Chapter 16
Two Rounds of Whole-Genome Duplication: Evidence and Impact on the Evolution of Vertebrate Innovations

Cristian Cañestro

Abstract The origin and evolution of the vertebrates have been linked to the study of genome duplications since Susumo Ohno ventured the 2R-hypothesis, suggesting that the successful diversification of complex vertebrates was facilitated by polyploidization in the stem vertebrate ancestor due to two rounds of whole-genome duplication (2R-WGD). This chapter first reviews evidence supporting Ohno's 2R-hypothesis and gathers information about the timing and mechanisms underlying the 2R-WGD. Second, this chapter describes the impact of the 2R-WGD on the evolution of the vertebrate genome structure, gene number, and the evolutionary dynamics of the functional fate of vertebrate ohnologs (paralogous genes that originated by WGD) in comparison with non-vertebrate chordate gene homologs. Finally, this review discusses the functional consequences of the 2R-WGD on the origin and evolution of vertebrate innovations compared with urochordates and cephalochordates, paying special attention to the origin of neural crest cells, placodes, and the big complex brain, key features that probably facilitated the transition from ancestral filter-feeding non-vertebrate chordates to voracious vertebrate predators. Currently available data, however, seem to suggest that these putative key features were present to at least some extent in stem Olfactores; hence, the impact of the 2R-WGD may not be related to the immediate origin of vertebrate innovations, but to the subsequent diversification of a wide variety of complex structures that facilitated the successful radiation of vertebrates.

C. Cañestro (✉)
Departament de Genètica, Facultat de Biologia, Universitat de Barcelona,
Av. Diagonal, 643, edifici Prevosti, 2a planta, 08028 Barcelona, Spain
e-mail: canestro@ub.edu

16.1 Introduction

The study of the origin and evolution of vertebrates, the subphylum to which we belong, has stood at the crossroad between genome evolution and molecular developmental biology since the late 1960s, when Susumo Ohno published his famous work on *Evolution by Gene Duplication* and proposed his hypothesis about the pivotal role of genome duplication in the origin of vertebrates and their diversification (Ohno et al. 1968; Ohno 1970). Vertebrates comprise all animals that have a backbone and include mammals, birds, reptiles, amphibians, fishes, and agnathans—the jawless lampreys and hagfishes. Vertebrates together with urochordates (tunicates) form the Olfactores, which together with cephalochordates (amphioxus or lancelets) constitute the chordates (Fig. 16.1). All chordates share a common basic body plan at least during the larval stage of their life cycle, consisting of a notochord running through a post-anal tail, with a dorsal hollow nerve cord, longitudinal blocks of muscle along the notochord, and ciliated pharyngeal gill slits (Brusca and Brusca 2002). Recent phylogenomic analyses have dethroned cephalochordates from the long-assumed position as sister group of the vertebrates; this position is now occupied by urochordates, which include ascidians, larvaceans and thaliaceans (Wada et al. 2006; Oda et al. 2002; Bourlat et al. 2006; Delsuc et al. 2006; Putnam et al. 2008) (Fig. 16.1).

Ohno's 2R-hypothesis was based on comparative analyses of genome sizes and isozyme complexity among chordate taxa. Ohno found that basally divergent chordate subphyla had smaller genomes and less isozyme complexity than vertebrate lineages. This observation led him to suggest that the combination of tandem gene duplication and in particular an octoploidization event involving two rounds of whole-genome duplication were key to the invertebrate–vertebrate transition, and for the subsequent successful vertebrate diversification (Ohno et al. 1968).

Ohno was one of the pioneers in conceiving the evolutionary significance of whole-genome duplication (see also Chap. 1, this volume). Ohno emphasized the importance of gene duplication as probably the main source of raw genetic material for the evolution of new gene functions (reviewed in Taylor and Raes 2004). In Ohno's classical model, one of the duplicated genes retains the original function whereas its duplicate either disappears by accumulation of detrimental mutations (called pseudogenization or nonfunctionalization) or it is preserved after gaining advantageous mutations that confer positively selected novel functions (neofunctionalization) (Ohno 1970; Nowak et al. 1997; Force et al. 1999). The duplication, degeneration, complementation hypothesis (or DDC model) suggests a third possibility for duplicate gene preservation: subfunctionalization, the complementary partitioning of ancestral structural and/or regulatory subfunctions between two duplicate genes, so that the sum of their functions provides at least that of the original pre-duplication gene (Force et al. 1999). The DDC model predicts that subfunctionalized genes will have lower pleiotropy than the original pre-duplicated gene and lower evolutionary constraints, and thereby will be more permissive to the
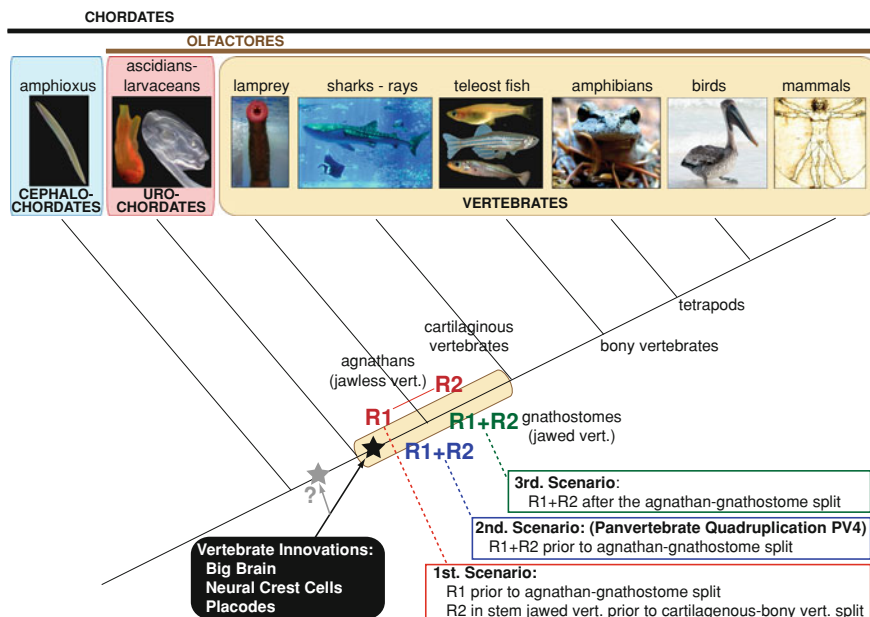
Fig. 16.1 Evolutionary tree of chordates representing the three possible scenarios for the timing of the 2R-WGD and the origins for vertebrate innovations. Evidence suggests that the 2R-WGD might have had a significant impact on the diversification (*black star*) of vertebrate innovations, including structures derived from neural crest cells and placodes, as well as the development of a big complex brain in the stem vertebrate. But whether the 2R-WGD was crucial for the evolutionary origin of these structures remains unclear, and that the hypothesis that the origin of these innovations dated back to at least stem olfactores (*gray star*) cannot be dismissed. (Lamprey picture courtesy of Juan Pascual-Anaya)

accumulation of mutations that might confer novel functions. The acquisition of new functions is favored if the duplication affects the entire genome at once, as opposed to multiple individual gene duplications, because when entire gene networks are duplicated, gene stoichiometry is maintained, and therefore deleterious gene dosage effects can be counteracted (Birchler and Veitia 2007, 2010; Van de Peer et al. 2009; Makino and McLysaght 2010; see also Chap. 2, this volume). Genes that originated by gene duplication are called paralogs. Genes that have been duplicated via genome duplication, however, are a special type of paralogs referred as *ohnologs*, a term suggested by Wolfe (2000) in honor of Ohno's contribution. This term is useful because of the special properties that ohnologs possess at their birth compared to duplications that arise by other local mechanisms such as unequal crossing-over, tandem gene duplication, or retrotransposition.

The sudden creation and evolution of ohnologs by the 2R-WGD that occurred in the stem vertebrate lineage has been suggested as one of the potential key events underlying the increase of morphological complexity, facilitating the acquisition of genetic and developmental innovations of vertebrates (Shimeld and Holland 2000; Aburomia et al. 2003). Genome duplication doubles the number of genes,

many of which have the chance of evolving 'novel' functions that might provide new selectable advantages promoting species diversification (Lynch et al. 2001; Van de Peer et al. 2009). Some vertebrate species that have undergone recent polyploidization, such as the frog *Xenopus laevis* that experienced tetraploidization ~ 40 million years ago (Hellsten et al. 2007), show a higher adaptability to a variety of different environments, such as drought, salt, cold, and disease resistance, than closely related diploid species, such as *Silurana tropicalis* (for further details on frog polyploidization see also Chap. 18, this volume). Interestingly, only a limited number of retained ohnologs present evidence of neofunctionalization or subfunctionalization in *X. laevis*, suggesting that additional selective mechanisms might act on preserving gene duplicates that could promote species diversification (Chain and Evans 2006; Semon and Wolfe 2008).

In addition to the evolutionary significance of novel fates of duplicate genes on species biology, another mechanism that might contribute to species diversification after genome duplication is reciprocal ohnolog loss between different populations, a concept known as 'divergent resolution' that can lead to reproductive isolation (Lynch and Force 2000). Reciprocal ohnolog loss is likely to occur in the period of relaxed selection that duplicate genes experience while they are functionally redundant (Werth and Windham 1991; Lynch and Conery 2000; Lynch and Force 2000; Scannell et al. 2006; Taylor et al. 2001; Semon and Wolfe 2007). The divergent resolution of gene redundancies, such that one population loses one ohnolog copy while the second population loses the other ohnolog copy, leads to chromosomal restructuring such that gametes produced by hybrid individuals can be completely lacking in functional genes for a duplicate pair. In addition to the isolation due to reciprocal gene losses, this model can be further expanded to isolation due to independent processes of gene duplicate subfunctionalization between two populations, in which hybrids will lack one or more subfunctions (Force et al. 1999; Lynch and Force 2000). Hence, large-scale reciprocal ohnolog loss and independent subfunctionalization of ohnologs can be the cause of reproductive isolation of two populations after polyploidization, favoring genetic divergence of these newly incipient future species. This hypothesis is supported by the analyses of both fish and angiosperm lineages that have undergone polyploidization and include more species diversity (e.g. salmonids, catostomids, eudicots, grasses) than their sister groups that did not go through polyploidization and include a lower number of species (Nelson 1994; Ferris et al. 1979; Soltis et al. 2009). Recent integrated approaches of comparative genomics and gene expression analyses in teleosts, however, provide limited evidence supporting the significance of differential ohnolog loss in reproductive isolation and diversification (Kassahn et al. 2009) (see also Chap. 17, this volume).

Many studies have tackled the central question of whether or not the 2R-WGD had a significant impact on the origin of vertebrate innovations and their subsequent diversification. This chapter first reviews evidence supporting the 2R-hypothesis and information regarding the timing and potential mechanisms underlying the 2R-WGD in vertebrates. Second, this chapter examines the impact that the 2R-WGD may have had on the evolution of vertebrate genome structure, number of

genes, and the fate of retained ohnologs in comparison with non-vertebrate chordate paralogs. Finally, this chapter discusses how the 2R-WGD might have affected the origin and evolution of vertebrate innovations, with special emphasis on the vertebrate big complex brain, and structures derived from neural crest cells and placodes. Recent data, however, suggest that neural crest cells and placodes could already have been present in stem chordates. Hence, the impact of the 2R-WGD cannot be related to the origin of neural crest cells and placodes, but it could be related to their subsequent diversification and development of a wide variety of complex structures.

