

OPINION

The evolutionary significance of ancient genome duplications

Yves Van de Peer, Steven Maere and Axel Meyer

Abstract | Many organisms are currently polyploid, or have a polyploid ancestry and now have secondarily 'diploidized' genomes. This finding is surprising because retained whole-genome duplications (WGDs) are exceedingly rare, suggesting that polyploidy is usually an evolutionary dead end. We argue that ancient genome doublings could probably have survived only under very specific conditions, but that, whenever established, they might have had a pronounced impact on species diversification, and led to an increase in biological complexity and the origin of evolutionary novelties.

Most species of flowering plants and vertebrates have descended from ancestors who doubled their genomes, either through autopolyploidy or allopolyploidy. Evidence from cytogenetic analyses, morphological studies of fossil and extant species and, more recently, whole-genome and EST analyses suggests that most (60–70%) flowering plants have a polyploid ancestry^{1,2–4}. In flowering plants, polyploids form at a frequency of 1 per 100,000 individuals⁵, and ~2–4% of speciation events involve polyploidization⁶. As a result, many plants, and most of our domesticated crop species, are polyploid⁷.

Although polyploidy is much rarer in animals than in plants, there are hundreds of known insects and vertebrate species that are polyploid, mainly amphibians and fish⁶. Whole-genome duplications (WGDs) have also been documented for unicellular organisms: the first ancient WGD to be discovered in eukaryotes was that of the yeast *Saccharomyces cerevisiae*⁸. More recently, it was shown that the unicellular ciliate *Paramecium tetraurelia* has also undergone several WGDs⁹.

Because ancient WGDs in plants and animals gave rise to some particularly species-rich groups, some have argued that polyploidy is not an evolutionary dead end but that it provides novel opportunities for evolutionary success^{10–13}. However,

most polyploidy events have occurred near the tips of the evolutionary tree of life rather than at deeper branches. Although many species are currently polyploid, few ancient polyploidy events have survived. During 500–600 million years of vertebrate evolution, no more than two (or three for teleosts) WGDs have persisted. Since the rise of the flowering plants 150–200 million years ago (mya)^{13,14}, the number of inferred ancient WGDs in any angiosperm lineage is at most four^{15,16}. In the fungal lineage, for which many more genome sequences are known, there is only evidence for a single ancient WGD event¹⁷. Paleopolyploidy events therefore seem to be exceedingly rare, and polyploids, or rather their descendants, have not been established tens or hundreds of times. However, all vertebrates seem to have shared two ancient WGD events, whereas all teleosts, and probably also eudicots, are derived from a lineage that experienced a WGD event^{15,18–21}. This would suggest that, although descendants of WGD events do not survive often, when they do survive their evolutionary lineage can be very successful.

The observation that these WGDs often gave rise to species-rich groups of organisms, such as >25,000 species of fish and >350,000 species of flowering plants, suggests that polyploidy can facilitate

diversification and speciation of organisms. Here, we discuss the relationship between WGDs and speciation and argue that most of the ancient WGDs that survived did so because they occurred at specific times: for instance, during major ecological upheavals and periods of extinction. At these times, competition with diploids was reduced and new ecological niches became available. Furthermore, when WGDs survive they can greatly enhance the diversification potential of a lineage through the preferential retention of regulatory genes.

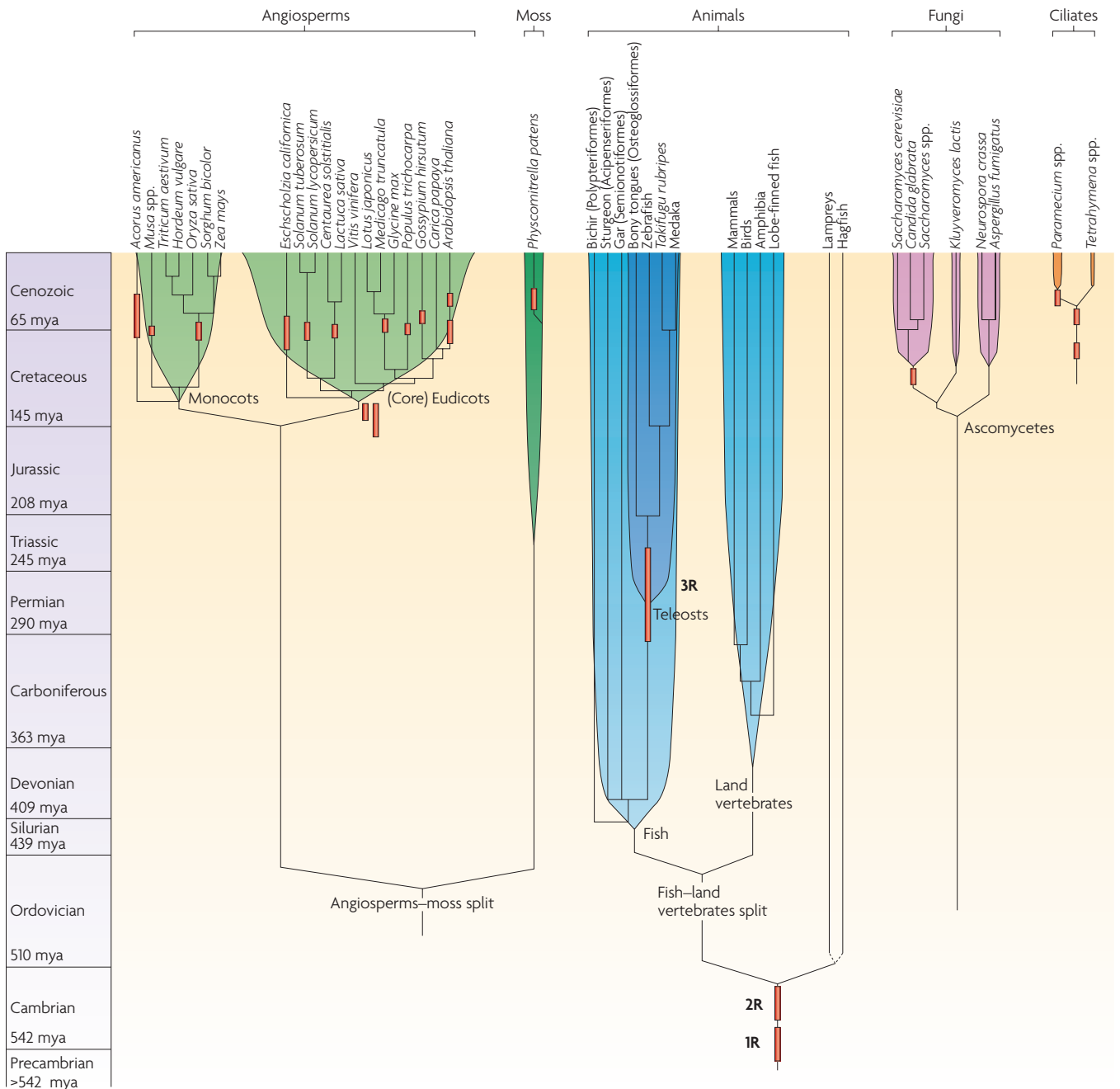
Competitive advantage of polyploids

In the short term, polyploidy may lead to transgressive segregation and increased vigour. In this section we argue that these properties might give newly established polyploids an edge over their diploid progenitors and a wider phenotypic range, thereby increasing their chances of survival.

