharl M (2003d). Evolutionary dynamics of the DM domain gene family in metazoans. *J Mol Evol* **57**: S241–S249.
- Werth CR, Windham MD (1991). A model for divergent allopatric speciation of polyploid pteridophytes resulting from silencing of duplicate gene expression. *Am Nat* **137**: 515–526.
- Winkler C, Elmasri H, Klamt B, Volff JN, Gessler M (2003a). Characterization of hey bHLH genes in teleost fish. *Dev Genes Evol* **213**: 541–553.
- Winkler C, Hornung U, Kondo M, Neuner C, Duschl J, Shima A *et al* (2004). Developmentally regulated and non-sex-specific expression of autosomal *dmrt* genes in embryos of the Medaka fish (*Oryzias latipes*). *Mech Dev* **121**: 997–1005.
- Winkler C, Schafer M, Duschl J, Scharl M, Volff JN (2003b). Functional divergence of two zebrafish midline growth factors following fish-specific gene duplication. *Genome Res* **13**: 1067–1081.
- Wittbrodt J, Meyer A, Scharl M (1998). More genes in fish? *BioEssays* **20**: 511–515.
- Wittbrodt J, Shima A, Scharl M (2002). Medaka – a model organism from the far East. *Nat Rev Genet* **3**: 53–64.
- Woods IG, Kelly PD, Chu F, Ngo-Hazelett P, Yan YL, Huang H *et al* (2000). A comparative map of the zebrafish genome. *Genome Res* **10**: 1903–1914.
- Woram RA, Gharbi K, Sakamoto T, Hoyheim B, Holm LE, Naish K *et al* (2003). Comparative genome analysis of the primary sex-determining locus in salmonid fishes. *Genome Res* **13**: 272–280.
- Woram RA, McGowan C, Stout JA, Gharbi K, Ferguson MM, Hoyheim B *et al* (2004). A genetic linkage map for Arctic char (*Salvelinus alpinus*): evidence for higher recombination rates and segregation distortion in hybrid *versus* pure strain mapping parents. *Genome* **47**: 304–315.
- Wu CI, Ting CT (2004). Genes and speciation. *Nat Rev Genet* **5**: 114–122.
- Yu WP, Brenner S, Venkatesh B (2003). Duplication, degeneration and subfunctionalization of the nested synapsin-*Timp* genes in Fugu. *Trends Genet* **19**: 180–183.
- Zarkower D (2002). Invertebrates may not be so different after all. *Novartis Found Symp* **244**: 115–126.