16.2 Supporting Evidence for the 2R-Hypothesis

If two rounds of genome duplication (2R-WGD) have occurred, we would expect the presence of many paralogs (ohnologs) in conserved, syntenic genomic regions, which are known as paralogons (Coulier et al. 2000) (or ohnologons (Gout et al. 2009)). Paralogons, therefore, consist of series of linked (but frequently functionally and phylogenetically unrelated) genes on one chromosome region, many of which have linked paralogs on at least one other chromosome region. The discovery of paralogy groups made of four paralogons in the genome of human and mouse was interpreted as remnants of the two events of tetraploidization that occurred early during vertebrate evolution and therefore provided the earliest strong evidence supporting Ohno's 2R-hypothesis (Lundin 1979, 1993; Pebusque et al. 1998). Possibly one of the best and first examples of a paralogy group supporting the 2R-hypothesis is the case of the four *Hox*-bearing regions on human chromosomes *Hsa*2, *Hsa*7, *Hsa*12, and *Hsa*17 (Fig. 16.2) (Kappen et al. 1989; Bailey et al. 1997; Larhammar et al. 2002; Lundin et al. 2003). In contrast, only a single *Hox* cluster is present in the cephalochordate amphioxus (Garcia-Fernàndez and Holland 1994). This 4:1 ratio is consistent with Ohno's hypothesis of two tetraploidization events after the cephalochordate–vertebrate split (Holland et al. 1994; Sidow 1996; Garcia-Fernandez 2005). In addition to the *Hox* paralogy group, several other similar examples have been identified (e.g. *MHC* (Katsanis et al. 1996), *Tbx* (Ruvinsky and Silver 1997), *G-protein-coupled* receptors (Fredriksson et al. 2003), *ParaHox* clusters (Ferrier et al. 2005), linked receptor *tyrosine kinases* (Siegel et al. 2007), *endothelin* ligands and receptors (Braasch et al. 2009), *Fox* cluster (Wotton and Shimeld 2006), and the *EGF* ligand paralogons (Laisney et al. 2010)).

Several databases of conserved syntenic chromosomal regions between different species are available, and these provide additional evidence of two rounds of genome duplication. Popovici et al. (2001), for instance, identified 14 paralogons containing more than 1600 genes assembled in a human genome paralogy map (http://u119.marseille.inserm.fr/Db/paralogy.html). The ParaDB (http://abi.marseille.inserm.fr/paradb) predicted that the human genome far exceeds 1000 paralogons that contain more than three pairs of duplicated genes (Leveugle et al. 2003).
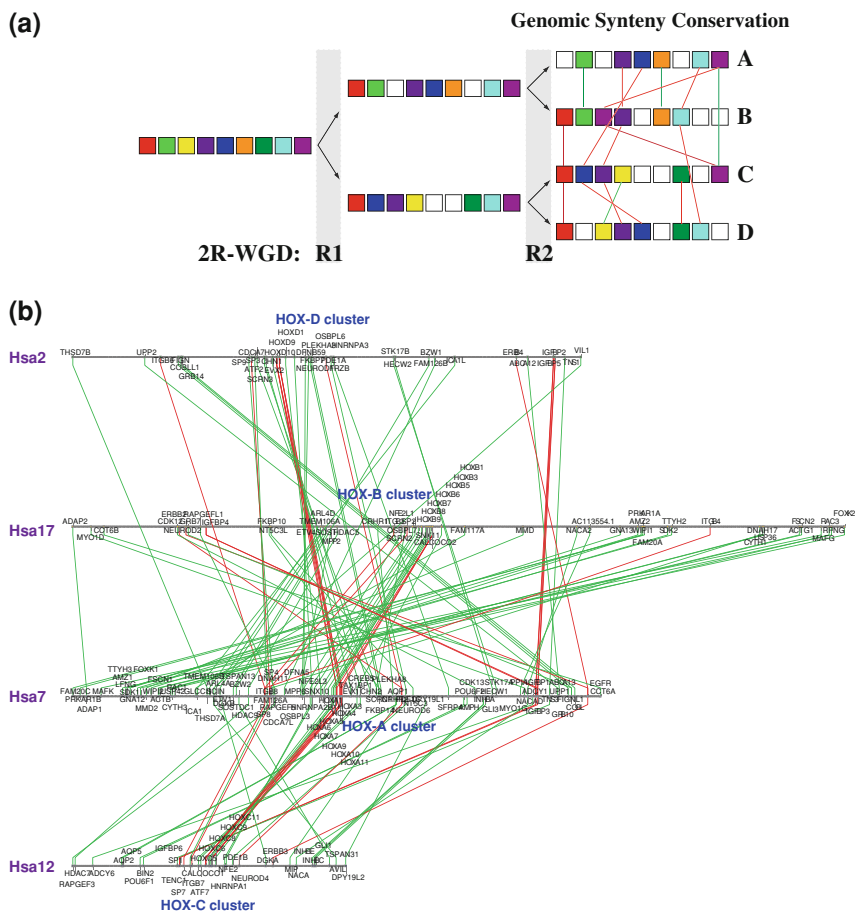
Fig. 16.2 Conserved synteny in the vertebrate genome generated by the 2R-WGD. **a** Simplified representation of a genomic region that has been amplified by the 2R-WGD (R1 and R2) showing conserved genes (*colored boxes*) in four syntenic regions (**a–d**), which have suffered genomic rearrangements and gene loss (*white boxes*) and different degrees of conservation (*green and red lines* label ohnologs preserved in two or more than two regions, respectively). **b** Representation of the four human paralogons containing *Hox* A-D clusters in chromosomes *Hsa*2, *Hsa*7, *Hsa*12, and *Hsa*17, displaying high amounts of conserved synteny between ohnologs in two (*green lines*) or at least three (*red lines*) different paralogons. This representation has been generated using a 100-gene sliding window in the Synteny Database (Catchen et al. 2009)

The Genomicus database v60.01 (http://www.dyogen.ens.fr/genomicus) predicted that 18,228 ancestral vertebrate genes were grouped in 2,642 conserved ancestral synteny blocks with a median N50 size of 5 genes (Muffato et al. 2010). The Synteny Database (http://teleost.cs.uoregon.edu/synteny_db) predicted 231 paralogy clusters with more than 5 genes, and 102 paralogy clusters with more than 10 genes, in a

more rigorous count using a 100-gene sliding window and taking amphioxus genes as the outgroup for paralogy assignment in human (Catchen et al. 2009, 2011).

Until recently, however, available evidence did not permit us to discard the possibility that these groups of paralogous genes originated by multiple independent block duplications, rather than two duplications of the entire genome (Skrabanek and Wolfe 1998; Wolfe 2001; Larhammar et al. 2002). An initial hypothesis, based on extensive phylogenetic analysis and dating of the duplications that produced hundreds of vertebrate gene families, proposed a 'big-bang mode' of sudden large-scale gene origin resulting from two waves of gene duplications, rather than the alternative hypothesis of a constant generation by small-scale duplications (Gu et al. 2002). Wave-I was suggested to consist of tandem or segmental duplications that occurred after the mammalian radiation, and wave-II was interpreted as a rapid increase of paralogs in the early stage of vertebrate evolution after their split from non-vertebrate chordates, consistent with one round of whole-genome duplication (Gu et al. 2002). The first analyses of the human genome draft led to the conclusion that the most parsimonious explanation of the current structure of the human paralogons was a 'big-bang' expansion event by a paleopolyploidy that included the whole genome or substantial sections of it. However, no specific evidence was found for two rounds of polyploidy as opposed to one (Venter et al. 2001; McLysaght et al. 2002; Panopoulou et al. 2003).

Recently, however, several analyses have provided definitive support for the 2R-hypothesis (reviewed in Kasahara 2007). Dehal and Boore (2005) developed an elegant, compelling approach to test the 2R-hypothesis by plotting the genomic map position of only those genes that were duplicated prior to the fish–tetrapod split, which rendered a clear global physical pattern of four-way paralogon organization covering most of the human genome. Dehal and Boore's work therefore provided unmistakable evidence of two distinct rounds of genome duplication during early vertebrate evolution.

Furthermore, the recent sequencing of the whole genome of the chordate amphioxus *Branchiostoma floridae* (sister to all other chordates) provided even more indisputable evidence supporting the 2R-hypothesis (Putnam et al. 2008). Despite the fact that small-region comparison between human, chicken, teleost fish, and amphioxus genomes revealed low gene-order conservation at the local level (microsynteny), striking extensive gene linkage conservation was observed when entire chromosomes were considered (macrosynteny). Syntenic analysis reconstructed 17 chordate linkage groups (CLG) that might represent the proto-chromosomes of the last common chordate ancestor (Putnam et al. 2008). Exhaustive evaluation of the 17 CLGs revealed that most of the human genome (112 segments spanning 2.68 Gb, which is the equivalent of 95 % of the euchromatic genome) was affected by large-scale duplication events that occurred on the stem vertebrate lineage before the teleost/tetrapod split. Analysis of the distribution of the human segments among the 17 CLGs showed that nearly all ancient chordate chromosomes were quadruplicated (Putnam et al. 2008) (Fig. 16.3). This result robustly demonstrated the occurrence of two rounds of genome duplication, corroborating previous lines of evidence based on analysis of

specific regions of interest, such as the *Hox*-bearing regions (Garcia-Fernàndez and Holland 1994; Larhammar et al. 2002) and the major histocompatibility regions (Vienne et al. 2003; Danchin and Pontarotti 2004). Spring (1997) proposed the term "tetralogs" to refer groups of quadruplicated vertebrate genes at four different chromosomal locations formed by the 2R-WGD corresponding to a single invertebrate gene, with all four more similar to each other than to members of the other tetralogy group.

16.2.1 Timing of the Vertebrate 2R-WGD

Analyses of the completely sequenced genome of the cephalochordate amphioxus (Putnam et al. 2008; Holland et al. 2008) and the genomes from various urochordates (Dehal et al. 2002; Small et al. 2007; Denoeud et al. 2010) validated Ohno's hypothesized lower-bound timing for the 2R-WGD as after the split between vertebrates and non-vertebrate chordates. Regarding the upper-bound timing, extensive analysis of gene duplicates (Robinson-Rechavi et al. 2004) and the identification of the four clusters in the genome of the elephant shark suggested that the 2R-WGD took place before the cartilaginous/bony vertebrate split (Venkatesh et al. 2007; Ravi et al. 2009).