Reducing the risk of extinction. Crow and Wagner²² have argued that genome duplications could reduce the risk of extinction through several means: by functional redundancy, mutational robustness, and increased rates of evolution and adaptation. Based on the work of Donoghue and Purnell²³, these authors observed that genome duplication events in vertebrate history seem to have been preceded by multiple extinct lineages, resulting in pre-duplication gaps in the phylogeny of extant taxa. By analysing the numbers of families in extinct and extant vertebrate lineages, they concluded that extinction rates were considerably higher for pre-duplication lineages than for post-duplication lineages.

The most compelling evidence that genome duplications might aid in avoiding extinction probably comes from flowering plants. Fawcett *et al.*²⁴ showed that various plants — including legumes, cereals, Solanaceae (such as tomatoes and potatoes), lettuce and cotton — independently underwent a WGD ~60–70 mya. This wave of WGDs occurred close in time to the K–T boundary (BOX 1), suggesting that polyploid plants coped better with the markedly changed environment than their diploid progenitors. Although many

Box 1 | Whole-genome duplications across the phylogeny of eukaryotes



Whole-genome duplications (WGDs) seem to have been followed by a substantial increase in morphological complexity (see the figure, paleopolyploidy events are indicated as red bars and are based on studies published previously for plants^{3,16,24}, fish^{20,55,56}, vertebrates¹⁰¹, fungi¹⁷ and ciliates⁹). The two rounds of genome duplication (1R and 2R) in the vertebrate stem were followed by a period of rapid morphological innovation, which led to: enhanced nervous, endocrine and circulatory systems; enhanced sensory organs; more complex brains; and the skull, vertebrae, the endoskeleton and teeth. These were followed in the jawed vertebrate lineage by innovations such as paired appendages, hinged jaws and an adaptive immune system^{40,102–105}. Many of these innovations are related to the emergence in vertebrates of the neural crest^{40,102}. Since Ohno first suggested that these innovations are facilitated by genome duplications¹⁰⁶, a causal link between the 2R duplication and the emergence of vertebrates has been suggested (see, for example, REFS 78,107).

Similarly, early polyploidization events in one or more angiosperm plant lineages might explain the rapid rise and diversification of angiosperms in the Early Cretaceous period^{6,11,13,14,52,108}. Fundamental innovations that occurred early in angiosperm evolution are the invention of the closed carpel, and the emergence of flowers and of double fertilization¹⁰⁹. These early innovations were elaborated on to create specialized pollination strategies and fruits. The evolution of xylem vessels is also believed to have been important in early angiosperm diversification⁸⁵, but their origin is less clear. Some basal angiosperms, such as *Amborella* spp., lack vessels, whereas vessel structures have been discovered in members of the Gnetales order and in ferns^{85,110,111}. However, the diversity of vasculature in angiosperms is unparalleled. The timing of the early angiosperm polyploidizations is unclear. It is possible that they contributed to the elaboration and diversification of the aforementioned inventions rather than to their establishment.

changes associated with polyploidization are probably disadvantageous or deleterious^{6,11,12}, it seems that many K–T polyploids outcompeted their diploid progenitors, probably owing to a higher tolerance of a wider range of environmental conditions^{25,26}. Alternatively, in a more ‘neutral’ scenario, one could assume that environmental stress leads to an increased incidence of polyploid formation: for instance, through the production of unreduced, 2n gametes²⁷. In this case, the cataclysmic events that were responsible for the K–T extinction could have increased the establishment of polyploid lineages by chance. However, it is unclear whether such an increase alone could explain the extent to which polyploid plants replaced or overshadowed their diploid relatives.

Increased vigour. In the adaptive scenario, heterotic effects and rapid genomic and epigenetic changes underlie the ability of polyploids to quickly adapt to more extreme environments. In allopolyploids and autopolyploids, increased heterozygosity can lead to increased variation in gene expression and in regulatory wiring²⁸, which may result in increased vigour and faster adaptation to novel conditions^{29,30}. Rapid genomic and epigenetic changes after WGD may similarly lead to increased variation and transgressive traits²⁸.

Transgressive segregation in polyploids might serve as a pre-adaptation for survival in habitats that were not accessible to their diploid parent species^{22,31}. Several studies have suggested that polyploid plants are more tolerant to a wider range of environmental conditions compared with their diploid relatives^{25,26}. Furthermore, many polyploids are invasive^{32,33} and can exploit habitats that their diploid progenitors cannot^{26,34}. Polyploid insects also have a wider geographical distribution than their diploid progenitors, often colonizing northern and mountain regions³⁵. One of the rare examples of relatively recent polyploidy establishment in vertebrates is given by the tetraploid frog *Xenopus laevis*, which is a highly invasive species that colonizes disturbed and man-made habitats. It is also extremely tolerant to salt, drought, cold and starvation, and is more disease resistant than its diploid relative *Silurana tropicalis*^{36,37}.

In summary, increased phenotypic variability and heterotic effects have the potential to enable polyploids to survive environmental conditions that do not favour their diploid ancestors³⁸ (FIG. 1). Polyploidy is also known to facilitate self-fertilization and the formation of asexually reproducing (apomictic) species^{35,39}, which might

be a selective advantage when sexual mates are scarce. Following this logic, environmental upheaval may have been a driving force in shaping survivorship probabilities associated with genome duplication²²; the clustered genome duplications in flowering plants at the K–T boundary provide a tantalizing example.

However, owing to uncertainties in the dating of most ancient WGDs, a more general link between WGDs and major extinctions cannot be ascertained. The 2R WGD event in vertebrates may date from 520–550 mya^{19,40}, close to the mass extinction at the dawn of the Cambrian explosion (542 mya^{41–43}), and the genome duplication in teleosts, which according to the most recent estimate happened 226–316 mya⁴⁴, may have occurred close to the Permian–Triassic (P–T) mass extinction event (250 mya). For other paleopolyploidies — for example, in *S. cerevisiae* and the core eudicots — there is no indication that they are linked to mass extinction events.

Increased species diversity

Genome duplications often seem to be accompanied by marked and sudden increases in species richness. Although a link between any specific genome duplication event and increased species diversity remains correlational rather than causal, several mechanisms might explain how gene duplication facilitates the formation of novel species.

Reciprocal gene loss. Both Werth and Windham⁴⁵ and Lynch and Force⁴⁶ proposed that the loss of different copies of a duplicated gene in separated populations might genetically isolate these populations (FIG. 2). Divergent resolution of the thousands to tens of thousands of genes and regulatory RNAs that are produced by a genome duplication event could therefore potentially facilitate speciation.

Scannell *et al.*⁴⁷ studied gene loss in three yeast species that have undergone a WGD and showed that, at many loci, different species lost different members of a duplicated pair, so that 4–7% of single-copy genes compared between any two species are not orthologues but paralogues. Such a pattern provides strong evidence for speciation through the reciprocal gene loss (RGL) model^{45,46} (FIG. 2a). Similar findings have been reported for duplicated fish genomes, in which it is estimated that ~1,700 (8%) ancestral loci of *Tetraodon nigroviridis* and zebrafish underwent RGL⁴⁸. Because RGL at only a few pairs of loci that encode

essential genes would be sufficient to result in reproductive isolation, it was concluded that RGL at duplicated loci might contribute to speciation events that occurred after the teleost WGD⁴⁸.