Within this time window, the most prevalent hypothesis suggests a scenario in which the first round (R1) occurred before the split between gnathostome and jawless vertebrates, and the second (R2) occurred in the stem jawed vertebrates after their divergence from jawless vertebrates (Fig. 16.1). However, a second scenario proposes that both rounds (R1 + R2) of genome duplication took place before the split between gnathostome and jawless vertebrates (pan-vertebrate quadruplication (PV4) hypothesis (Kuraku et al. 2009)) (Fig. 16.1). Comparative analysis of 55 gene families revealed a common expansion in both jawless and jawed vertebrates, which has been interpreted as evidence supporting this second scenario (Kuraku et al. 2009). Available information from sea lampreys and hagfish does not permit us to discern between these two hypothetical scenarios, because these organisms also appear to have suffered lineage-specific duplications and reciprocal gene losses compared to vertebrates, which together obscure the assessments of orthology/paralogy (reviewed in Kuraku 2008, 2010). For instance, multiple *Hox* gene surveys in different species of sea lampreys and hagfish suggested that extensive independent duplications of *Hox* genes might have occurred during the evolution of jawless vertebrates (Pendleton et al. 1993; Sharman and Holland 1998; Takio et al. 2004; Force et al. 2002; Irvine et al. 2002; Fried et al. 2003; Stadler et al. 2004; Kuraku et al. 2009). Some of the jawless *Hox* clusters might have disintegrated, casting doubt as to the usefulness of *Hox* genes as reliable markers to trace duplications during genome evolution in stem vertebrates (Kuraku 2011). Finally, recent phylogenetic analysis of the degenerated *ParaHox* cluster in hagfish has opened the possibility of a third scenario, in which both rounds (R1 + R2) occurred in stem jawed vertebrates after their divergence
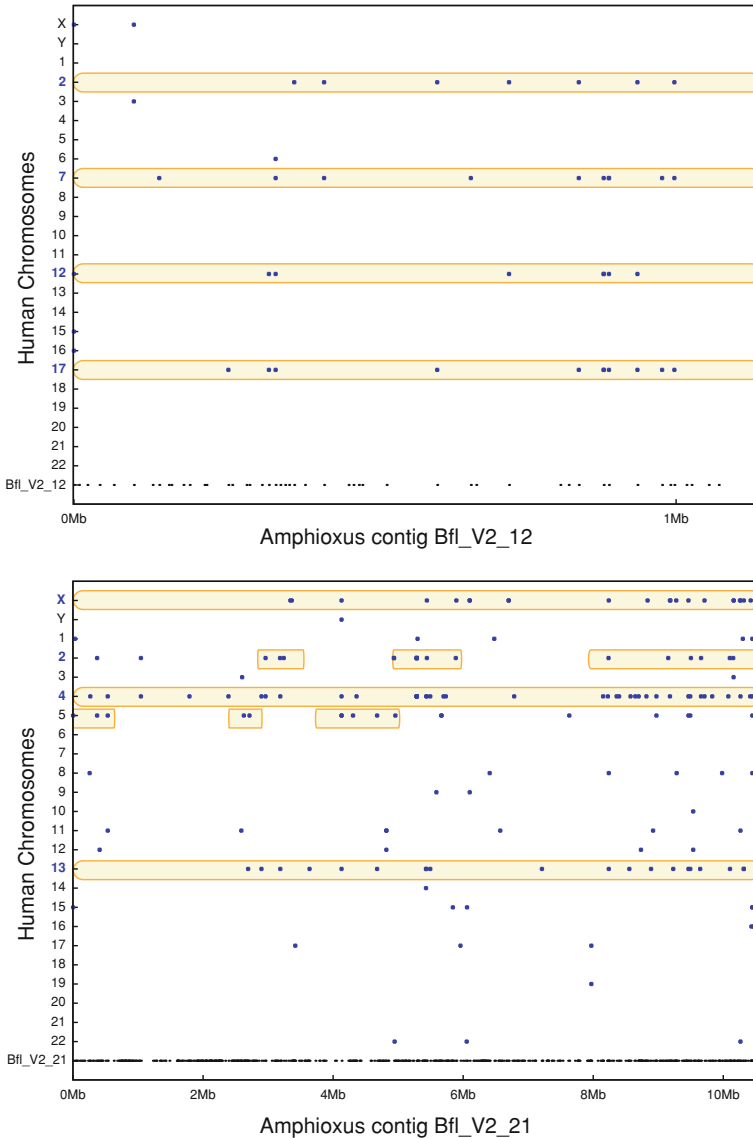
Fig. 16.3 Quadruplicated conserved syntenic pattern between the amphioxus and the human genome as a result of the 2R-WGD. *Dot-plots* display the distribution throughout human chromosomes (y-axes) of human orthologs (*blue dots*) of amphioxus genes (*black dots*) located in two arbitrarily selected genomic regions of approximately 1 Mb (**a**) and 10 Mb (**b**) (x-axes). The *dot-plots* reveal four major human chromosomes (*yellow shadow*) of conserved synteny as the product of the two rounds of whole-genome duplication. In panel (**a**), the four paralogons coincide with the *Hox*-cluster bearing chromosomes 2, 7, 12, and 17, whereas in panel B the four paralogons coincide with the endothelin receptors and *ParaHox*-cluster bearing chromosomes 2/5, 4, 13, and X. The *dot-plots* were generated as described in Canestro et al. (2009) using the Synteny Database (Catchen et al. 2009)

from jawless vertebrates (Furlong et al. 2007) (Fig. 16.1). The validation of this third scenario could have a significant impact on our understanding of vertebrate evolution, because it would imply that the 2R-WGD would have not been important for the origin of vertebrate innovations (i.e. big brain, neural crest cells, and placodes, which clearly exist in jawless vertebrates (Kuratani and Ota 2008; Kuratani 2009)). According to this third scenario, however, the 2R-WGD would have been important for the radiation of gnathostomes into cartilaginous fish, bony fishes, and tetrapods. A solid answer about the timing of the 2R-WGD may have to wait until larger-scale comparisons of the whole-genome organization of hagfish and lampreys are available.

16.2.2 Mechanisms Underlying the Vertebrate 2R-WGD

A question that still remains is how did the stem vertebrate genome become octoploid by two rounds of tetraploidization. There are two main mechanisms of tetraploidization observed in many species of plants and animals (Van de Peer et al. 2009). The first mechanism is allotetraploidy, which occurs when two related but not identical genomes are combined by hybridization of closely related species and associated (often subsequent) genome duplication. In the case of allotetraploidy, the pairs of distinct 'homologous' chromosomes that are sufficiently different due to their separated origin are called *homeologs*. The second mechanism is autotetra-ploidy, which occurs when the genomes are not sufficiently diverged into homeologous sets; autotetraploidy therefore ranges from the combination of genomes of two conspecific individuals (perhaps from different populations) to the combination of identical genomes from a single individual. The genetic attributes of allo- and autotetraploids differ and may have substantial effects at individual, population, and species levels (see also Chap. 2, this volume). Both allotetraploidy and autotetraploidy could be generated by several processes such as: (i) an abnormal non-disjunction of sister chromatids at meiosis; (ii) the uncoupling of mitotic DNA replication and cell division during early development of the germ line (this process, for instance, occurs normally during the endoreplication of megakaryocytic bone-marrow precursors of blood platelets, or during the devel-opment of the oikoblastic epithelia that secrete the house in basal urochordate larvaceans); (iii) potential cell fusion during early embryo development or in germ-line precursors in syncytial gametogenesis (cell fusion is observed naturally, for instance, in skeletal muscle cells and placenta) (reviewed in Storchova and Pellman 2004; Shemer and Podbilewicz 2000).

In the case of allotetraploids, each pair of homologous chromosomes should segregate normally during meiosis, and genetic interchange between homeologous chromosomes is rare. If two consecutive events of allotetraploidization occurred in stem vertebrates, we would predict that in the ideal situation in the absence of gene losses, a phylogenetic tree of homeologs will render a symmetrical (A,B) (C,D) topology (Furlong and Holland 2002).

In autopolyploids, however, meiotic pairing might occur between any of the four identical chromosomes at meiosis I, facilitating genetic interchanges freely among the four alleles, and leading to 'tetrasomic inheritance'. Eventually the alleles, and chromosomes, might diverge, starting a process of diploidization that reestablishes diploidy. Randomly one of the chromosomes will diverge first and no longer form homologous structures, while the other three will keep pairing until another further divergence. Hence, if two consecutive events of autotetraploidy occurred in quick succession (pseudo-octoploidy) in stem vertebrates, we would predict that in the absence of gene losses, gene family phylogenetic trees will likely render asymmetrical (((A,B),C),D) topologies (Furlong and Holland 2002). Because many vertebrate gene families do render asymmetrical tree topologies (Friedman and Hughes 2001; Hughes 1999; Hughes and Friedman 2003), two quick consecutive events of autotetraploidy have been considered a likely mechanism for the 2R-WGD in stem vertebrates (Furlong and Holland 2002; Lynch and Wagner 2009).

16.3 Consequences of the 2R-WGD on the Evolution of Vertebrate Genome Structure

It has been suggested that polyplodization events, at least in plants, can trigger genomic stress associated with major genomic rearrangements, in many cases mediated by a burst of mobilization of transposable elements (Matzke and Matzke 1998; Comai 2000). Transposable elements can be substrates for unequal and illegitimate recombination and can be responsible for a variety of genome reorganizations associated with the transposition, including chromosomal insertions, deletions, inversions, translocations, and duplications. Lineage-specific genome rearrangements mediated by transposable elements might facilitate rapid evolution, reproductive isolation of different populations, and consequently species diversification (Parisod et al. 2010).

Contrary to possible genome reorganization after polyploidization, as noted above for plants, in stem vertebrates, recent work based on proximate gene pair methods and measurement of syntenic clustering conservation found that the 2R-WGD in vertebrates were not followed by an increase of genome rearrangement (Hufton et al. 2008). Unexpectedly, this work measured massive genome rearrangements prior to the 2R-WGD, which has been interpreted as a pre-existing 'disposition' toward genomic structural change (Hufton et al. 2008). Interestingly, in contrast to the archetypal condition that has been described in the organization of particular genomic regions (e.g. *Hox*-cluster region (Garcia-Fernàndez and Holland 1994)), the amphioxus genome structure is not exceptionally well conserved, evolving its own particular type of repetitive elements (e.g. '*mirage*' minisatellites (Cañestro et al. 2002b; Ebner et al. 2010)), undergoing extensive local tandem gene duplications (see section below), and experiencing a moderate rate of synteny loss similar to that of sea urchin or sea

anemone (Hufton et al. 2008). Therefore, the amphioxus genome structure cannot be considered a fossil genome representing the pre-duplication condition, at least in terms of genome structure (Garcia-Fernàndez et al. 2001; Hufton et al. 2008), although it is far less divergent from the vertebrate genome structure than is any known urochordate genome (Dehal et al. 2002; Denoeud et al. 2010; Louis et al. 2012).

There have been several attempts to infer the karyotype and genome structure from common chordate ancestors and to reconstruct the evolutionary history leading to present chromosome structures. The first comparisons of conserved syntenic associations in different vertebrate karyotypes, using an in silico chromosome painting approach, allowed reconstructions of the ancestral vertebrate genome containing 10–13 ancestral proto-chromosomes (Kohn et al. 2006; Nakatani et al. 2007). Recently, the sequencing of the amphioxus genome has allowed researchers to reconstruct the ancestral chordate genome as consisting of 17 conserved syntenic blocks, which might represent the ancient chordate proto-chromosomes (Putnam et al. 2008).

After the 2R-WGD, under the naive assumption of absence of loss or fusions of chromosomes, we would expect 68 (17×4) proto-vertebrate segments, but parsimonious reconstruction of chromosome history revealed that numerous chromosomal fusions and translocations have occurred. These reconstructions predict at least 20 fusions that led to 37–49 chromosomes in the bony vertebrate ancestor, which became 12 chromosomes in the stem teleost ancestor due to many additional fusions, and 33–45 chromosomes in the stem tetrapod ancestor due to at least 4 fusions shared between human and chicken genomes (Putnam et al. 2008; Naruse et al. 2004; Nakatani et al. 2007). An excellent example of chromosomal rearrangement after the 2R-WGD has been recently provided by a phylogenetic analysis of members of the four *Hox* paralogons that resulted in a (B(A(C,D))) topology. These results suggest that two chromosomal rearrangements between protochromosomes 11 and 4, and 7 and 5 occurred after the clusters duplicated but before the diversification of extant vertebrates 450 million years ago (Lynch and Wagner 2009). These chromosomal rearrangements resolve conflicting data regarding the order of linked genes and support the hypothesis that the 2R-WGD occurred by two consecutive events of autotetraploidy, and thereby the ancestral vertebrate might have been "pseudo-octoploid". Interestingly, the asymmetrical (B(A(C,D))) topology of the vertebrate *Hox* cluster (Lynch and Wagner 2009) contrasts with the symmetrical (A,B) (C,D) topology inferred from the cartilaginous elephant shark using the amphioxus *Hox* cluster as the outgroup (Ravi et al. 2009). Further extensive analyses including *HoxA-D* clusters from a broader representation of cartilaginous and bony vertebrates will be required to resolve these conflicting topologies, which could suggest that the *Hox*-cluster rearrangement took place after the cartilaginous/bony vertebrate split and not immediately subsequent to the 2R-WGD.

16.4 Consequences of the 2R-WGD on the Evolution of Vertebrate Gene Fate

16.4.1 Function of Gene Duplicates After 2R-WGD

After polyploidization, a period of transilience may follow in which genes might enjoy extra 'degrees of freedom' to mutate without selective penalty (reviewed in Soltis and Soltis 1999; Otto 2007). Understanding the processes by which genome duplication might influence the fate of duplicated genes is crucial to evaluate how the 2R-WGD might have impacted the evolution of vertebrate innovations. Neofunctionalization and subfunctionalization are the two main processes driving the functional fate of newly generated ohnologs after the 2R-WGD and have been extensively discussed in the literature (Hughes 1994; Force et al. 1999; Lynch and Conery 2000; Durand 2003; Postlethwait et al. 2004; Hoekstra and Coyne 2007; Conant and Wolfe 2008; Semon and Wolfe 2007, 2008; Jimenez-Delgado et al. 2009). A prominent example of neofunctionalization related to the 2R-WGD occurred during the expansion of the vertebrate retinoic acid receptor (RAR) family, which acquired new functions in both their expression domains and in their structural protein activities (Escriva et al. 2006). There are also examples, however, in which neofunctionalization and subfunctionalization are related to both 2R-WGD and local tandem duplications (e.g. the expansion of the vertebrate globin superfamily, which promotes the vertebrate innovation related to oxygen transport and storage (Hoffmann et al. 2011)), or merely related to local tandem duplications, and not the 2R-WGD, such as the expansion of the vertebrate Alcohol Dehydrogenase (*Adh*) family, which promotes the acquisition of new enzymes for the synthesis of retinoic acid (Cañestro et al. 2000, 2002a, 2003b). Therefore, not all vertebrate innovations can be exclusively attributed to the 2R-WGD, and the global weight of the impact of the 2R-WGD on the evolution of vertebrate gene functions remains unknown.

16.4.2 Gene Network Rewiring by Tranposons After 2R-WGD

In many cases, neofunctionalization and subfunctionalization can be due to alterations in *cis*-regulatory elements that might lead to adaptive changes in duplicated genes (Force et al. 1999). Many of the *cis*-regulatory elements appear to be embedded in distinct repeat families, especially in transposable elements (TE) (Thornburg et al. 2006; Polak and Domany 2006; Bourque et al. 2008). Analysis of the distribution of 10,000 TEs in the human genome, for instance, revealed that most TEs are concentrated under strong purifying selection near regulatory and developmental genes (Lowe et al. 2007). Most of the described examples of TE mobilization and rewiring of gene regulatory networks have been associated with relatively recent events of TE mobilizations. For instance, TE-mediated rewiring

for neofunctionalization after gene duplication has been recently described for the sex-determining gene *dmrt*1*bY* in medaka fish, in which a novel regulatory element driving a negative feedback on *dmrt1bY* has been acquired due to the insertion of an *Izanagi* transposon (Herpin et al. 2010), or for the origin of a novel gene regulatory network dedicated to pregnancy in placental mammals, which was due to a transposition of the MER20 TE (Lynch et al. 2011).

A massive expansion of TEs appears, therefore, as a powerful mechanism that could boost a vast redeployment of *cis*-regulatory elements into new gene regulatory networks (Feschotte 2008), promoting large-scale events of neofunctionalization and subfunctionalization (van de Lagemaat et al. 2003; Bennetzen 2005; Bejerano et al. 2006). Polyploidization can trigger the mobilization of transposable elements (Matzke and Matzke 1998; Parisod et al. 2010), because recently duplicated genomes contain many redundant genes and substantial repetitive DNA, which serve as buffer against TE insertional mutagenesis (Matzke et al. 2000). According to this expectation, bursts of TE mobilization have been described after polyploidization in different organisms (Matzke and Matzke 1998; Comai 2000; SanMiguel et al. 1996, 1998).

A question that remains unclear is whether there was or was not a massive TE mobilization after the 2R-WGD that could have favored a significant redeployment of *cis*-regulatory elements into new gene regulatory networks in the stem vertebrate lineage. Recent comparison of the diversity and content of TEs between vertebrates and amphioxus has provided some insights that might help to answer this question (Canestro and Albalat 2012). The dynamics of the TE content within a genome follows a competition model in which the expansion of a particular TE might cause the reduction of other types of TEs, consequently reducing the TE diversity, until a new equilibrium that preserves the functionality of the genome is reached (Abrusán and Krambeck 2006). According to this model, if a massive expansion of TEs occurred after the 2R-WGD, we expect that the diversity of TEs shared among vertebrates should be smaller than in cephalochordates. Consistent with this prediction, a recent comparative study reveals that the shared TE diversity of vertebrates (14 superfamilies in lampreys, 28 in ray-finned fishes, 20 in amphibians, 14 in reptiles, 10 in birds, and 15 in mammals) is lower than the TE diversity in amphioxus (33 superfamilies), which makes plausible the hypothesis that a TE burst could have occurred after the 2R-WGD in the stem vertebrate lineage (Canestro and Albalat 2012). Further comparative genomic analysis between different vertebrates and cephalochordates will be required to test this hypothetical burst of TEs, and especially to evaluate its putative impact on the evolution of gene functions after the 2R-WGD.

16.4.3 Ohnologs Gone Missing After 2R-WGD and Impact on Surviving Ohnologs

While several studies focus on the functional fate of retained gene duplicates, less attention has been paid to how losses of paralogs or ohnologs might impact the evolution of the functions of other genes (reviewed in Cañestro et al. 2007). Loss of one copy of two fully redundant gene duplicates should not usually have significant impact, but loss of one of the paralogs after functional divergence likely has evolutionary consequences. Recent analyses of gene losses by comparative genomics have led to the unexpected finding that significant components of the developmental toolkit might be lost without major changes to the body plan (Cañestro and Postlethwait 2007; Holland 2007), which suggests the presence of compensatory mechanisms or the acquisition of innovations that have preserved unaltered the ancestral condition (Cañestro et al. 2007). Tracing the evolution of gene families throughout ancestral proto-chromosomes using blocs of conserved synteny has become a powerful tool to clarify uncertain phylogenies, to detect ohnologs gone missing (OGM) (Postlethwait 2007; Catchen et al. 2009, 2011), to provide robust assessments of orthology and paralogy between different species, and to discern evolutionary innovations from losses of ancestral features in sister lineages (Canestro et al. 2009).

There are cases in which different ohnologs in different species acquire the same expression pattern, which has been called function shuffling (McClintock et al. 2001) and synfunctionalization (Gitelman 2007), and in some cases the convergence of expression patterns between paralogs can be related to OGM (Postlethwait 2007; Canestro et al. 2009). The evolution of the vertebrate retinaldehyde dehydrogenease *Aldh*1*a* family provides a paradigmatic example of how uncovering the evolution of gene family members through the 2R-WGD has been fundamental to illuminating how gene functions evolve among newly generated paralogs after genome duplications in the face of loss of ohnologs (Canestro et al. 2009). For instance, analysis of conserved synteny revealed that the presence of *Aldh*1*a*1 in tetrapods and its absence in teleost fish was not due to a tetrapod innovation, but to an OGM in the teleost stem lineage, which was accompanied by a re-acquisition of ancestral functions by surviving paralogs (Canestro et al. 2009). Medaka provides a more radical example in which *aldh*1*a*2, the only survivor of the *aldh*1*a* family in this species, recapitulates the expression pattern of all other *aldh*1*a* paralogs that have been lost in medaka. This result is in agreement with a model of functional evolution in which surviving genes re-acquire ancestral gene family roles in the face of loss of ohnologs. Other examples that illustrate the importance of identifying OGMs ohnologs are shown in the *endothelin* and *agouti* systems, in which the exclusive presence of *endothelin* 4 (*edn*4) and the *agouti-signaling protein* 2 genes (*asip*2*a/ b*) in teleost fish was not due to a fish innovation related to the teleost-specific whole-genome duplication, but instead to a loss of ohnologs that originated in the 2R-WGD in the tetrapod lineage (Braasch et al. 2009; Braasch and Postlethwait 2011). To understand acquisition of functions of vertebrate ohnologs that were generated by

the 2R-WGD, both the impact of the retention of neo- or subfunctionalized ohnologs, as well as the impact of OGM, on the functions of other survivor gene family members should be studied.

16.5 Consequences of the 2R-WGD on Vertebrate Gene Number and Functional Evolution

How many of the genes that were part of the original fourfold increase in genes generated by 2R-WGD in the stem vertebrate have actually survived nonfunctionalization? And importantly, how significant have the functional consequences of those retained genes been for promoting the origin and evolution of vertebrate innovations? Estimates on gene retention in other organisms that have experienced a WGD have reported ~13 % retention over ~100 million years (MY) in yeast (Wolfe and Shields 1997), ~72 % in maize over ~11 MY (Ahn and Tanksley 1993; Gaut and Doebley 1997), and ~77 % in *Xenopus* over ~40 MY (Hellsten et al. 2007). In vertebrates, a ~33 % retention of divergent functional genes after the 2R-WGD over ~500 MY was inferred initially based on theoretical models applied to 270 gene families of the human genome (Nadeau and Sankoff 1997). More recent and broader analyses based in the complete catalog of human ohnologs estimated a rate of retention between 20 and 30 % (Putnam et al. 2008; Huminiecki and Heldin 2010; Makino and McLysaght 2010).

But how can we assess the impact of the 2R-WGD on the origin of vertebrate complex features? A naive approach to estimating this impact could be to perform a comparison of the total number of retained paralogs and their distribution among functional categories in vertebrates and non-vertebrate chordates that did not undergo any WGD since their split from our last common chordate ancestor. Comparison of the gene catalog of the three chordate subphyla (i.e. cephalochordates, urochordates, and vertebrates) has allowed us to identify a lower bound of 8,437 gene families with members that descend from a single gene in the last common chordate ancestor (Putnam et al. 2008). Through subsequent genome or local duplication, these families account for 13,610 amphioxus genes, 13,401 human genes, and 7,216 ascidian genes, the latter being a significantly lower number due to the extensive gene losses that have occurred in urochordate lineages (Dehal et al. 2002; Cañestro et al. 2003a; Edvardsen et al. 2005; Denoeud et al. 2010). Although it is frequently true that the multiple ohnologs of a vertebrate gene family are represented by a single gene in amphioxus, the total number of paralogs derived from a single-copy gene in the last common ancestor is surprisingly similar between amphioxus (13,610) and human (13,401) (Putnam et al. 2008). Therefore the mere total numbers of retained genes after the 2R-WGD duplicates might not be the key to explain the gain of complexity during the evolution of the vertebrate lineage in comparison with amphioxus.

In vertebrates, analysis of the functional categories of the gene families that have expanded after the 2R-WGD revealed that cell signalers and transcriptional

regulators of developmental pathways are generally retained as multiple ohnologs (Roux and Robinson-Rechavi 2008; Putnam et al. 2008; Hufton et al. 2008; Huminiecki and Heldin 2010). Genes associated with basic cellular functions (i.e. translation, replication, splicing, and recombination, with the important exception of cell cycle), however, have been less successfully retained after the 2R-WGD (Huminiecki and Heldin 2010) (although see Gout et al. (2009) for different results in other organisms that have also undergone WGD). Analysis of the human genome reveals that dosage-balance constraints act on the retention of ohnologs, resulting in an enrichment of dosage-balanced genes, an observation predicted following WGD (Birchler and Veitia 2007, 2010) and also reported for other vertebrates, plants, and yeast (e.g., Paterson et al. 2006). Interestingly, many of these retained ohnologs in humans are refractory to copy number variation, have rarely experienced subsequent small-scale duplication, and are frequently associated with diseases related to dosage-imbalance such as down syndrome (Makino and McLysaght 2010). Analysis of retained genes that have originated in vertebrates by local duplications revealed a strong underrepresentation of genes related to cell communication, cell cycle, and embryo development (Huminiecki and Heldin 2010).

In amphioxus, although a thorough analysis of amphioxus-specific gene family expansions has not been performed, Table 16.1 shows an extensive list of amphioxus-specific duplicated genes reported in the literature (this list is probably not complete, and may be biased toward the research with which I am most familiar). This list shows numerous retained duplicates from a broad array of functional categories, including metabolic enzymes, members of transduction and signaling cascades, members of the immunity system, as well as pivotal transcription factors of developmental pathways. Awaiting a more exhaustive analysis, including different amphioxus species to infer the ancestral cephalochordate condition, the list in Table 16.1 shows no obvious bias toward any particular functional category, although it is noticeable that duplicated developmental transcription factors do not account for more than two paralogs (with the exception of the eight hairy amphioxus paralogs (Minguillon et al. 2003)).