No similar studies have been performed for plants, but recent experimental work has provided evidence that reciprocal silencing or loss of duplicated genes provides an important source of epistatic interactions that follow the Bateson–Dobzhansky–Muller model. Bikard *et al.*⁴⁹ show that, in crosses between different accessions of *Arabidopsis thaliana*, loci interact in an epistatic manner to control a recessive embryo lethality. This effect is explained by divergent evolution occurring among paralogues of an essential duplicated gene when the functional copy is not located at the same locus in the different accessions; this results in lowered fitness in the first or second filial (F₁ or F₂) generations of certain crosses, which contributes to reproductive isolation. By demonstrating the link between gene duplication and genetic incompatibility, the authors provide direct evidence for duplicate gene loss as a neutral mechanism that generates post-zygotic isolating barriers within existing species or populations.

Subfunctionalization. Other neutral scenarios might also promote speciation. One example would be a case in which both copies of a gene that has multiple functions (for instance, it is expressed at different stages in development or in different tissues) are retained in different populations after a duplication event. Should the populations become geographically isolated, the two duplicate genes in each population could subfunctionalize⁴⁶ and the orthologues in the different populations might evolve different functions. The resulting F₁ hybrids from the two populations would develop correctly because each subfunction is performed by one of the genes from each population. However, one-eighth of the F₂ zygotes will lack one of the subfunctions and will die if this function is essential^{50,51} (FIG. 2b). As a result, lineage-specific subfunction partitioning could accelerate rates of speciation.

Speciation. There seems to be a correlation between WGDs in plants and increased rates of speciation or divergence. First, there seems to be a correlation between the older WGDs and the early and fast diversification of flowering plants^{52,53}. Second, Soltis *et al.*¹³ found a strong correlation between diversification rates and polyploidy following recent WGDs in many plant lineages. For instance, the WGD in the Poaceae lineage

possibly coincides with the origin and divergence of the core Poaceae, a large clade containing ~10,000 species. Early-branching subclades of the Poaceae, as well as closely related non-Poaceae families, contain only a small number of species. Whole-genome duplications have also been reported for the Brassicaceae (3,700 species), Asteraceae (23,000 species), the Fabaceae (19,400 species) and the Solanaceae (>3,000 species), to name but a few, and these WGDs also seem to correlate with species-rich plant families, although the precise phylogenetic placement of these WGDs is unclear¹³. Furthermore, the rate of diversification is also high in these families compared with other families in the same orders⁵⁴.

In fish, the correlation between WGD and species diversification rates is less clear. Fish constitute half of all vertebrate species and are a highly successful and diverse evolutionary lineage²¹. The fish-specific genome duplication (3R) in the teleost lineage is estimated to have occurred 226–350 mya^{44,55–57}. The inferred phylogenetic timing of 3R seems to separate the species-poor, early-branching lineages of ray-finned fish from the species-rich teleost lineage, and therefore seems to provide evidence that 3R might be causally related to an increase in species and biological diversity. However, there is a large period of time between 3R and the main teleost radiations, which, according to fossil evidence, did not occur until the

Late Cretaceous and Tertiary periods, more than 150 million years later. This observation could be taken to indicate that genome duplication was not an important factor in the rapid radiation of teleosts. However, both RGL and subfunction partitioning can occur over tens of millions of years after a WGD and can continue to promote speciation over long periods of time^{47,48}. It is conceivable that 3R continued to increase the propensity for speciation until a suitable ecological occasion presented itself, such as the K–T mass extinction. As an example of such stored diversifying potential, *X. laevis* still maintains ~32–47%⁵⁸ of its genes in duplicate, some 40 million years after its most recent polyploidization event, and its genome

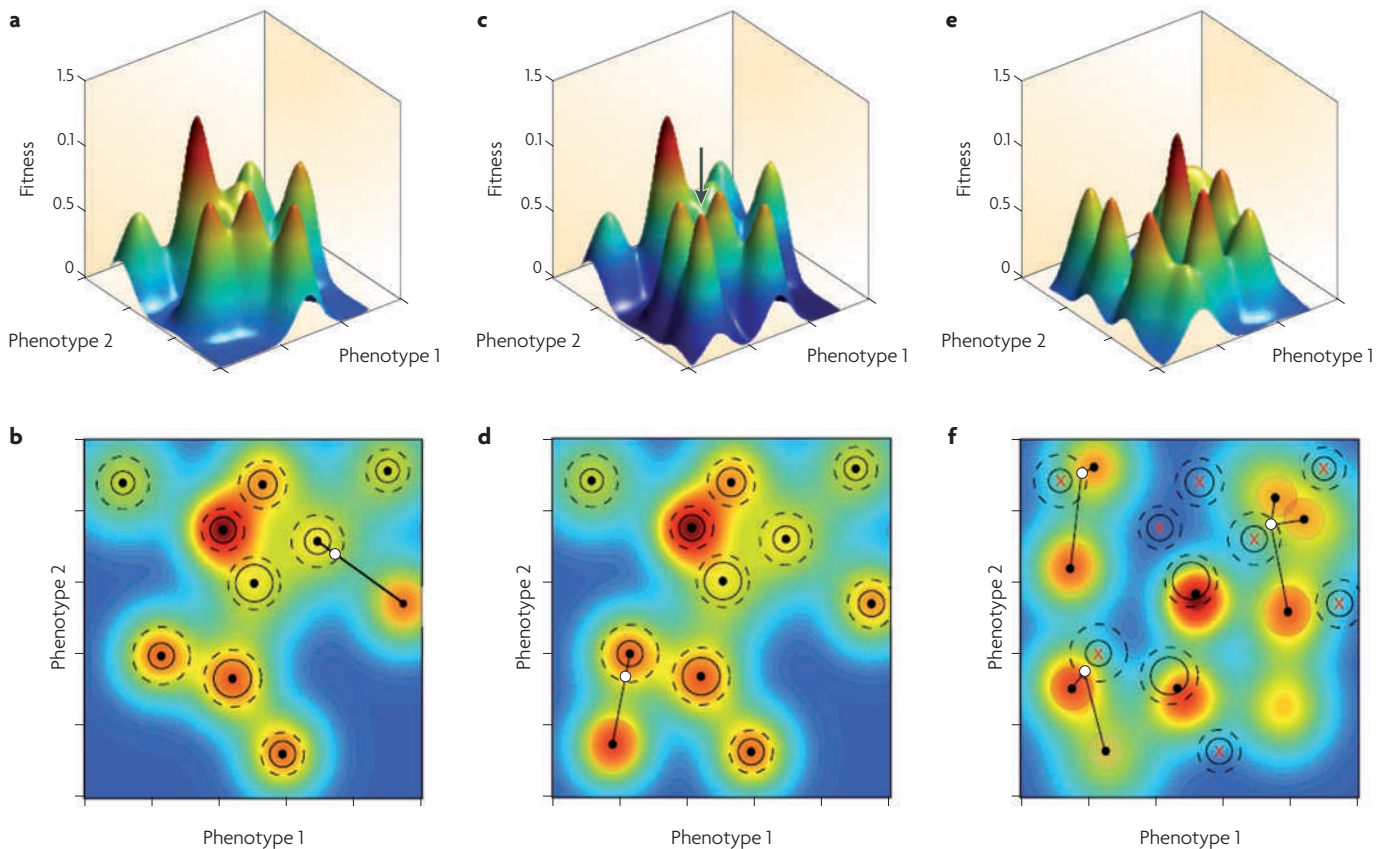


Figure 1 | Survival of the fittest. The figure illustrates one of many^{92,95,112–115} simplified fitness landscape models. The upper and lower panels show the fitness landscape with two imaginary phenotype axes, 1 and 2. These axes do not represent single quantitative traits but rather a flattened version of phenotype space. The black dots represent well-adapted organisms that occupy the peaks in phenotype space (red indicates the most well adapted, blue the least well adapted), which correspond to niches in which that particular combination of phenotypic characters is advantageous. The full circles represent the phenotypes accessible to the organisms, whereas the dashed circles are a simplified representation of the phenotype space of their polyploid relatives. Blue regions of the phenotype space are not viable, so there is little room for successful genome duplication events. **a–d** | In one scenario, there is an unoccupied peak in the fitness landscape (**a,b**) or a new fitness peak emerges (**c,d**), for instance, through evolution of