Remarkably, the main difference between the newly acquired paralogs in amphioxus and ohnologs in vertebrates is the mechanism of duplication. While approximately 25 % of the ancestral chordate gene families have two or more ancient vertebrate ohnologs generated by the 2R-WGD, there is strong evidence that most amphioxus paralogs originated by local tandem duplications rather than large-scale chromosomal duplications (Table 16.1). Therefore, considering the functional bias of retention of genes duplicated by WGD or local duplication, it is reasonable to speculate that the key influence of the 2R-WGD promoting the successful diversification of vertebrate features resides in the fact that whole networks were duplicated, in contrast to local duplications such as those that occurred in amphioxus, an organism that seems to have maintained morphological and genetic stasis during the last 200 million years (Garcia-Fernàndez and Holland 1994; Cañestro et al. 2002a; Somorjai et al. 2008; Canestro and Albalat 2012; Paps et al. 2012). Duplication of whole gene networks is dosage-balanced and increases the

Table 16.1 List of paralogs originated independently in the amphioxus lineage

Functional category	Gene	Number of amphioxus paralogs	References
Metabolic proteins	*Aldh*1*a*	5	Cañestro et al. (2006)
	*Cyp*26	3	Albalat (2009)
	*Hsd*11β1	22	Albalat et al. (2011)
	Rdh cluster	12	
	*Bdh*1	14	
	*Rdh*11/12	22	
	*Hsd*17β8	6	
	ApoD	5	
	*Cyp*2	10	
	*Cyp*11/24/27	8	
	*Hsd*3β	7	
	Fabp	8	
	Nos	3	Andreakis et al. (2011)
Transduction and signaling cascades	*Opsins*	20	Holland et al. (2008)
	*GPR*54	18	
	GnRH	3	
	CRHR	5	
	Somatostatine receptors	12	
	*LGR*7/8	6	
	PTHR	4	
	Secretin receptors	5	
	Estrogen receptors	2	Bridgham et al. (2008)
	NOK	22	D'Aniello et al. (2008)
	ACK	3	
	TIE	7	
	MARTK	8	
	EXTK	47	
	HNK-Ras	2	Bertrand et al. (2009)
	Calmoduline-like	3	Karabinos and Bhattacharya (2000)
Immunity system	*Toll-like receptors*	28	Holland et al. (2008)
	C1q-like	39	
	LRR-containing gene models	406	
Developmental transcription factors	*Hairy*	8	Minguillon et al. (2003)
	Brachyury	2	Holland et al. (1995)
	bHLH	2	Araki et al. (1996)
	Emx	2	Minguillon et al. (2002)
	*Hox*13–14	2	Ferrier et al. (2000)
	Evx	2	Ferrier et al. (2001)
	Mnx	2	Holland et al. (2008)
	Uncx	2	
	*Pou*3	2	
	*Lhx*2/9	2	
	Iro	2	

evolvability to generate novel functions, which in the case of the vertebrate 2R-WGD could have led to an increase in complexity of the signaling and developmental regulatory networks that facilitated the acquisition of innovations.

In addition to the evolutionary role of coding genes in the acquisition of innovations, microRNAs (miRNAs) also play crucial roles during development and have been postulated as important players for the evolution of organismal complexity (Lee et al. 2007; Sempere et al. 2006). Analysis of miRNAs in chordate species showed that the 2R-WGD has increased the diversity of the inventory of miRNAs in vertebrates, which correlated with the increase of complex patterns of tissue specificity of miRNAs (Heimberg et al. 2008; Campo-Paysaa et al. 2011). However, the finding of 41 vertebrate-specific miRNA families, absent in non-vertebrate chordates, suggests that their origin must have occurred in stem vertebrates after their separation from urochordates and is not explained by the 2R-WGD (Heimberg et al. 2008). The appearance of these 41 vertebrate-specific miRNA families has been proposed as a potential key evolutionary force lying behind the dramatic increase of vertebrate complexity (Heimberg et al. 2008). Future exhaustive analysis of the expression patterns of the members of these 41 vertebrate-specific families, and an understanding of their roles, will allow a reevaluation of the importance that this innovation could have had on the origin of vertebrate features.

16.6 Consequences of the 2R-WGD on the Innovation of Vertebrate Features

A major question not resolved yet is the precise impact of the 2R-WGD on the innovation of particular vertebrate features. Three vertebrate features are perhaps the most prominent innovations: derivatives from neural crest cells, sensory organs concentrated in the head derived from ectodermal placodes, and a big complex brain. When taken together, these features probably allowed the transition from ancestral, peaceful, filter-feeding, non-vertebrate chordates to active, voracious, vertebrate predators (Northcutt and Gans 1983; Gans and Northcutt 1983); reviewed in Yu et al. 2008; Holland 2009) (Fig. 16.1).

Vertebrate neural crest cells are a transient population of developmental cells that delaminate at the border of the neural plate through an epithelial–mesenchymal transition, migrate, and differentiate at their final destination into a variety of structures such as sensory neurons, glial cells, peripheral nervous system, pigment cells, smooth muscle cells, connective tissue, cranio-facial cartilage, skeletal bones, and teeth (Weston 1970). Vertebrate crest development depends on four crucial sets of genes that form what is called the neural crest gene regulatory network (NC-GRN): (1) patterning signal genes establish the expression of (2) neural plate border specifier genes, which activate (3) crest specifier genes, which turn on (4) neural crest effector genes that provide differentiated products (Meulemans and Bronner-Fraser 2004, 2005; Ota and Kuratani 2007; Sauka-Spengler et al. 2007). Analysis of the amphioxus genome has revealed the presence of cephalochordate

orthologs from all of these four sets of genes, including (1): *Fgf*, *Wnt*, *Bmp*, *Notch Dlx*, *AP*2, *SoxB, Zic*, and *islet*; (2): *Pax*3/7, *Msx*, *Dlx*5, and *Zic*; (3): *Snail*, *SoxE*, *AP*2, *Twist, Id, FoxD*, and *Myc*; and (4): *Rho, cRet, Erbb*3, *Mitf, tyrosinase*, and *tyrosinase-related* genes, with the remarkable exception of the tyrosine kinase *c-Kit* essential for migration and survival of crest cells, and the gene for myelin protein P0, consistent with the notion that the glial myelin sheath is a vertebrate innovation (Meulemans and Bronner-Fraser 2007; Holland et al. 2008; Holland and Short 2008; Nikitina et al. 2009).

The fact that most of the specifier genes are present as single copy in amphioxus, but multiple paralogs in vertebrates, presumably due to the 2R-WGD, has led to the hypothesis that neofunctionalization and subfunctionalization of paralogs may have facilitated the co-option of ancestral genes into the NC-GRN (Sauka-Spengler et al. 2007; Meulemans and Bronner-Fraser 2007; Holland et al. 2008; Holland and Short 2008). Gene ontology (GO) analysis estimates that 91 % of the neural crest genes in vertebrates have been co-opted from genes already present in basal metazoans, while the remaining 9 % of the neural crest genes are vertebrate innovations (Martinez-Morales et al. 2007), including the assembly of new signaling pathways like the endothelin system (Braasch et al. 2009). The evolution of the vertebrate NC-GRN, therefore, appears as the result of a combination of ancestral gene co-option, newly evolving genes, and amplification of these components by the 2R-WGDs (Braasch et al. 2009).

Interestingly, however, the recent description in urochordates of neural crest-like cells that express typical vertebrate crest marker orthologs, migrate, and differentiate into pigments challenges the idea that neural crest cells are a vertebrate innovation (Jeffery et al. 2004, 2008; Jeffery 2007). Thus, it cannot be discounted that some types of neural crest cells might have been present in the last common ancestor of olfactores (urochordates + vertebrates), followed by losses during the significant morphological and genetic simplification suffered by urochordate lineages (Seo et al. 2004; Cañestro et al. 2005; Cañestro and Postlethwait 2007; Holland 2007). Therefore, it seems plausible to consider that the 2R-WGD might have not been crucial for the origin of the neural crest cells, but the 2R-WGD might have been important for increasing the evolvability of the NC-GRN and the diversification of derivative structures.

Similar conclusions have been reached through studies of the gene regulatory network underlying placode and brain development. Analysis of placode-marking genes (e.g. *Eye*, *Pitx*, *Six*, and *Pax*) in ascidian and larvacean urochordates suggested that the last common olfactore ancestor already presented multiple placode derivatives, such as olfactory and adenohypophyseal. Additional and independent proliferation and loss of a variety of placodes probably occurred in both urochordate and vertebrate lineages (Bassham and Postlethwait 2005; Mazet et al. 2005), in some cases recruiting paralogs that had been independently duplicated in both urochordates and vertebrates (Bassham et al. 2008).

Despite the fact that non-vertebrate chordates have a simple brain lacking a midbrain and a midbrain–hindbrain organizer (MHB), most brain-making gene orthologs are present in non-vertebrate chordates, suggesting that vertebrate brain

features were built on a foundation already present in the ancestral chordate probably facilitated by the new ohnologs created by the vertebrate 2R-WGD (reviewed in Holland 2009). Recent analysis of developmental genes in the ascidian brain revealed that the expression of *Fgf*8 can reorganize the expression of other brain genes and transform hindbrain structures into an expanded mesencephelon, recapitulating the organizing activity of the vertebrate MHB and therefore suggesting that the MHB was already present at least in the last common ancestor of olfactores (Imai et al. 2009). Analysis of urochordate genomes revealed that important genes (i.e. *Gbx*) for the positioning of the MHB have been lost in stem urochordates (Cañestro et al. 2005), as has the retinoic acid dependent anterior–posterior axial patterning of the central nervous system (Cañestro et al. 2006), making plausible the hypothesis that the absence of midbrain in urochordates is not due to a vertebrate innovation of a midbrain, but a simplification in urochordates of an ancestral tripartite brain structure (Cañestro and Postlethwait 2007; Cañestro et al. 2007). Evolutionary analysis of the origin of the complex Nova-regulated splice variants of the vertebrate brain genes revealed that many of these variants were already present in the last common olfactore ancestor (Irimia et al. 2011). It is possible that the 2R-WGD promoted the increase of the complexity of Nova-dependent splice variants in the vertebrate brain, although a simplification of this system during urochordate evolution cannot be discarded.

In conclusion, it is likely that the origin of vertebrate features such as neural crest cells, placodes, and a complex tripartite brain are not related to the 2R-WGD, but that these features were already present to some extent in stem non-vertebrate chordates (Fig. 16.1) (reviewed in Donoghue et al. 2008). However, it is likely that the subsequent evolution of these three features has been strongly influenced by the new ohnologs that originated after the 2R-WGD, due to processes of neofunctionalization, subfunction partitioning and subsequent refinement, recruitment of *cis*-regulatory elements driven by genome rearrangement and transposable element activity, inventions of novel miRNA families, and evolution of novel splice variants, which overall increased the complexity of duplicated developmental gene regulatory networks after the 2R-WGD. Future integrative analysis of comparative genomics, functional evo-devo, and examinations of gene regulatory networks in a wide variety of non-vertebrate chordates as well as basally divergent jawless vertebrates will help to narrow down the precise timing of the 2R-WGD and evaluate its actual impact on the origin and evolution of vertebrate features.

Probably the new '2R, or not 2R' question (Hughes and Friedman 2003) is now to ascertain whether the origins of vertebrate innovations were, or were not, the consequence of the 2R-WGD, and to understand the mechanisms by which the 2R-WGD increased the evolvability of developmental gene regulatory networks that facilitated the diversification of complex vertebrate features.