a new niche (the new peak is indicated by an arrow in **c**). None of the existing species has the evolutionary potential to fill this niche, but a polyploid species (white dot in **b** and **d**) may be able to develop the necessary phenotypic innovations. **e,f** | In another scenario, the fitness landscape changes drastically, for example, through a catastrophic event. Most organisms cannot adapt to the changed environment and perish (red crosses). Some organisms (near the centre of the landscape) live in relatively unaltered niches and can adapt enough to survive. Others may manage to survive initially through polyploidization (white dots), outcompeting their diploid parents because of, for example, heterotic effects. These polyploids also harbour the potential to develop innovations that in time may enable them to colonize empty niches in phenotype space that cannot be reached by other organisms. Differential realization of this potential among the polyploid offspring may lead to phenotype diversification and speciation.

shows little evidence of duplicate subfunctionalization or neofunctionalization^{58–60}.

Any theory that attempts to link WGD to species diversity should take into account the fact that radiations are not always preceded by genome duplications. Invertebrates and vertebrates have diversified at similar rates⁶¹, despite the fact that the vertebrates underwent two rounds of genome duplication and the invertebrates none.

Evolutionary innovations

In the longer run, polyploidy may pave the way for evolutionary innovations or elaborations of existing morphological structures that allow exploration of fundamentally different regions of phenotype space.

Genome duplication favours gene retention.

One of the prerequisites for developing more complex systems is an increase in the number of gene regulators⁶². Intriguingly, genome duplications are the preferred way to accomplish such an increase. Transcriptional and developmental regulators and signal transducers have been preferentially retained in duplicate after all genome duplications in *Arabidopsis thaliana*^{63–65}, after the 1R and 2R WGDs in vertebrates^{19,66}, after 3R in fish^{66,67}, and after the WGD in yeast^{68,69}. Moreover, these regulatory gene classes cannot be expanded easily through single-gene duplications, which accentuates the importance of genome duplications in expanding the regulatory gene repertoire. Maere *et al.*⁶³ estimated that more than 90% of the increase in regulatory genes in the *Arabidopsis* lineage in the last ~150 million years is caused by genome duplications. Both the under-retention of regulators after single-gene duplications and their over-retention after genome duplications can be explained by dosage balance effects^{70,71}. Freeling and Thomas⁷² and Freeling⁷³ argue that, after genome duplication, entire functional modules are inherently retained in duplicate through non-adaptive dosage balance effects, after which they can adaptively evolve novel functions and might ultimately cause an increase in morphological complexity.

The study of individual gene families also points to the importance of genome duplications in expanding the regulatory gene repertoire of an organism. In plants, important developmental regulators, such as the AUX/IAA family of auxin response regulators⁷⁴ and certain MADS-box transcription factor subfamilies^{75,76}, seem to have expanded mainly through genome duplications. In vertebrates, 1R and 2R are thought to be responsible for the expansion of the number of homeobox

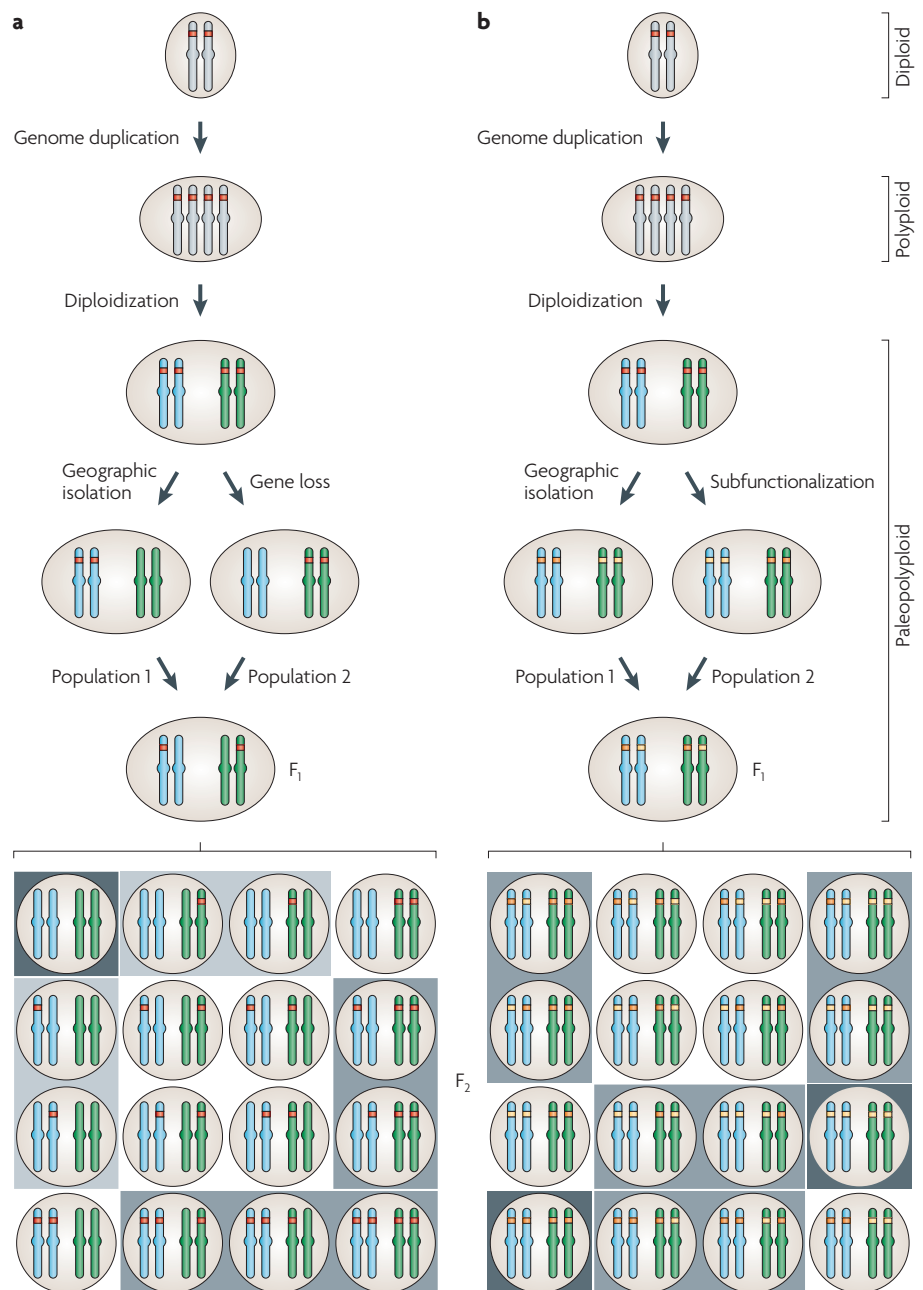


Figure 2 | Reciprocal gene loss or subfunctionalization facilitates speciation. Red bands on chromosomes represent a locus that is duplicated (along with all other loci) during a tetraploidization event. **a** | After diploidization, the duplicated gene is present on two different chromosomes. After geographic isolation, both populations have lost one of the duplicates on different chromosomes. If individuals from isolated populations mate, their ‘hybrid’ progeny would be heterozygous, possessing a functional allele at each locus of the duplicated gene. However, one-sixteenth (approximately 6%) of crosses between the first filial (F_1) individuals produce second filial (F_2) individuals that have null alleles at both loci in question (dark grey square) and therefore lack viability and/or fertility. Others might receive one allele (light grey squares), which might reduce functionality when a gene is haploinsufficient, or might receive three or four functional alleles (mid-grey squares), which might have a negative dosage effect. All these outcomes might lead to post-mating reproductive isolation⁴⁶. **b** | In this scenario, after diploidization and geographic isolation, the duplicated genes in the different populations have subfunctionalized (orange and yellow bands on chromosomes). Hybrids between the two populations should in general develop normally, but one-sixteenth of the F_2 generation will be homozygous for alleles lacking one essential subfunction, and another one-sixteenth will be homozygous for alleles lacking the other essential subfunction (dark grey squares), thus reducing the fitness of hybrids. Other F_2 individuals might, as in **a**, show reduced fitness caused by dosage or haploinsufficiency effects.

(Hox) clusters and other Hox genes^{40,77,78}, transforming growth factor- β pathway genes⁷⁹, insulin receptors⁸⁰, nuclear receptors⁸¹ and genes that specify the neural crest⁴⁰.

Increase in complexity. It is unclear whether or not polyploidy caused the evolution of the defining innovations in angiosperm and vertebrate lineages (BOX 1). Rather than facilitating innovation from scratch, the power of genome duplications may be their ability to perfect primitive versions or precursors of innovative features and fully exploit their potential, for example, by lifting constraints on pleiotropic genes and facilitating their co-option for specialized purposes^{40,82}. It is conceivable that an increase in regulatory gene complexity fuelled by WGD would tend to increase the potential of an organism to become more complex, providing a 'drive' towards more complex organisms⁷². But the

fact that genome duplications provide the raw material for increased complexity does not imply that they should always lead to more complex organisms⁸³. For instance, there is no indication that morphological complexity increased substantially after the WGD in *S. cerevisiae* and, despite undergoing three genome duplications, *P. tetraurelia* is still a unicellular (although well evolved) organism. Accordingly, the over-retention of regulatory duplicates after the WGDs in these organisms is less pronounced⁹.

In this respect, the duplicated genome, although maybe not immediately useful, could be regarded as a genomic 'spandrel'⁸⁴ that occasionally might have been used for adaptive or complexity-increasing embellishments. An increased rate of speciation after polyploidy could have facilitated this process by providing a lineage with more opportunities to sample phenotype space. If

enough species roam the fitness landscape, for some species further changes are likely to become adaptive or previous changes may be co-opted for a novel purpose.

Donoghue and Purnell²³ argued against a link between genome duplications and increasing complexity based on the observation that when extinct lineages are taken into account, there are no bursts in morphological innovation or jumps in complexity in post-WGD clades. However, morphological evolution after WGD does not need to be saltational. Genome duplication merely enhances the diversification potential of a lineage; the ensuing process of morphological elaboration is likely to take time and to spawn intermediate forms that might go extinct because they are later outcompeted by more derived relatives^{43,85}. As with species diversity, a better indicator of the diversifying force of WGDs is obtained by comparing morphological innovations in WGD clades with those of their non-WGD sister clades (or their closest living non-WGD relatives). Basal chordates (such as urochordates and cephalochordates) do not exhibit the morphological evolution of vertebrates; for example, amphioxus is considered a living fossil⁸⁶ (for other examples, see REFS 4,21).

Concerning the more recent wave of genome duplications around the K–T boundary in plants, the phylogenetic placement of the WGD events is too uncertain to allow the accurate identification of WGD and non-WGD sister clades¹³, so an assessment of the correlation or causation between the occurrence of these WGDs and morphological innovations is difficult. However, the fact that these WGDs have occurred in many of the most species-rich and morphologically diverse angiosperm families is probably no coincidence^{2,4,13,87–89}. It is not known whether genome duplications are also involved in the diversification of other large families, such as the orchids, although the morphological innovations in orchid flowers have been linked to an expansion of DEF-like MADS-box genes that could have been caused by genome duplication^{90,91}.

Conclusions and perspectives

Arguably the greatest consequence of polyploidy is an increase in the attainable 'morphospace'. But the potential for phenotypic enhancement provided by WGDs is less useful when there are no niches in which the newly available phenotypes are advantageous. In stable ecosystems, newly formed polyploids are probably not able to compete with the highly adapted occupants of existing niches, including their diploid

Glossary

Accession

A sample of a plant variety collected at a specific location and time. This term is used to describe the *Arabidopsis thaliana* laboratory lines collected initially from the wild.

Allopolyploidy

The generation of the polyploid state by the fusion of nuclei from different species. For example, two fertilized diploid oocytes can fuse such that the newly formed single egg has two complete sets of chromosomes.

Autopolyploidy

In contrast to allopolyploidy, different sets of chromosomes are derived from the same species. This can occur in the fertilized oocyte if the nucleus divides but the cell does not.

Bateson–Dobzhansky–Muller model

Describes incompatibilities between organisms on the basis of the synergistic interaction of genes that have functionally diverged among the respective parents. Such incompatibilities can lead to speciation.

Carpel

A leaf-like structure that encloses the ovules and seeds and is the defining characteristic of flowering plants. In some species, multiple carpels might be present in a compound structure called an ovary.

Dosage balance effects

The components of macromolecular complexes must be balanced to avoid dominant fitness defects. Therefore, both under- and overexpression of individual protein subunits within a complex — for example, through duplication — tend to lower fitness.

Haploinsufficient

Describes the situation in which a lower than normal amount of a wild-type gene product confers a detectable phenotype.

Heterosis

The greater fitness of a hybrid individual carrying different alleles of genes relative to either of the two corresponding homozygous parents. Also called hybrid

vigour. A more precise definition is non-additive

inheritance, in which a trait in the first filial generation transgresses both parental values.

K–T boundary

The K–T event — which occurred ~65 million years ago at the end of the Cretaceous period and the beginning of the Tertiary — is the most recent large-scale mass extinction of animal and plant species. There is general consensus that the K–T extinction was caused by one or more catastrophic events, such as a massive asteroid impact and increased volcanic activity.

Mutational robustness

Describes the extent to which the phenotype of an organism remains constant in spite of mutations. If an organism has an extra copy of a gene through gene or genome duplication, the effect of the loss of one copy might be limited.

Neural crest

A migratory cell population that gives rise to numerous differentiated cell types in vertebrates.

Orthologues

Loci in two species that are derived from a common ancestral locus by a speciation event.

Paralogues

Genes in the same organism that have evolved from a gene duplication, usually with a subsequent, sometimes subtle, divergence of function.

Phenotype space

A multi-dimensional continuum of all possible phenotypes.

Pleiotropic gene

A gene that is responsible for several distinct and seemingly unrelated phenotypic effects.