References

Abrusán G, Krambeck H-J (2006) Competition may determine the diversity of transposable elements. Theor Popul Biol 70(3):364–375

Aburomia R, Khaner O, Sidow A (2003) Functional evolution in the ancestral lineage of vertebrates or when genomic complexity was wagging its morphological tail. J Struct Funct Genomics 3(1–4):45–52

Ahn S, Tanksley SD (1993) Comparative linkage maps of the rice and maize genomes. Proc Natl Acad Sci U S A 90(17):7980–7984

Albalat R (2009) The retinoic acid machinery in invertebrates: ancestral elements and vertebrate innovations. Mol Cell Endocrinol 313(1–2):23–35

Albalat R, Brunet F, Laudet V, Schubert M (2011) Evolution of retinoid and steroid signaling: vertebrate diversification from an amphioxus perspective. Genome Biol Evol 3:985–1005

Andreakis N, D'Aniello S, Albalat R, Patti FP, Garcia-Fernandez J, Procaccini G, Sordino P, Palumbo A (2011) Evolution of the nitric oxide synthase family in metazoans. Mol Biol Evol 28(1):163–179

Araki I, Terazawa K, Satoh N (1996) Duplication of an amphioxus myogenic bHLH gene is independent of vertebrate myogenic bHLH gene duplication. Gene 171(2):231–236

Bailey W, Kim J, Wagner G, Ruddle F (1997) Phylogenetic reconstruction of vertebrate *Hox* cluster duplications. Mol Biol Evol 14:843–853

Bassham S, Cañestro C, Postlethwait JH (2008) Evolution of developmental roles of Pax2/5/8 paralogs after independent duplication in urochordate and vertebrate lineages. BMC Biol 6:35

Bassham S, Postlethwait JH (2005) The evolutionary history of placodes: a molecular genetic investigation of the larvacean urochordate *Oikopleura dioica*. Development 132(19):4259–4272

Bejerano G, Lowe CB, Ahituv N, King B, Siepel A, Salama SR, Rubin EM, Kent WJ, Haussler D (2006) A distal enhancer and an ultra conserved exon are derived from a novel retroposon. Nature 441(7089):87–90

Bennetzen JL (2005) Transposable elements, gene creation and genome rearrangement in flowering plants. Curr Opin Genet Dev 15(6):621–627

Bertrand S, Campo-Paysaa F, Camasses A, Garcia-Fernandez J, Escriva H (2009) Actors of the tyrosine kinase receptor downstream signaling pathways in amphioxus. Evol Dev 11(1):13–26

Birchler JA, Veitia RA (2007) The gene balance hypothesis: from classical genetics to modern genomics. Plant Cell 19(2):395–402

Birchler JA, Veitia RA (2010) The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution. New Phytol 186(1):54–62

Bourlat SJ, Juliusdottir T, Lowe CJ, Freeman R, Aronowicz J, Kirschner M, Lander ES, Thorndyke M, Nakano H, Kohn AB, Heyland A, Moroz LL, Copley RR, Telford MJ (2006) Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. Nature 444(7115):85–88

Bourque G, Leong B, Vega VB, Chen X, Lee YL, Srinivasan KG, Chew JL, Ruan Y, Wei CL, Ng HH, Liu ET (2008) Evolution of the mammalian transcription factor binding repertoire via transposable elements. Genome Res 18(11):1752–1762

Braasch I, Postlethwait JH (2011) The teleost agouti-related protein 2 gene is an ohnolog gone missing from the tetrapod genome. Proc Natl Acad Sci U S A 108(13):E47–E48

Braasch I, Volff JN, Schartl M (2009) The endothelin system: evolution of vertebrate-specific ligand–receptor interactions by three rounds of genome duplication. Mol Biol Evol 26(4):783–799

Bridgham JT, Brown JE, Rodriguez-Mari A, Catchen JM, Thornton JW (2008) Evolution of a new function by degenerative mutation in cephalochordate steroid receptors. PLoS Genet 4(9):e1000191

Brusca RC, Brusca GJ (2002) Invertebrates. Sinauer Associates, Sunderland

Campo-Paysaa F, Semon M, Cameron RA, Peterson KJ, Schubert M (2011) microRNA complements in deuterostomes: origin and evolution of microRNAs. Evol Dev 13(1):15–27

Cañestro C, Albalat R (2012) Transposon diversity is higher in amphioxus than in vertebrates: functional and evolutionary inferences. Brief Funct Genomics 11(2):131–141

Cañestro C, Albalat R, Hjelmqvist L, Godoy L, Jornvall H, Gonzalez-Duarte R (2002a) Ascidian and amphioxus Adh genes correlate functional and molecular features of the ADH family expansion during vertebrate evolution. J Mol Evol 54(1):81–89

Cañestro C, Bassham S, Postlethwait JH (2003a) Seeing chordate evolution through the *Ciona* genome sequence. Genome Biol 4(3):208–211

Cañestro C, Bassham S, Postlethwait JH (2005) Development of the central nervous system in the larvacean *Oikopleura dioica* and the evolution of the chordate brain. Dev Biol 285(2):298–315

Cañestro C, Catchen JM, Rodriguez-Mari A, Yokoi H, Postlethwait JH (2009) Consequences of lineage-specific gene loss on functional evolution of surviving paralogs: ALDH1A and retinoic acid signaling in vertebrate genomes. PLoS Genet 5(5):e1000496

Cañestro C, Godoy L, Gonzàlez-Duarte R, Albalat R (2003b) Comparative expression analysis of Adh3 during arthropod, urochordate, cephalochordate and vertebrate development challenges its predicted housekeeping role. Evol Dev 5(2):157–162

Cañestro C, Gonzàlez-Duarte R, Albalat R (2002b) Minisatellite instability at the Adh locus reveals somatic polymorphism in amphioxus. Nucleic Acids Res 30(13):2871–2876

Cañestro C, Hjelmqvist L, Albalat R, Garcia-Fernàndez J, Gonzàlez-Duarte R, Jörnvall H (2000) Amphioxus alcohol dehydrogenase is a class 3 form of single type and of structural conservation but with unique developmental expression. Eur J Biochem 267:6511–6518

Cañestro C, Postlethwait JH (2007) Development of a chordate anterior–posterior axis without classical retinoic acid signaling. Dev Biol 305(2):522–538

Cañestro C, Postlethwait JH, Gonzàlez-Duarte R, Albalat R (2006) Is retinoic acid genetic machinery a chordate innovation? Evol Dev 8(5):394–406

Cañestro C, Yokoi H, Postlethwait JH (2007) Evolutionary developmental biology and genomics. Nat Rev Genet 8(12):932–942

Catchen JM, Braasch I, Postlethwait JH (2011) Conserved synteny and the zebrafish genome. Methods Cell Biol 104:259–285

Catchen JM, Conery JS, Postlethwait JH (2009) Automated identification of conserved synteny after whole-genome duplication. Genome Res 19(8):1497–1505

Chain FJ, Evans BJ (2006) Multiple mechanisms promote the retained expression of gene duplicates in the tetraploid frog *Xenopus laevis*. PLoS Genet 2(4):e56

Comai L (2000) Genetic and epigenetic interactions in allopolyploid plants. Plant Mol Biol 43 (2–3):387–399

Conant GC, Wolfe KH (2008) Turning a hobby into a job: how duplicated genes find new functions. Nat Rev Genet 9(12):938–950

Coulier F, Popovici C, Villet R, Birnbaum D (2000) MetaHox gene clusters. J Exp Zool 288(4):345–351

D'Aniello S, Irimia M, Maeso I, Pascual-Anaya J, Jimenez-Delgado S, Bertrand S, Garcia-Fernandez J (2008) Gene expansion and retention leads to a diverse tyrosine kinase superfamily in amphioxus. Mol Biol Evol 25(9):1841–1854

Danchin EG, Pontarotti P (2004) Towards the reconstruction of the bilaterian ancestral pre-MHC region. Trends Genet 20(12):587–591

Dehal P, Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol 3(10):e314

Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomoso A, Davidson B, Di Gregorio A, Gelpke M, Goodstein DM, Harafuji N, Hastings KEM, Ho I, Hotta K, Huang W, Kawashima T, Lemaire P, Martinez D, Meinertzhagen IA, Necula S, Nonaka M, Putnam N, Rash S, Saiga H, Satake M, Terry A, Yamada L, Wang H-G, Awazu S, Azumi K, Boore J, Branno M, Chin-bow S, DeSantis R, Doyle S, Francino P, Keys DN, Haga S, Hayashi H, Hino K, Imai KS, Inaba K, Kano S, Kobayashi K, Kobayashi M, Lee B-I, Makabe KW, Manohar C, Matassi G, Medina M, Mochizuki Y, Mount S, Morishita T, Miura S, Nakayama A, Nishizaka S, Nomoto H, Ohta F, Oishi K, Rigoutsos I, Sano M, Sasaki A, Sasakura Y, Shoguchi E, Shin-i T, Spagnuolo A, Stainier D, Suzuki MM, Tassy O, Takatori N, Tokuoka M, Yagi K, Yoshizaki F, Wada S, Zhang C, Hyatt PD, Larimer F, Detter C, Doggett N, Glavina T, Hawkins T, Richardson P, Lucas S, Kohara Y, Levine M, Satoh N, Rokhsar DS (2002) The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. Science 298(5601):2157–2167

Delsuc F, Brinkmann H, Chourrout D, Philippe H (2006) Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature 439(7079):965–968

Denoeud F, Henriet S, Mungpakdee S, Aury JM, Da Silva C, Brinkmann H, Mikhaleva J, Olsen LC, Jubin C, Cañestro C, Bouquet JM, Danks G, Poulain J, Campsteijn C, Adamski M, Cross I, Yadetie F, Muffato M, Louis A, Butcher S, Tsagkogeorga G, Konrad A, Singh S, Jensen MF, Cong EH, Eikeseth-Otteraa H, Noel B, Anthouard V, Porcel BM, Kachouri-Lafond R, Nishino A, Ugolini M, Chourrout P, Nishida H, Aasland R, Huzurbazar S, Westhof E, Delsuc F, Lehrach H, Reinhardt R, Weissenbach J, Roy SW, Artiguenave F, Postlethwait JH, Manak JR, Thompson EM, Jaillon O, Du Pasquier L, Boudinot P, Liberles DA, Volff JN, Philippe H, Lenhard B, Roest Crollius H, Wincker P, Chourrout D (2010) Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. Science 330(6009):1381–1385

Donoghue PC, Graham A, Kelsh RN (2008) The origin and evolution of the neural crest. BioEssays 30(6):530–541

Durand D (2003) Vertebrate evolution: doubling and shuffling with a full deck. Trends Genet 19(1):2–5

Ebner B, Panopoulou G, Vinogradov SN, Kiger L, Marden MC, Burmester T, Hankeln T (2010) The globin gene family of the cephalochordate amphioxus: implications for chordate globin evolution. BMC Evol Biol 10:370

Edvardsen RB, Seo HC, Jensen MF, Mialon A, Mikhaleva J, Bjordal M, Cartry J, Reinhardt R, Weissenbach J, Wincker P, Chourrout D (2005) Remodelling of the homeobox gene complement in the tunicate *Oikopleura dioica*. Curr Biol 15(1):R12–R13

Escriva H, Bertrand S, Germain P, Robinson-Rechavi M, Umbhauer M, Cartry J, Duffraisse M, Holland L, Gronemeyer H, Laudet V (2006) Neofunctionalization in vertebrates: the example of retinoic acid receptors. PLoS Genet 2(7):e102

Ferrier DE, Dewar K, Cook A, Chang JL, Hill-Force A, Amemiya C (2005) The chordate *ParaHox* cluster. Curr Biol 15(20):R820–R822

Ferrier DE, Minguillon C, Holland PW, Garcia-Fernandez J (2000) The amphioxus *Hox* cluster: deuterostome posterior flexibility and *Hox*14. Evol Dev 2(5):284–293

Ferrier DEK, Minguillón C, Cebrián C, Garcia-Fernàndez J (2001) Amphioxus *Evx* genes: implications for the evolution of the midbrain–hindbrain boundary and the chordate tailbud. Dev Biol 237(2):270–281

Ferris SD, Portnoy SL, Whitt GS (1979) The roles of speciation and divergence time in the loss of duplicate gene expression. Theor Popul Biol 15(1):114–139

Feschotte C (2008) Transposable elements and the evolution of regulatory networks. Nat Rev Genet 9(5):397–405

Force A, Amores A, Postlethwait JH (2002) *Hox* cluster organization in the jawless vertebrate *Petromyzon marinus*. J Exp Zool 294(1):30–46

Force A, Lynch M, Pickett FB, Amores A, Yan Y-L, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerate mutations. Genetics 151(4):1531–1545

Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol 63(6):1256–1272

Fried C, Prohaska SJ, Stadler PF (2003) Independent *Hox*-cluster duplications in lampreys. J Exp Zoolog Part B Mol Dev Evol 299(1):18–25

Friedman R, Hughes AL (2001) Pattern and timing of gene duplication in animal genomes. Genome Res 11(11):1842–1847

Furlong R, Holland P (2002) Were vertebrates octoploid? Philos T Roy Soc B 357:531–544

Furlong RF, Younger R, Kasahara M, Reinhardt R, Thorndyke M, Holland PW (2007) A degenerate *ParaHox* gene cluster in a degenerate vertebrate. Mol Biol Evol 24(12): 2681–2686

Gans C, Northcutt RG (1983) Neural crest and the origin of vertebrates: a new head. Science 220:268–274

Garcia-Fernandez J (2005) The genesis and evolution of homeobox gene clusters. Nat Rev Genet 6(12):881–892

Garcia-Fernàndez J, Ferrier DE, Minguillón C, Cebrián C (2001) The amphioxus genome in evo-devo: archetype or "cul de sac". Int J Dev Biol 45(S1):S137–S138

Garcia-Fernàndez J, Holland PW (1994) Archetypal organization of the amphioxus *Hox* gene cluster. Nature 370(6490):563–566

Gaut BS, Doebley JF (1997) DNA sequence evidence for the segmental allotetraploid origin of maize. Proc Natl Acad Sci U S A 94(13):6809–6814

Gitelman I (2007) Evolution of the vertebrate twist family and synfunctionalization: a mechanism for differential gene loss through merging of expression domains. Mol Biol Evol 24(9): 1912–1925

Gout JF, Duret L, Kahn D (2009) Differential retention of metabolic genes following whole-genome duplication. Mol Biol Evol 26(5):1067–1072

Gu X, Wang Y, J G (2002) Age distribution of human gene families shows significant roles of both large- and small-scale duplications in vertebrate evolution. Nat Genet 31:205–209

Heimberg AM, Sempere LF, Moy VN, Donoghue PC, Peterson KJ (2008) microRNAs and the advent of vertebrate morphological complexity. Proc Natl Acad Sci U S A 105(8):2946–2950

Hellsten U, Khokha MK, Grammer TC, Harland RM, Richardson P, Rokhsar DS (2007) Accelerated gene evolution and subfunctionalization in the pseudotetraploid frog *Xenopus laevis*. BMC Biol 5:31

Herpin A, Braasch I, Kraeussling M, Schmidt C, Thoma EC, Nakamura S, Tanaka M, Schartl M (2010) Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. PLoS Genet 6(2):e1000844

Hoekstra HE, Coyne JA (2007) The locus of evolution: evo devo and the genetics of adaptation. Evolution 61(5):995–1016

Hoffmann FG, Opazo JC, Storz JF (2012) Whole-genome duplications spurred the functional diversification of the globin gene superfamily in vertebrates. Mol Biol Evol 29:303−312

Holland LZ (2007) Developmental biology: a chordate with a difference. Nature 447(7141):153–155

Holland LZ (2009) Chordate roots of the vertebrate nervous system: expanding the molecular toolkit. Nat Rev Neurosci 10(10):736–746

Holland LZ, Albalat R, Azumi K, Benito-Gutierrez E, Blow MJ, Bronner-Fraser M, Brunet F, Butts T, Candiani S, Dishaw LJ, Ferrier DE, Garcia-Fernandez J, Gibson-Brown JJ, Gissi C, Godzik A, Hallbook F, Hirose D, Hosomichi K, Ikuta T, Inoko H, Kasahara M, Kasamatsu J, Kawashima T, Kimura A, Kobayashi M, Kozmik Z, Kubokawa K, Laudet V, Litman GW, McHardy AC, Meulemans D, Nonaka M, Olinski RP, Pancer Z, Pennacchio LA, Pestarino M, Rast JP, Rigoutsos I, Robinson-Rechavi M, Roch G, Saiga H, Sasakura Y, Satake M, Satou Y, Schubert M, Sherwood N, Shiina T, Takatori N, Tello J, Vopalensky P, Wada S, Xu A, Ye Y, Yoshida K, Yoshizaki F, Yu JK, Zhang Q, Zmasek CM, de Jong PJ, Osoegawa K, Putnam

NH, Rokhsar DS, Satoh N, Holland PW (2008) The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Res 18(7):1100–1111

Holland LZ, Short S (2008) Gene duplication, co-option and recruitment during the origin of the vertebrate brain from the invertebrate chordate brain. Brain Behav Evol 72(2):91–105

Holland PWH, Garcia-Fernàndez J, Williams NA, Sidow A (1994) Gene duplications and the origins of vertebrate development. Development (Suppl.):125–133

Holland PWH, Koschorz B, Holland LZ, Herrmann BG (1995) Conservation of *Brachyury* (*T*) genes in amphioxus and vertebrates: developmental and evolutionary implications. Development 121:4283–4291

Hufton AL, Groth D, Vingron M, Lehrach H, Poustka AJ, Panopoulou G (2008) Early vertebrate whole genome duplications were predated by a period of intense genome rearrangement. Genome Res 18(10):1582–1591

Hughes AL (1994) The evolution of functionally novel proteins after gene duplication. Proc Roy Soc Lond B 256:119–124

Hughes AL (1999) Phylogenies of developmentally important proteins do not support the hypothesis of two rounds of genome duplication early in vertebrate history. J Mol Evol 48(5):565–576

Hughes AL, Friedman R (2003) 2R or not 2R: testing hypotheses of genome duplication in early vertebrates. J Struct Funct Genomics 3(1–4):85–93

Huminiecki L, Heldin CH (2010) 2R and remodeling of vertebrate signal transduction engine. BMC Biol 8:146

Imai KS, Stolfi A, Levine M, Satou Y (2009) Gene regulatory networks underlying the compartmentalization of the Ciona central nervous system. Development 136(2):285–293

Irimia M, Denuc A, Burguera D, Somorjai I, Martin-Duran JM, Genikhovich G, Jimenez-Delgado S, Technau U, Roy SW, Marfany G, Garcia-Fernandez J (2011) Stepwise assembly of the Nova-regulated alternative splicing network in the vertebrate brain. Proc Natl Acad Sci U S A 108(13):5319–5324

Irvine SQ, Carr JL, Bailey WJ, Kawasaki K, Shimizu N, Amemiya CT, Ruddle FH (2002) Genomic analysis of *Hox* clusters in the sea lamprey *Petromyzon marinus*. J Exp Zool 294:47–62

Jeffery WR (2007) Chordate ancestry of the neural crest: new insights from ascidians. Semin Cell Dev Biol 18:481−491

Jeffery WR, Chiba T, Krajka FR, Deyts C, Satoh N, Joly JS (2008) Trunk lateral cells are neural crest-like cells in the ascidian *Ciona intestinalis*: insights into the ancestry and evolution of the neural crest. Dev Biol 324(1):152–160

Jeffery WR, Strickler AG, Yamamoto Y (2004) Migratory neural crest-like cells form body pigmentation in a urochordate embryo. Nature 431(7009):696–699

Jimenez-Delgado S, Pascual-Anaya J, Garcia-Fernandez J (2009) Implications of duplicated cis-regulatory elements in the evolution of metazoans: the DDI model or how simplicity begets novelty. Brief Funct Genomic Proteomic 8(4):266–275

Kappen C, Schughart K, Ruddle FH (1989) Two steps in the evolution of Antennapedia-class vertebrate homeobox genes. Proc Natl Acad Sci U S A 86(14):5459–5463

Karabinos A, Bhattacharya D (2000) Molecular evolution of calmodulin and calmodulin-like genes in the cephalochordate *Branchiostoma*. J Mol Evol 51(2):141–148

Kasahara M (2007) The 2R hypothesis: an update. Curr Opin Immunol 19(5):547–552

Kassahn KS, Dang VT, Wilkins SJ, Perkins AC, Ragan MA (2009) Evolution of gene function and regulatory control after whole-genome duplication: comparative analyses in vertebrates. Genome Res 19(8):1404–1418

Katsanis N, Fitzgibbon J, Fisher EMC (1996) Paralogy mapping: identification of a region in the human MHC triplicated onto human chromosomes 1 and 9 allows the prediction and isolation of novel *PBX* and *NOTCH* loci. Genomics 35:101–108

Kohn M, Hogel J, Vogel W, Minich P, Kehrer-Sawatzki H, Graves JA, Hameister H (2006) Reconstruction of a 450-My-old ancestral vertebrate protokaryotype. Trends Genet 22(4):203–210

Kuraku S (2008) Insights into cyclostome phylogenomics: pre-2R or post-2R. Zoolog Sci 25(10):960–968

Kuraku S (2010) Palaeophylogenomics of the vertebrate ancestor—impact of hidden paralogy on hagfish and lamprey gene phylogeny. Integr Comp Biol 50(1):124–129

Kuraku S (2011) *Hox* gene clusters of early vertebrates: do they serve as reliable markers for genome evolution? Genomics Proteomics Bioinformatics 9(3):97–103

Kuraku S, Meyer A, Kuratani S (2009) Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? Mol Biol Evol 26(1):47–59

Kuratani S (2009) Insights into neural crest migration and differentiation from experimental embryology. Development 136(10):1585–1589

Kuratani S, Ota KG (2008) Hagfish (cyclostomata, vertebrata): searching for the ancestral developmental plan of vertebrates. BioEssays 30(2):167–172

Laisney JA, Braasch I, Walter RB, Meierjohann S, Schartl M (2010) Lineage-specific co-evolution of the Egf receptor/ligand signaling system. BMC Evol Biol 10:27

Larhammar D, Lundin L, Hallbook F (2002) The human *Hox*-bearing chromosome regions did arise by block or chromosome (or even genome) duplications. Genome Res 12(12):1910–1920

Lee CT, Risom T, Strauss WM (2007) Evolutionary conservation of microRNA regulatory circuits: an examination of microRNA gene complexity and conserved microRNA-target interactions through metazoan phylogeny. DNA Cell Biol 26(4):209–218

Leveugle M, Prat K, Perrier N, Birnbaum D, Coulier F (2003) ParaDB: a tool for paralogy mapping in vertebrate genomes. Nucleic Acids Res 31(1):63–67

Louis A, Roest Crollius H, Robinson-Rechavi M (2012) How much does the amphioxus genome represent the ancestor of chordates? Brief Funct Genomics 11(2):89–95

Lowe CB, Bejerano G, Haussler D (2007) Thousands of human mobile element fragments undergo strong purifying selection near developmental genes. Proc Natl Acad Sci U S A 104(19):8005–8010

Lundin LG (1979) Evolutionary conservation of large chromosomal segments reflected in mammalian gene maps. Clin Genet 16(2):72–81

Lundin LG (1993) Evolution of the vertebrate genome as relected in paralogous chromosomal regions in man and the house mouse. Genomics 16:1–19

Lundin LG, Larhammar D, Hallbook F (2003) Numerous groups of chromosomal regional paralogies strongly indicate two genome doublings at the root of the vertebrates. J Struct Funct Genomics 3(1–4):53–63

Lynch M, Conery J (2000) The evolutionary fate and consequences of gene duplication. Science 290(5494):1151–1155

Lynch M, Force A (2000) The origin of interspecific genomic incompatibility via gene duplication. Am Nat 156:590–605

Lynch M, O'Hely M, Walsh B, Force A (2001) The probability of preservation of a newly arisen gene duplicate. Genetics 159(4):1789–1804

Lynch VJ, Leclerc RD, May G, Wagner GP (2011) Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. Nat Genet 43:1154−1159

Lynch VJ, Wagner GP (2009) Multiple chromosomal rearrangements structured the ancestral vertebrate *Hox*-bearing protochromosomes. PLoS Genet 5(1):e1000349

Makino T, McLysaght A (2010) Ohnologs in the human genome are dosage balanced and frequently associated with disease. Proc Natl Acad Sci U S A 107(20):9270–9274

Martinez-Morales JR, Henrich T, Ramialison M, Wittbrodt J (2007) New genes in the evolution of the neural crest differentiation program. Genome Biol 8(3):R36

Matzke MA, Matzke AJ (1998) Polyploidy and transposons. Trends Ecol Evol 13(6):241

Matzke MA, Mette MF, Matzke AJ (2000) Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant Mol Biol 43(2–3):401–415

Mazet F, Hutt JA, Milloz J, Millard J, Graham A, Shimeld SM (2005) Molecular evidence from *Ciona intestinalis* for the evolutionary origin of vertebrate sensory placodes. Dev Biol 282(2):494–508

McClintock JM, Carlson R, Mann DM, Prince VE (2001) Consequences of *Hox* gene duplication in the vertebrates: an investigation of the zebrafish *Hox* paralogue group 1 genes. Development 128(13):2471–2484

McLysaght A, Hokamp K, Wolfe KH (2002) Extensive genomic duplication during early chordate evolution. Nat Genet 31:200–204

Meulemans D, Bronner-Fraser M (2004) Gene-regulatory interactions in neural crest evolution and development. Dev Cell 7(3):291–299

Meulemans D, Bronner-Fraser M (2005) Central role of gene cooption in neural crest evolution. J Exp Zoolog B Mol Dev Evol 304(4):298–303

Meulemans D, Bronner-Fraser M (2007) Insights from amphioxus into the evolution of vertebrate cartilage. PLoS One 2(8):e787

Minguillon C, Ferrier DE, Cebrian C, Garcia-Fernandez J (2002) Gene duplications in the prototypical cephalochordate amphioxus. Gene 287(1–2):121–128

Minguillon C, Jimenez-Delgado S, Panopoulou G, Garcia-Fernandez J (2003) The amphioxus Hairy family: differential fate after duplication. Development 130(24):5903–5914

Muffato M, Louis A, Poisnel CE, Roest Crollius H (2010) Genomicus: a database and a browser to study gene synteny in modern and ancestral genomes. Bioinformatics 26(8):1119–1121

Nadeau JH, Sankoff D (1997) Comparable rates of gene loss and functional divergence after genome duplications early in vertebrate evolution. Genetics 147(3):1259–1266

Nakatani Y, Takeda H, Kohara Y, Morishita S (2007) Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. Genome Res 17(9):1254–1265

Naruse K, Tanaka M, Mita K, Shima A, Postlethwait J, Mitani H (2004) A medaka gene map: the trace of ancestral vertebrate proto-chromosomes revealed by comparative gene mapping. Genome Res 14(5):820–828

Nelson JS (1994) Fishes of the world, 3rd edn. Wiley-Interscience, New York

Nikitina N, Sauka-Spengler T, Bronner-Fraser M (2009) Chapter 1. Gene regulatory networks in neural crest development and evolution. Curr Top Dev Biol 86:1–14

Northcutt RG, Gans C (1983) The genesis of neural crest and epidermal placodes: a reinterpretation of vertebrate origins. Quart Rev Biol 58:1–28

Nowak MA, Boerlijst MC, Cooke J, Smith JM (1997) Evolution of genetic redundancy. Nature 388(6638):167–171

Oda H, Wada H, Tagawa K, Akiyama-Oda Y, Satoh N, Humphreys T, Zhang S, Tsukita S (2002) A novel amphioxus cadherin that localizes to epithelial adherens junctions has an unusual domain organization with implications for chordate phylogeny. Evol Dev 4(6):426–434

Ohno S (1970) Evolution by gene duplication. Springer, New York

Ohno S, Wolf U, Atkins NB (1968) Evolution from fish to mammals by gene duplication. Hereditas 59(1):169–187

Ota KG, Kuratani S (2007) Cyclostome embryology and early evolutionary history of vertebrates. Integr Comp Biol 47(3):329–337

Otto SP (2007) The evolutionary consequences of polyploidy. Cell 131(3):452–462

Panopoulou G, Hennig S, Groth D, Krause A, Poustka AJ, Herwig R, Vingron M, Lehrach H (2003) New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes. Genome Res 13(6A):1056–1066

Paps J, Holland PW, Shimeld SM (2012) A genome-wide view of transcription factor gene diversity in chordate evolution: less gene loss in amphioxus? Brief Funct Genomics 11(2):177–186

Parisod C, Alix K, Just J, Petit M, Sarilar V, Mhiri C, Ainouche M, Chalhoub B, Grandbastien MA (2010) Impact of transposable elements on the organization and function of allopolyploid genomes. New Phytol 186(1):37–45

Paterson AH, Chapman BA, Kissinger JC, Bowers JE, Feltus FA, Estill JC (2006) Many gene and domain families have convergent fates following independent whole-genome duplication events in Arabidopsis, Oryza, Saccharomyces and Tetraodon. Trends Genet 22(11):597–602

Pebusque MJ, Coulier F, Birnbaum D, Pontarotti P (1998) Ancient large-scale genome duplications: phylogenetic and linkage analyses shed light on chordate genome evolution. Mol Biol Evol 15(9):1145–1159

Pendleton JW, Nagai BK, Murtha MT, Ruddle FH (1993) Expansion of the *Hox* gene family and the evolution of chordates. Proc Natl Acad Sci U S A 90:6300–6304

Polak P, Domany E (2006) Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes. BMC Genomics 7:133

Popovici C, Leveugle M, Birnbaum D, Coulier F (2001) Coparalogy: physical and functional clusterings in the human genome. Biochem Biophys Res Commun 288(2):362–370

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(10):481–490

Postlethwait JH (2007) The zebrafish genome in context: ohnologs gone missing. J Exp Zoolog B Mol Dev Evol 308(5):563–577

Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, Benito-Gutierrez EL, Dubchak I, Garcia-Fernandez J, Gibson-Brown JJ, Grigoriev IV, Horton AC, de Jong PJ, Jurka J, Kapitonov VV, Kohara Y, Kuroki Y, Lindquist E, Lucas S, Osoegawa K, Pennacchio LA, Salamov AA, Satou Y, Sauka-Spengler T, Schmutz J, Shin IT, Toyoda A, Bronner-Fraser M, Fujiyama A, Holland LZ, Holland PW, Satoh N, Rokhsar DS (2008) The amphioxus genome and the evolution of the chordate karyotype. Nature 453(7198):1064–1071

Ravi V, Lam K, Tay BH, Tay A, Brenner S, Venkatesh B (2009) Elephant shark (*Callorhinchus milii*) provides insights into the evolution of *Hox* gene clusters in gnathostomes. Proc Natl Acad Sci U S A 106(38):16327–16332

Robinson-Rechavi M, Boussau B, Laudet V (2004) Phylogenetic dating and characterization of gene duplications in vertebrates: the cartilaginous fish reference. Mol Biol Evol 21(3): 580–586

Roux J, Robinson-Rechavi M (2008) Developmental constraints on vertebrate genome evolution. PLoS Genet 4(12):e1000311

Ruvinsky I, Silver LM (1997) Newly identified paralogous groups on mouse chromosomes 5 and 11 reveal the age of a T-box cluster duplication. Genomics 40(2):262–266

SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. Nat Genet 20(1):43–45

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274(5288):765–768

Sauka-Spengler T, Meulemans D, Jones M, Bronner-Fraser M (2007) Ancient evolutionary origin of the neural crest gene regulatory network. Dev Cell 13(3):405–420

Scannell DR, Byrne KP, Gordon JL, Wong S, Wolfe KH (2006) Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. Nature 440(7082):341–345

Semon M, Wolfe KH (2007) Consequences of genome duplication. Curr Opin Genet Dev 17(6):505–512

Semon M, Wolfe KH (2008) Preferential subfunctionalization of slow-evolving genes after allopolyploidization in *Xenopus laevis*. Proc Natl Acad Sci U S A 105(24):8333–8338

Sempere LF, Cole CN, McPeek MA, Peterson KJ (2006) The phylogenetic distribution of metazoan microRNAs: insights into evolutionary complexity and constraint. J Exp Zool B Mol Dev Evol 306(6):575–588

Seo HC, Edvardsen RB, Maeland AD, Bjordal M, Jensen MF, Hansen A, Flaat M, Weissenbach J, Lehrach H, Wincker P, Reinhardt R, Chourrout D (2004) *Hox* cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. Nature 431(7004):67–71

Sharman AC, Holland PWH (1998) Estimation of *hox* gene cluster number in lampreys. Int J Dev Biol 42:617–620

Shemer G, Podbilewicz B (2000) Fusomorphogenesis: cell fusion in organ formation. Dev Dyn 218(1):30–51

Shimeld SM, Holland PW (2000) Vertebrate innovations. Proc Natl Acad Sci U S A 97(9): 4449–4452

Sidow A (1996) Gen(om)e duplications in the evolution of early vertebrates. Curr Opin Genet Dev 6:715–722

Siegel N, Hoegg S, Salzburger W, Braasch I, Meyer A (2007) Comparative genomics of *ParaHox* clusters of teleost fishes: gene cluster breakup and the retention of gene sets following whole genome duplications. BMC Genomics 8:312

Skrabanek L, Wolfe KH (1998) Eukaryote genome duplication—where's the evidence? Curr Opin Genet Dev 8:694–700

Small KS, Brudno M, Hill MM, Sidow A (2007) A haplome alignment and reference sequence of the highly polymorphic *Ciona savignyi* genome. Genome Biol 8(3):R41

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, Depamphilis CW, Wall PK, Soltis PS (2009) Polyploidy and angiosperm diversification. Am J Bot 96(1):336–348

Soltis DE, Soltis PS (1999) Polyploidy: recurrent formation and genome evolution. Trends Ecol Evol 14(9):348–352

Somorjai I, Bertrand S, Camasses A, Haguenauer A, Escriva H (2008) Evidence for stasis and not genetic piracy in developmental expression patterns of *Branchiostoma lanceolatum* and *Branchiostoma floridae*, two amphioxus species that have evolved independently over the course of 200 Myr. Dev Genes Evol 218(11–12):703–713

Spring J (1997) Vertebrate evolution by interspecific hybridization—are we polyploid? Fed Eur Biol Soc Lett 400:2–8

Stadler PF, Fried C, Prohaska SJ, Bailey WJ, Misof BY, Ruddle FH, Wagner GP (2004) Evidence for independent *Hox* gene duplications in the hagfish lineage: a PCR-based gene inventory of *Eptatretus stoutii*. Mol Phylogenet Evol 32(3):686–694

Storchova Z, Pellman D (2004) From polyploidy to aneuploidy, genome instability and cancer. Nat Rev Mol Cell Biol 5(1):45–54

Takio Y, Pasqualetti M, Kuraku S, Hirano S, Rijli FM, Kuratani S (2004) Evolutionary biology: lamprey *Hox* genes and the evolution of jaws. Nature 429(6989):1–262

Taylor JS, Raes J (2004) Duplication and divergence: the evolution of new genes and old ideas. Annu Rev Genet 38:615–643

Taylor JS, van de Peer Y, Meyer M (2001) Genome duplication, divergent resolution and speciation. Trends Genet 17:299–301

Thornburg BG, Gotea V, Makalowski W (2006) Transposable elements as a significant source of transcription regulating signals. Gene 365:104–110

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(10):530–536

Van de Peer Y, Maere S, Meyer A (2009) The evolutionary significance of ancient genome duplications. Nat Rev Genet 10(10):725–732

Venkatesh B, Kirkness EF, Loh YH, Halpern AL, Lee AP, Johnson J, Dandona N, Viswanathan LD, Tay A, Venter JC, Strausberg RL, Brenner S (2007) Survey sequencing and comparative analysis of the elephant shark (*Callorhinchus milii*) genome. PLoS Biol 5(4):e101

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA et al (2001) The sequence of the human genome. Science 291(5507):1304–1351

Vienne A, Shiina T, Abi-Rached L, Danchin E, Vitiello V, Cartault F, Inoko H, Pontarotti P (2003) Evolution of the proto-MHC ancestral region: more evidence for the plesiomorphic organisation of human chromosome 9q34 region. Immunogenetics 55(7):429–436

Wada H, Okuyama M, Satoh N, Zhang S (2006) Molecular evolution of fibrillar collagen in chordates, with implications for the evolution of vertebrate skeletons and chordate phylogeny. Evol Dev 8(4):370–377

Werth CR, Windham MD (1991) A model for divergent, allopatric speciation of polypoid pteridophytes resulting from silencing of duplicate-gene expression. Am Nat 137:515–526

Weston JA (1970) The migration and differentiation of neural crest cells. Adv Morphog 8:41–114

Wolfe K (2000) Robustness—it's not where you think it is. Nat Genet 25(1):3–4

Wolfe KH (2001) Yesterday's polyploids and the mystery of diploidization. Nat Rev Genet 2(5):333–341

Wolfe KH, Shields DC (1997) Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387(6634):708–713

Wotton KR, Shimeld SM (2006) Comparative genomics of vertebrate Fox cluster loci. BMC Genomics 7:271

Yu JK, Meulemans D, McKeown SJ, Bronner-Fraser M (2008) Insights from the amphioxus genome on the origin of vertebrate neural crest. Genome Res 18(7):1127–1132