Transgressive segregation

Refers to the formation of extreme phenotypes that are observed in segregating hybrid populations when compared with parental lines.

ancestors^{11,85,92–95}. Therefore, we argue that the availability of ecological niches or severely perturbed ecosystems could be the single most important determinant for the survival and long-term evolutionary success of a WGD.

Mass extinctions are among the most drastic events by which old or new niches become available for colonization (FIG. 1e,f). WGDs occurring close to these extinctions probably contributed to the radiation of vertebrates in the Cambrian period and of several angiosperm families at the K–T boundary. The K–T mass extinction might also have played an important part in unlocking the diversification potential of 3R in teleosts, even >150 million years after 3R occurred. Conversely, the reason why the teleosts did not diversify right after the P–T mass extinction may have been because the few survivors from the Triassic period still occupied most of the relevant niche space⁹⁵.

However, new niches may also become available through biotic evolution. For instance, the rise of angiosperm plants led to the emergence of sugar-rich fruits. Conant and Wolfe⁹³ have suggested that the success of the genome duplication in budding yeast, approximately 100 mya, may be linked to the emergence of this new ecological niche. They showed that the retention of glycolytic pathway genes after the WGD in yeasts supported an increase in glycolytic flux that gave post-WGD yeast species a growth advantage in glucose-rich environments (FIG. 1c,d).

The angiosperms also did not rise to ecological dominance by filling niches that became available after a mass extinction event⁸⁵. It is possible that angiosperms filled niches in phenotype space that already existed but that had remained largely unoccupied because the necessary phenotypic characteristics had not yet been developed (FIG. 1a,b). Specifically, angiosperm–insect interactions may have been important in angiosperm niche diversification. Although insect pollination evolved in several non-angiosperm plants, such as *Welwitschia mirabilis* (a gnetophyte), angiosperms developed several innovations that dramatically increased the effectiveness of insect pollination, such as the association of male and female reproductive organs on the same axis and the development of colourful perianth organs^{96–98}. These specialized angiosperm–insect associations allow efficient pollination over large distances, which might have enabled angiosperm plants to colonize previously unoccupied habitats, such as dispersed microhabitats, or disturbed or resource-poor habitats. Indeed, some of the major classes

of pollinating insects, Lepidoptera and bees, appear in the fossil record at approximately the same time as angiosperms^{96,99,100}. Early angiosperm polyploidizations occurring at this time might have helped plants to conquer these newly unclosed niches.

It may prove difficult to determine whether polyploidy enabled organisms to survive extinctions or whether polyploidy facilitated evolutionary transitions and increased biological complexity. Sequencing more genomes and developing tools that are more able to detect and correctly date ancient polyploidy events may unveil correlations between polyploidy and evolutionary changes that are currently unknown, and studying the genes remaining after a WGD and their interactors at a systems level may provide clues as to why polyploids occasionally might have had a selective advantage over their diploid sister species.

Yves Van de Peer and Steven Maere are at the Department of Plant Systems Biology, VIB (Flanders Institute of Biotechnology),

B-9052 Ghent, Belgium, and the Bioinformatics and Evolutionary Genomics research group, Department of Plant Biotechnology and Genetics, Ghent University, B-9052 Ghent, Belgium.

Axel Meyer is at the Department of Biology, University of Konstanz, D-78457 Konstanz, Germany, and the Institute for Advanced Study, Wallotstraße 19, D-14193 Berlin, Germany.

Correspondence to Y.V.d.P.
e-mail: Yves.Vandeppeer@psb.vib-ugent.be

doi:10.1038/nrg2600

Published online 4 August 2009

1. Masterston, J. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* **264**, 421–423 (1994).
2. Blanc, G. & Wolfe, K. H. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* **16**, 1667–1678 (2004).
3. Tang, H. *et al.* Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. *Genome Res.* **18**, 1944–1954 (2008).
4. Cui, L. *et al.* Widespread genome duplications throughout the history of flowering plants. *Genome Res.* **16**, 738–749 (2006).
5. Ramsey, J. & Schemske, D. W. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annu. Rev. Ecol. Syst.* **29**, 467–501 (1998).
6. Otto, S. P. & Whitton, J. Polyploid extinction and evolution. *Annu. Rev. Genet.* **34**, 401–437 (2000).
7. Soltis, P. S. & Soltis, D. E. The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.* **60**, 561–588 (2009).
8. Wolfe, K. H. & Shields, D. C. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708–713 (1997).
9. Aury, J. M. *et al.* Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature* **444**, 171–178 (2006).
10. Wittbrodt, J., Meyer, A. & Schartl, M. More genes in fish? *Bioessays* **20**, 511–515 (1998).
11. Otto, S. P. The evolutionary consequences of polyploidy. *Cell* **131**, 452–462 (2007).
12. Comai, L. The advantages and disadvantages of being polyploid. *Nature Rev. Genet.* **6**, 836–846 (2005).
13. Soltis, D. E. *et al.* Polyploidy and angiosperm diversification. *Am. J. Bot.* **96**, 336–348 (2009).
14. Soltis, D. E., Bell, C. D., Kim, S. & Soltis, P. S. Origin and early evolution of angiosperms. *Ann. NY Acad. Sci.* **1133**, 3–25 (2008).
15. Tang, H. *et al.* Synteny and collinearity in plant genomes. *Science* **320**, 486–488 (2008).
16. Jaillon, O. *et al.* The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**, 463–467 (2007).
17. Scannell, D. R., Butler, G. & Wolfe, K. H. Yeast genome evolution — the origin of the species. *Yeast* **24**, 929–942 (2007).
18. Dehal, P. & Boore, J. L. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* **3**, e314 (2005).
19. Putnam, N. H. *et al.* The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453**, 1064–1071 (2008).
20. Jaillon, O. *et al.* Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **431**, 946–957 (2004).
21. Meyer, A. & Van de Peer, Y. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *Bioessays* **27**, 937–945 (2005).
22. Crow, K. D. & Wagner, G. P. What is the role of genome duplication in the evolution of complexity and diversity? *Mol. Biol. Evol.* **23**, 887–892 (2006).
23. Donoghue, P. C. J. & Purnell, M. A. Genome duplication, extinction and vertebrate evolution. *Trends Ecol. Evol.* **20**, 312–319 (2005).
24. Fawcett, J. A., Maere, S. & Van de Peer, Y. Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. *Proc. Natl Acad. Sci. USA* **106**, 5737–5742 (2009).
25. Levin, D. A. Polyploidy and novelty in flowering plants. *Am. Nat.* **122**, 1–25 (1983).
26. Thompson, J. D. & Lumaret, R. The evolutionary dynamics of polyploid plants — origins, establishment and persistence. *Trends Ecol. Evol.* **7**, 302–307 (1992).
27. Bretagnolle, F. & Thompson, J. D. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytol.* **129**, 1–22 (1995).
28. Osborn, T. C. *et al.* Understanding mechanisms of novel gene expression in polyploids. *Trends Genet.* **19**, 141–147 (2003).
29. Rieseberg, L. H. *et al.* Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica* **129**, 149–165 (2007).
30. Rieseberg, L. H. *et al.* Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**, 1211–1216 (2003).
31. Hegarty, M. J. *et al.* Changes to gene expression associated with hybrid speciation in plants: further insights from transcriptomic studies in *Senecio*. *Philos. Trans. R. Soc. Lond. B* **363**, 3055–3069 (2008).
32. Ellstrand, N. C. & Schierenbeck, K. A. Hybridization as a stimulus for the evolution of invasiveness in plants? *Proc. Natl Acad. Sci. USA* **97**, 7043–7050 (2000).
33. Pandit, M. K., Tan, H. T. W. & Bisht, M. S. Polyploidy in invasive plant species of Singapore. *Bot. J. Linn. Soc.* **151**, 395–403 (2006).
34. Soltis, D. E., Soltis, P. S. & Tate, J. A. Advances in the study of polyploidy since plant speciation. *New Phytol.* **161**, 173–191 (2003).
35. Lokki, J. & Saura, A. Polyploidy in insect evolution. *Basic Life Sci.* **13**, 277–312 (1979).
36. Weldon, C., du Preez, L. H., Hyatt, A. D., Muller, R. & Spears, R. Origin of the amphibian chytrid fungus. *Emerg. Infect. Dis.* **10**, 2100–2105 (2004).
37. Parker, J. M., Mikaelian, I., Hahn, N. & Diggs, H. E. Clinical diagnosis and treatment of epidermal chytridiomycosis in African clawed frogs (*Xenopus tropicalis*). *Comp. Med.* **52**, 265–268 (2002).
38. Hegarty, M. & Hiscock, S. Polyploidy: doubling up for evolutionary success. *Curr. Biol.* **17**, R927–R929 (2007).
39. Bicknell, R. A. & Koltunow, A. M. Understanding apomixis: recent advances and remaining conundrums. *Plant Cell* **16**, S228–S245 (2004).
40. Holland, L. Z. *et al.* The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res.* **18**, 1100–1111 (2008).
41. Xiao, S. & Laflamme, M. On the eve of animal radiation: phylogeny, ecology and evolution of the Ediacara biota. *Trends Ecol. Evol.* **24**, 31–40 (2009).
42. Wille, M., Nagler, T. F., Lehmann, B., Schroder, S. & Kramers, J. D. Hydrogen sulphide release to surface waters at the Precambrian/Cambrian boundary. *Nature* **453**, 767–769 (2008).
43. Knoll, A. H. & Carroll, S. B. Early animal evolution: emerging views from comparative biology and geology. *Science* **284**, 2129–2137 (1999).
44. Hurley, I. A. *et al.* A new time-scale for ray-finned fish evolution. *Proc. R. Soc. B* **274**, 489–498 (2007).

45. Werth, C. R. & Windham, M. D. A model for divergent, allopatric speciation of polyploid pteridophytes resulting from silencing of duplicate-gene expression. *Am. Nat.* **137**, 515–526 (1991).
46. Lynch, M. & Force, A. The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**, 459–473 (2000).
47. Scannell, D. R., Byrne, K. P., Gordon, J. L., Wong, S. & Wolfe, K. H. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature* **440**, 341–345 (2006).
48. Semon, M. & Wolfe, K. H. Reciprocal gene loss between Tetraodon and zebrafish after whole genome duplication in their ancestor. *Trends Genet.* **23**, 108–112 (2007).
49. Bikard, D. *et al.* Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science* **323**, 623–626 (2009).
50. Postlethwait, J., Amores, A., Cresko, W., Singer, A. & Yan, Y. L. Subfunction partitioning, the teleost radiation and the annotation of the human genome. *Trends Genet.* **20**, 481–490 (2004).
51. Volff, J. N. Genome evolution and biodiversity in teleost fish. *Heredity* **94**, 280–294 (2005).
52. De Bodt, S., Maere, S. & Van de Peer, Y. Genome duplication and the origin of angiosperms. *Trends Ecol. Evol.* **20**, 591–597 (2005).
53. Bowers, J. E., Chapman, B. A., Rong, J. & Paterson, A. H. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**, 433–438 (2003).
54. Magallón, S. & Castillo, A. Angiosperm diversification through time. *Am. J. Bot.* **96**, 349–365 (2009).
55. Vandepoelle, K., De Vos, W., Taylor, J. S., Meyer, A. & Van de Peer, Y. Major events in the genome evolution of vertebrates: paraneome age and size differ considerably between ray-finned fishes and land vertebrates. *Proc. Natl Acad. Sci. USA* **101**, 1638–1643 (2004).
56. Christoffels, A., Koh, E. G., Brenner, S., Aparicio, S. & Venkatesh, B. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. *Mol. Biol. Evol.* **21**, 1146–1151 (2004).
57. Hoegg, S., Brinkmann, H., Taylor, J. S. & Meyer, A. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. *J. Mol. Evol.* **59**, 190–203 (2004).
58. Semon, M. & Wolfe, K. H. Preferential subfunctionalization of slow-evolving genes after allopolyploidization in *Xenopus laevis*. *Proc. Natl Acad. Sci. USA* **105**, 8353–8358 (2008).
59. Chain, F. J. & Evans, B. J. Multiple mechanisms promote the retained expression of gene duplicates in the tetraploid frog *Xenopus laevis*. *PLoS Genet.* **2**, e56 (2006).
60. Chain, F. J., Iliava, D. & Evans, B. J. Duplicate gene evolution and expression in the wake of vertebrate allopolyploidization. *BMC Evol. Biol.* **8**, 43 (2008).
61. McPeck, M. A. & Brown, J. M. Clade age and not diversification rate explains species richness among animal taxa. *Am. Nat.* **169**, E97–E106 (2007).
62. Sole, R. V., Fernandez, P. & Kauffman, S. A. Adaptive walks in a gene network model of morphogenesis: insights into the Cambrian explosion. *Int. J. Dev. Biol.* **47**, 685–693 (2003).
63. Maere, S. *et al.* Modeling gene and genome duplications in eukaryotes. *Proc. Natl Acad. Sci. USA* **102**, 5454–5459 (2005).
64. Seoighe, C. & Gehring, C. Genome duplication led to highly selective expansion of the *Arabidopsis thaliana* proteome. *Trends Genet.* **20**, 461–464 (2004).
65. Blanc, C. & Wolfe, K. H. Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* **16**, 1679–1691 (2004).
66. Blomme, T. *et al.* The gain and loss of genes during 600 million years of vertebrate evolution. *Genome Biol.* **7**, R43 (2006).
67. Brunet, F. G. *et al.* Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. *Mol. Biol. Evol.* **23**, 1808–1816 (2006).
68. Seoighe, C. & Wolfe, K. H. Yeast genome evolution in the post-genome era. *Curr. Opin. Microbiol.* **2**, 548–554 (1999).
69. Davis, J. C. & Petrov, D. A. Preferential duplication of conserved proteins in eukaryotic genomes. *PLoS Biol.* **2**, E55 (2004).
70. Papp, B., Pal, C. & Hurst, L. D. Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**, 194–197 (2003).
71. Birchler, J. A., Riddle, N. C., Auger, D. L. & Veitia, R. A. Dosage balance in gene regulation: biological implications. *Trends Genet.* **21**, 219–226 (2005).
72. Freeling, M. & Thomas, B. C. Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. *Genome Res.* **16**, 805–814 (2006).
73. Freeling, M. Bias in plant gene content following different sorts of duplication: tandem, whole-genome segmental, or by transposition. *Annu. Rev. Plant Biol.* **60**, 433–453 (2009).
74. Remington, D. L., Vision, T. J., Guilfoyle, T. J. & Reed, J. W. Contrasting modes of diversification in the *Aux/IAA* and *ARF* gene families. *Plant Physiol.* **135**, 1738–1752 (2004).
75. Veron, A. S., Kaufmann, K. & Bornberg-Bauer, E. Evidence of interaction network evolution by whole-genome duplications: a case study in MADS-box proteins. *Mol. Biol. Evol.* **24**, 670–678 (2007).
76. Zahn, L. M. *et al.* The evolution of the *SEPALLATA* subfamily of MADS-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history. *Genetics* **169**, 2209–2223 (2005).
77. Holland, P. W. & Garcia-Fernandez, J. *Hox* genes and chordate evolution. *Dev. Biol.* **173**, 382–395 (1996).
78. Holland, P. W. More genes in vertebrates? *J. Struct. Funct. Genomics* **3**, 75–84 (2003).
79. Huminięki, L. *et al.* Emergence, development and diversification of the TGF- β signalling pathway within the animal kingdom. *BMC Evol. Biol.* **9**, 28 (2009).
80. Hernandez-Sanchez, C., Mansilla, A., de Pablo, F. & Zardoya, R. Evolution of the insulin receptor family and receptor isoform expression in vertebrates. *Mol. Biol. Evol.* **25**, 1043–1053 (2008).
81. Bertrand, S. *et al.* Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. *Mol. Biol. Evol.* **21**, 1923–1937 (2004).
82. Conant, G. C. & Wolfe, K. H. Turning a hobby into a job: how duplicated genes find new functions. *Nature Rev. Genet.* **9**, 938–950 (2008).
83. Semon, M. & Wolfe, K. H. Consequences of genome duplication. *Curr. Opin. Genet. Dev.* **17**, 505–512 (2007).
84. Gould, S. J. & Lewontin, C. R. The spandrels of San Marco and the panglossian paradigm: a critique of the adaptationist programme. *Proc. Roy. Soc. Lond. B* **205**, 581–598 (1979).
85. Feild, T. S. & Arens, N. C. The ecophysiology of early angiosperms. *Plant Cell Environ.* **30**, 291–309 (2007).
86. Garcia-Fernandez, J. Amphioxus: a peaceful anchovy fillet to illuminate chordate evolution (I). *Int. J. Biol. Sci.* **2**, 30–31 (2006).
87. Schlueter, J. A. *et al.* Mining EST databases to resolve evolutionary events in major crop species. *Genome* **47**, 868–876 (2004).
88. Schranz, M. E. & Mitchell-Olds, T. Independent ancient polyploidy events in the sister families Brassicaceae and Cleomeaceae. *Plant Cell* **18**, 1152–1165 (2006).
89. Hall, J. C., Iltis, H. H. & Sytsma, K. J. Molecular phylogenetics of core Brassicales, placement of orphan genera *Emblingia*, *Forchhammeria*, *Tirania*, and character evolution. *Syst. Bot.* **29**, 654–669 (2004).
90. Mondragon-Palmino, M. & Theissen, G. Why are orchid flowers so diverse? Reduction of evolutionary constraints by paralogues of class B floral homeotic genes. *Ann. Bot. (Lond.)* 13 Jan 2009 (doi:10.1093/abo/mcn258).
91. Mondragon-Palmino, M. & Theissen, G. MADS about the evolution of orchid flowers. *Trends Plant Sci.* **13**, 51–59 (2008).
92. Valentine, J. W. Determinants of diversity in higher taxonomic categories. *Paleobiology* **6**, 444–450 (1980).
93. Conant, G. C. & Wolfe, K. H. Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Mol. Syst. Biol.* **3**, 129 (2007).
94. Hegarty, M. J. & Hiscock, S. J. Genomic clues to the evolutionary success of polyploid plants. *Curr. Biol.* **18**, R435–R444 (2008).
95. Marshall, C. S. Explaining the Cambrian ‘explosion’ of animals. *Annu. Rev. Earth Planet. Sci.* **34**, 355–384 (2006).
96. Glover, B. J. *Understanding Flowers and Flowering: An Integrated Approach* (Oxford Univ. Press, New York, 2007).
97. Regal, P. J. Ecology and evolution of flowering plant dominance. *Science* **196**, 622–629 (1977).
98. Theissen, G. & Melzer, R. Molecular mechanisms underlying origin and diversification of the angiosperm flower. *Ann. Bot. (Lond.)* **100**, 603–619 (2007).
99. Hu, S., Dilcher, D. L., Jarzen, D. M. & Winship Taylor, D. Early steps of angiosperm pollinator coevolution. *Proc. Natl Acad. Sci. USA* **105**, 240–245 (2008).
100. Poinar, G. O. Jr & Danforth, B. N. A fossil bee from Early Cretaceous Burmese amber. *Science* **314**, 614 (2006).
101. Kuraku, S., Meyer, A. & Kuratani, S. Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? *Mol. Biol. Evol.* **26**, 47–59 (2009).
102. Shimeld, S. M. & Holland, P. W. Vertebrate innovations. *Proc. Natl Acad. Sci. USA* **97**, 4449–4452 (2000).
103. Holland, N. D. & Chen, J. Origin and early evolution of the vertebrates: new insights from advances in molecular biology, anatomy, and palaeontology. *Bioessays* **23**, 142–151 (2001).
104. Wagner, A. Gene duplications, robustness and evolutionary innovations. *Bioessays* **30**, 367–373 (2008).
105. Khaner, O. Evolutionary innovations of the vertebrates. *Int. Zool.* **2**, 60–67 (2007).
106. Ohno, S. *Evolution by Gene Duplication* (Springer, New York, 1970).
107. Aburomia, R., Khaner, O. & Sidow, A. Functional evolution in the ancestral lineage of vertebrates or when genomic complexity was wagging its morphological tail. *J. Struct. Funct. Genomics* **3**, 45–52 (2003).
108. Crepet, W. L. Progress in understanding angiosperm history, success, and relationships: Darwin’s abominably ‘perplexing phenomenon’. *Proc. Natl Acad. Sci. USA* **97**, 12939–12941 (2000).
109. Stuessy, T. F. A transitional-combinatorial theory for the origin of angiosperms. *Taxon* **53**, 3–16 (2004).
110. Carquist, S. & Schneider, E. L. The tracheid-vessel element transition in angiosperms involves multiple independent features: cladistic consequences. *Am. J. Bot.* **89**, 185–195 (2002).
111. Muhammad, A. F. & Sattler, R. Vessel structure of *Gnetum* and the origin of angiosperms. *Am. J. Bot.* **69**, 1004–1021 (1982).
112. Valentine, J. W. & Walker, D. Diversity trends within a model of taxonomic hierarchy. *Physica D* **22**, 31–42 (1986).
113. Niklas, K. J. Computer models of early land plant evolution. *Annu. Rev. Earth Planet. Sci.* **32**, 47–66 (2004).
114. Niklas, K. J. Morphological evolution through complex domains of fitness. *Proc. Natl Acad. Sci. USA* **91**, 6772–6779 (1994).
115. Kauffman, S. A. *The Origins of Order* (Oxford Univ. Press, New York, 1993).

Acknowledgements

Y.V.d.P. acknowledges support from the IUAP P6/25 (BioMaGNet). A.M. thanks the Deutsche Forschungsgemeinschaft, University of Konstanz and the Institute for Advanced Study Berlin for support. S.M. is a fellow of the Fund for Scientific Research — Flanders (FWO). We thank two anonymous reviewers for valuable comments and suggestions and apologize to those whose work could not be cited because of space limitations.

FURTHER INFORMATION

Van de Peer laboratory homepage:

<http://bioinformatics.psb.ugent.be>

Meyer laboratory homepage:

<http://www.evolutionbiologie.uni-konstanz.de>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF