PHILOSOPHICAL TRANSACTIONS OF — THE ROYAL SOCIETY B | BIOLOGICAL SCIENCES

Tangled up in two: a burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution

Kevin Vanneste, Steven Maere and Yves Van de Peer

References	**This article cites 145 articles, 59 of which can be accessed free** http://rstb.royalsocietypublishing.org/content/369/1648/20130353.full.html#ref-list-1
🔓 **open access**	This article is free to access
Subject collections	Articles on similar topics can be found in the following collections computational biology (49 articles) evolution (693 articles) plant science (90 articles)
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. B* go to: **http://rstb.royalsocietypublishing.org/subscriptions**

Author for correspondence:
Yves Van de Peer
e-mail: yves.vandepeer@psb.vib-ugent.be

Tangled up in two: a burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution

Kevin Vanneste[1,2], Steven Maere[1,2] and Yves Van de Peer[1,2,3]

[1]Department of Plant Systems Biology, VIB, 9052 Ghent, Belgium
[2]Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium
[3]Genomics Research Institute (GRI), University of Pretoria, 0028 Pretoria, South Africa

Genome sequencing has demonstrated that besides frequent small-scale duplications, large-scale duplication events such as whole genome duplications (WGDs) are found on many branches of the evolutionary tree of life. Especially in the plant lineage, there is evidence for recurrent WGDs, and the ancestor of all angiosperms was in fact most likely a polyploid species. The number of WGDs found in sequenced plant genomes allows us to investigate questions about the roles of WGDs that were hitherto impossible to address. An intriguing observation is that many plant WGDs seem associated with periods of increased environmental stress and/or fluctuations, a trend that is evident for both present-day polyploids and palaeopolyploids formed around the Cretaceous–Palaeogene (K–Pg) extinction at 66 Ma. Here, we revisit the WGDs in plants that mark the K–Pg boundary, and discuss some specific examples of biological innovations and/or diversifications that may be linked to these WGDs. We review evidence for the processes that could have contributed to increased polyploid establishment at the K–Pg boundary, and discuss the implications on subsequent plant evolution in the Cenozoic.

1. Introduction

Flowering plants typically have large genome sizes and contain many genes, the majority of which evolved during the past 250–300 Myr through gene duplication [1]. A particularly striking feature of plant genomes, also explaining their large sizes, is the large number of whole genome duplications (WGDs) that have been uncovered [2–4]. It is now commonly accepted that one WGD occurred in the ancestor of all seed plants, and an extra one in the ancestor of all flowering plants, so that every extant angiosperm is in fact a palaeopolyploid containing the remnants of at least two WGDs [5]. Furthermore, a hexaploidy event pre-dates the origin of all core eudicots, which make up approximately 75% of extant angiosperm diversity [6–8], whereas traces of a WGD at the base of the monocots also suggest a WGD shared by most, if not all, monocots [9]. In addition, several more recent independent WGDs have been unveiled in many different plant lineages. As a result, the genomes of some extant plant species carry the remains of up to six successive genome duplications [10]. Here, we focus on the more 'recent' palaeopolyploidizations that occurred in the past 100 Myr, a large fraction of which seemingly took place around the Cretaceous–Palaeogene (K–Pg) extinction event at 66 Ma [11]. We have an in-depth look at this wave of WGDs associated with the K–Pg boundary, many of which pre-date lineage diversifications that resulted in some of the largest and arguably most successful present-day plant families, often characterized by particular biological innovations. Finally, we review processes that can explain these observations, and discuss how these palaeopolyploidizations could have influenced plant evolution in the Cenozoic.

2. A burst of genome duplications at the Cretaceous–Palaeogene boundary

In 2009, we described a tentative link between many of the known palaeopolyploidization events in plants and the K–Pg boundary, and speculated that WGD was linked to plant survival around that time [11]. Although many found this an interesting hypothesis [12], most remained sceptical, in particular because of the limited amount of data available at that time, and because dating ancient events that occurred tens of millions of years ago is often problematic. Only six complete genome sequences and a few transcriptome assemblies were available for analysis in 2009, limiting both the taxon sampling and possibility to implement proper primary fossil calibrations. Dating was performed using a penalized-likelihood inference method that incorporates an autocorrelated relaxed clock model, which assumes that branches that share a direct common ancestor also share similar evolutionary rates [13]. This assumption seems unlikely however, in the light of the sparse taxon sampling considered [14], and violation thereof may lead to inconsistent age estimates [15]. Calibrations were typically implemented as fixed secondary point calibrations, which may lead to illusionary precision of the time estimates [16].

Recent years have seen a huge increase in plant (whole genome) sequence data [17], in addition to the development of more powerful Bayesian methods for sequence divergence estimation [18–20], as well as more powerful high-performance computing systems that allow such intensive Bayesian algorithms to be run on a massive scale. We therefore recently revisited the hypothesized link between the K–Pg mass extinction and successful WGDs [21]. We used plant genome sequence information from a total of 41 species representing a broad coverage of the overall angiosperm phylogeny, incorporating 38 full genome sequences and three transcriptome assemblies, greatly improving taxon sampling with respect to the previous study [11]. In total, 20 independent WGDs could be dated compared with nine previously by dating all their identifiable homeologues created by the WGD event. For WGDs for which genome sequence information was available for several descendant species (e.g. WGDs preceding the divergence of Solanaceae, Fabaceae or Poaceae—see further), this WGD was dated independently for each species to assess their individual age estimates. Absolute age distributions were then constructed for each species WGD, for which a consensus WGD age estimate was obtained by taking the mode of its kernel density estimate, which is more flexible in comparison with traditional parametric distributions because it allows a better exploration of the true underlying shape of the distribution [22], whereas 90% confidence intervals were obtained through a bootstrapping procedure [23]. Dating itself was carried out with the BEAST package [20], using an uncorrelated relaxed clock model that assumes a lognormal distribution on evolutionary rates [19], and therefore should be better equipped to deal with rate shifts between different branches compared with autocorrelated relaxed clocks when taxon sampling is limited [24]. Proper calibration priors in Bayesian time estimation are of paramount importance as they can have a profound impact on the posterior age estimates [15,25–28]. Primary fossil calibrations were implemented as flexible lognormal calibration priors that represent the error associated with the age of the fossil in a more intuitive manner [27,29]. Fossils have a hard minimum bound corresponding to the earliest age to which the fossil can reliably be attributed to. The peak mass probability can be put at some distance after this earliest age to accommodate for the lag between first fossil occurrence and the actual divergence event the fossil is used to describe. Lastly, the lognormal distribution has an infinite extending, but small probability tail that can be used as a soft maximum bound to account for the uncertainty associated with choosing proper maximum bounds for fossil calibrations. More detailed information can be found in Vanneste et al. [21].

An updated overview of palaeopolyploidizations is summarized in figure 1 [21]. Although dating of such ancient events surely remains a challenging exercise, and WGD dates are subjected to change as more plant sequence data and powerful dating methods become available [12,30,31], many plant palaeopolyploidizations were again found to cluster at the K–Pg boundary [21], supporting our previous observations [11].

3. Implications of genome duplications associated with the Cretaceous–Palaeogene boundary

The increased long-term survival of WGDs around the K–Pg boundary appears indicative of enhanced polyploid plant establishment at that time, either because WGDs provided a selective advantage for polyploids compared with their diploid progenitors, or alternatively, because the cataclysmic events that took place 66 Ma were responsible for the production of an excess of polyploids (see further). However, whether cause or effect, many of these WGDs pre-date the radiation of some very large and successful plant families with particular biological innovations. Similar observations can be done in other parts of the tree of life, where WGDs are often found at branches leading to species-rich clades, such as more than 25 000 species of teleost fishes and more than 350 000 species of flowering plants [3,32]. On the other hand, one should be cautious not to over-interpret the importance of WGDs for species radiations. For instance, in vertebrates, it was suggested that the often-quoted correlation between the teleost fish WGD and increased post-WGD diversity and/or complexity does not hold when extinct basal lineages were considered [33]. Teleost fish evolution rather fits a more nuanced pattern of reduced extinction risk after WGD, resulting in a lag period between WGD and its effect on species diversity and/or complexity [34]. Additionally, it was recently demonstrated that an extended period of about 40–50 Myr passed between the salmonid-specific WGD and strong lineage diversification, suggesting the latter was probably mostly driven by climatic factors [35]. Below, we first examine a few examples of biological innovations (or better said, elaborations thereof [2]) that can reliably be traced back to WGDs located at the K–Pg boundary in plants, focusing on fleshy fruits in the Solanaceae and advanced nodulation characteristics in the papilionoids, before taking a deeper look at evidence whether or not these WGDs could have directly enhanced speciation.

(a) Biological novelty
(i) Fleshy fruits
The fleshy fruits observed in some plant lineages are an important biological innovation that serves to enhance seed
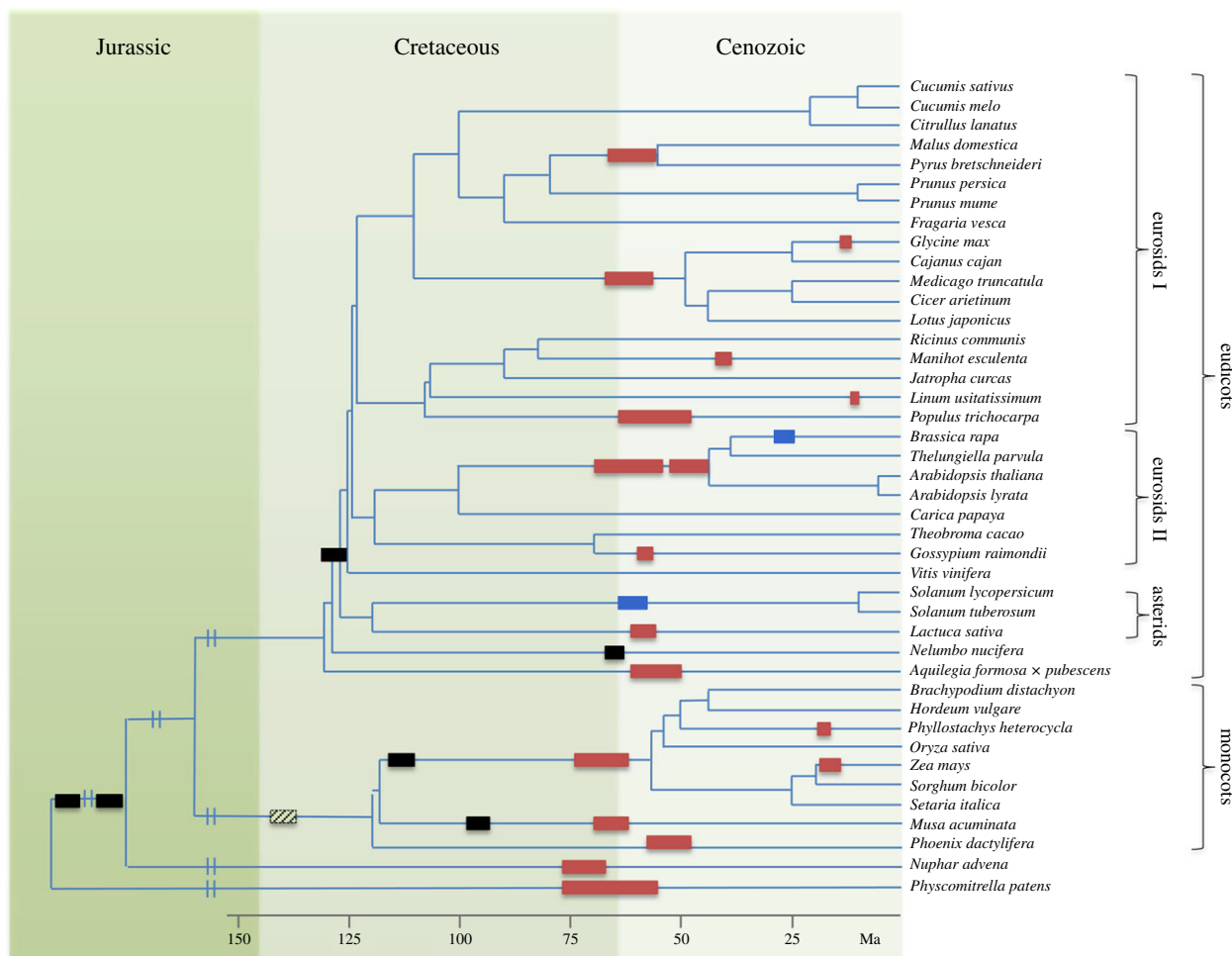
3

Figure 1. A wave of WGDs is associated with the K–Pg boundary approximately 66 Ma. The figure illustrates the tree topology for the green plants with all known WGDs indicated by bars. Red and blue bars represent 90% confidence intervals on dated tetraploidies and hexaploidies, respectively. Black bars represent WGD age estimates from literature [21]. A possible WGD at the base of the monocots is indicated by a dashed bar, because its exact phylogenetic placement remains unclear [9]. The WGD for *Populus trichocarpa* and the one shared by *M. domestica* and *P. bretschneideri* are corrected WGD age estimates based on fossil minimum boundaries and/or other dating studies [21]. Branch lengths are truncated after 150 Ma to allow a better overview. Figure adapted from Vanneste *et al.* [21].

distribution by attracting vertebrate frugivores for long-distance seed dispersal, and hence increases plant success [36]. Specialization of the fleshy fruit for particular (groups of) vertebrates may also enhance speciation [37]. Based on the recently published genome of tomato (*Solanum lycopersicum*), a genome triplication event in the Solanaceae shared with potato (*Solanum tuberosum*) was firmly established [38] and dated at the K–Pg boundary (figure 1). Many new gene family members with important fruit-specific functions were created through this WGD. Figure 2*a* illustrates several genes in the fruit-ripening control network that are paralogues with different physiological roles generated through the genome triplication. These include, for instance, the transcription factors and enzymes necessary for ethylene biosynthesis (*MADS1/RIN*, *CNR* and *ACS2/ACS6*), red light photoreceptors influencing fruit quality (*PHYB1/PHYB2*), and also some effector genes mediating lycopene biosynthesis (*PSY1/PSY2*) that control fruit pigmentation. Endogenous ethylene receptors (*ETR3/ETR4*) created by the eudicot-wide genome duplication also participate in this network. Similarly, fruit texture is controlled in part by over 50 genes that encode proteins involved in modification of cell wall structure and composition, and show differential expression during fruit development and ripening. Figure 2*b*, for instance, illustrates the expansion, through genome triplication and subsequent

tandem duplications, of a family of xyloglucan endotransglucosylase/hydrolases (*XTHs*) involved in determining fruit texture. Differential loss between tomato and potato of one of the triplicated members, *XTH10*, suggests that genetic specialization, and hence diversification between the different members of the Solanaceae, was facilitated by the triplication event [38]. It should however be noted that fleshy fruits exist in many different plant lineages, many of which are not marked by a specific polyploidy, emphasizing that the Solanaceae-shared WGD contributed several genes that were later incorporated into more elaborate fleshy fruit development, so that the latter represents an 'elaboration' rather than a true 'innovation' [2].

(ii) Rhizobial nodulation

A common feature of most papilionoid legumes is rhizobial nodulation, the formation of specialized organs called root nodules, which host nitrogen-fixing rhizobial symbionts. Nodulation is a biological innovation that allows growth on nitrogen-deprived soils, because plants receive fixed nitrogen from their symbionts, in return for a steady supply of carbon and energy sources [39]. Specialization for different rhizobial symbionts may also have aided papilionoid speciation [40]. Analysis of the genome sequence of *Medicago truncatula*
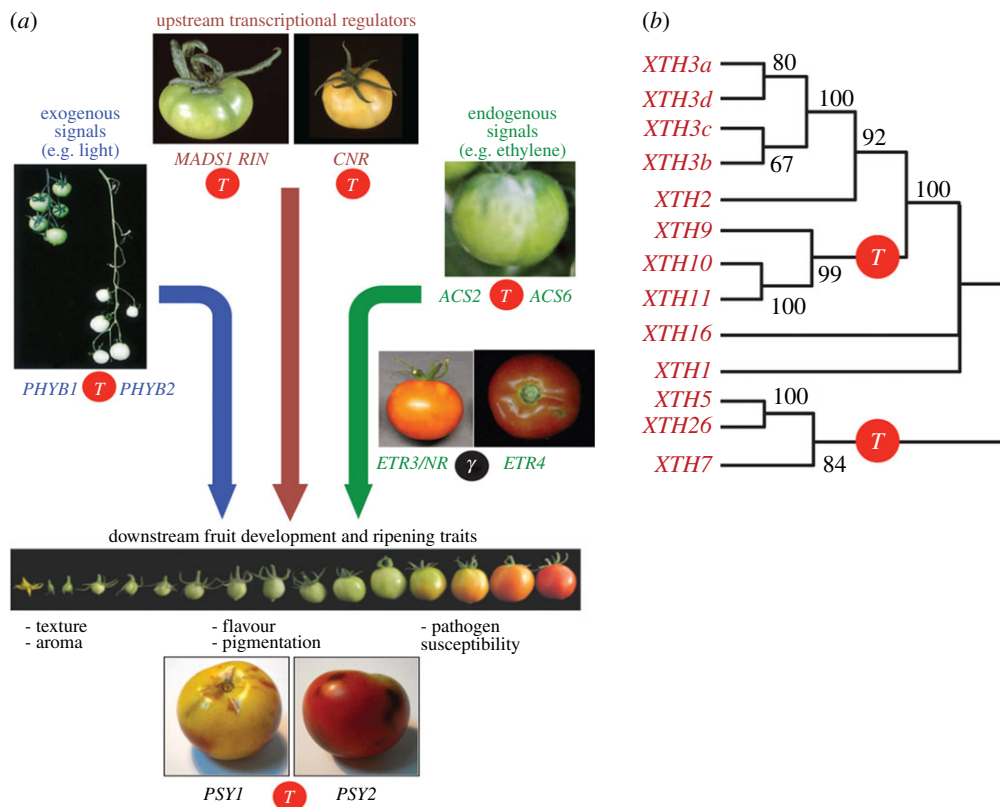
Figure 2. The Solanaceae-specific genome triplication contributed to the evolution of the tomato fruit. (a) Illustration of the fruit-ripening control network. The upstream transcriptional regulators MADS-RIN and CNR, in combination with the enzyme ACC synthase (ACS), control the production of the ripening hormone ethylene. Ethylene receptors (ETR) drive expression changes in several output genes, including phytoene synthase (PSY), which is the rate-limiting step in carotenoid biosynthesis. Light influences fruit pigmentation through an ethylene-independent pathway mediated by phytochromes (PHY). Several key component paralogous gene pairs (MADS1/RIN, PHYB1/PHYB2, ACS2/ACS6, PSY1/PSY2) were generated by the genome triplication (T, red circle), whereas ETR3/ETR4 was created by the core eudicot shared hexaploidy (γ, black circle). (b) Illustration of the expansion by both genome triplication (T, red circle) and tandem duplications of a family of xyloglucan endotransglucosylase/hydrolases (XTHs), which control fruit-ripening through modification of cell wall structure and composition. Figure adapted from Sato et al. [38].

confirmed that the papilionoid-shared WGD, also located at the K–Pg boundary (figure 1), has played an important role in the evolution and elaboration of rhizobial nodulation [41]. Nodulation is initiated when the plant signalling system comes into contact with specific bacterial Nod factors, which in papilionoids evolved a distinctly nodulation-specific function [42]. Analysis of the M. truncatula genome showed that both the Nod factor receptor NFP and transcription factor ERN1 have paralogues, LYR1 and ERN2, respectively, which originated through the papilionoid WGD. Figure 3 illustrates that both gene pairs show divergent expression patterns, reflecting functional specialization. NFP and ERN1 are expressed predominantly in the nodule and are known to be active in nodulation [43], whereas LYR1 and ERN2 are highly expressed during mycorrhizal colonization. This suggests that these nodulation-specific signalling components are derived from more ancient genes originally functional in mycorrhizal signalling that evolved new transcriptional functionality after the papilionoid WGD [41]. Additional support for this conclusion comes from the observation that the orthologue of NFP in a nodulating non-legume outgroup, Parasponia andersonii, functions both in nodulation and mycorrhizal signalling [44]. Interestingly, a nodulating legume outgroup that did not share the papilionoid WGD, Chamaecrista fasciculata, exhibits ancestral nodule characteristics in comparison with most nodulating papilionoids [45]. Parasponia diverged somewhere between 100 and 120 Ma

from the papilionoids [46], whereas Chamaecrista diverged approximately 60 Ma from the papilionoids [45]. Independent of whether their last common ancestor could already perform nodulation or whether this trait evolved independently in both lineages, this would suggest that the ability for advanced nodulation characteristics was not able to evolve for about 40–60 Ma, whereas it did so very rapidly after the papilionoid WGD [45]. This emphasizes that although the papilionoid WGD was not an absolute prerequisite for the evolution of nitrogen-fixing nodulation, it most likely facilitated the development of several elaborate papilionoid nodule forms.

To assess the contribution of the papilionoid WGD to M. truncatula nodulation in more detail, Young et al. [41] also investigated the expression of 618 homeologous gene pairs from six different organs based on RNA-seq data for one or both homeologues, to determine the number of genes showing organ-enhanced expression (defined as having expression in a single organ that is at least twice the level in any other). A large fraction of homeologues demonstrated organ-specific enhanced expression. Among homeologous gene pairs with nodule-enhanced expression, a single paralogue was nodule-enhanced in 43 of 51 gene pairs, with the other eight gene pairs showing nodule-enhanced expression for both gene pairs. Of 142 transcription factors derived from the papilionoid WGD for which RNA-seq data was available, 11 showed such enhanced nodule expression. These results indicate that many homeologous genes, in particular signalling
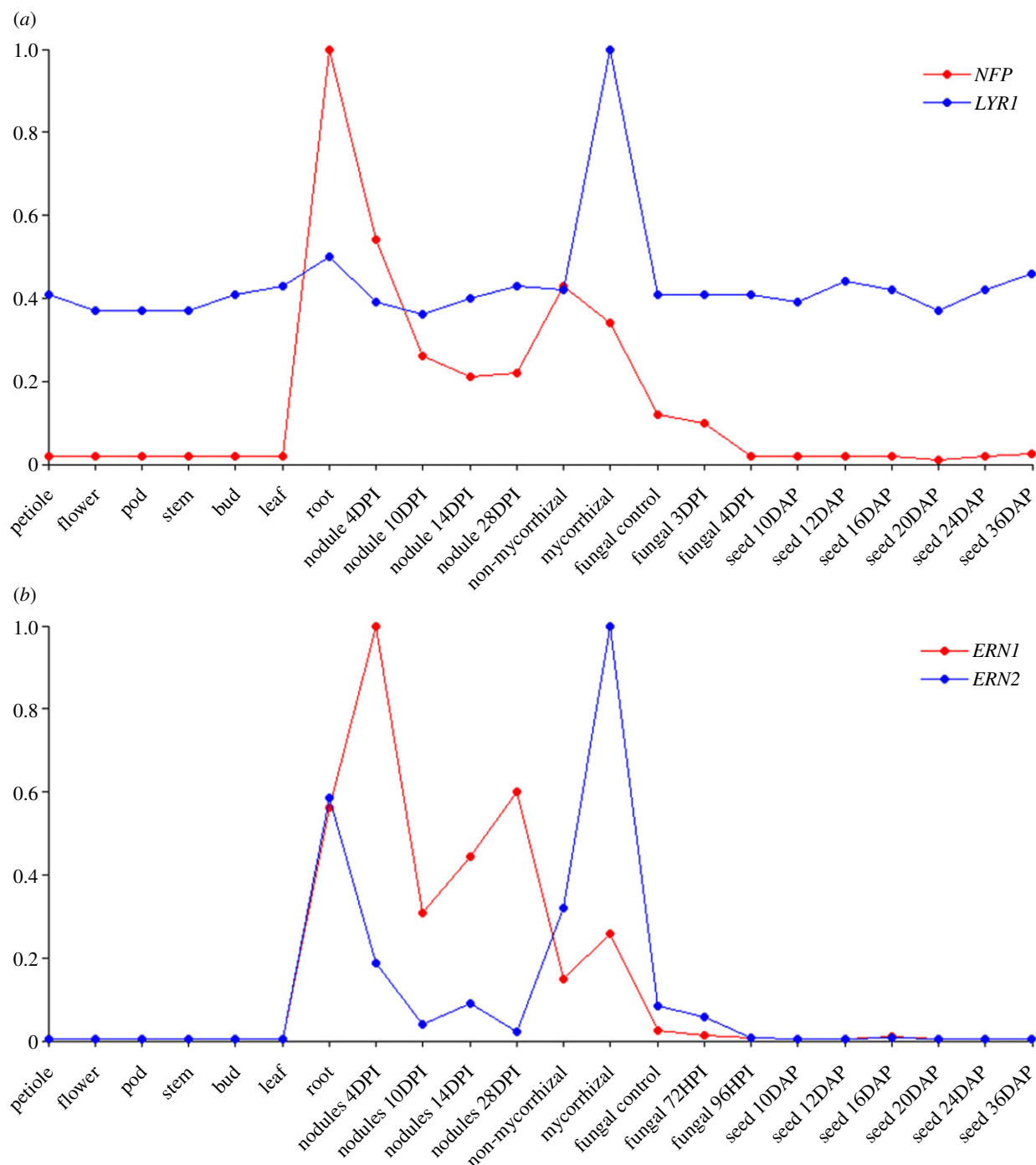
Figure 3. The papilionoid genome duplication contributed to the evolution of nodulation. Paralogues created by WGD, (*a*) *NFP* and *LYR1*, and (*b*) *ERN1* and *ERN2*, display contrasting expression patterns, suggesting functional specialization. *NFP* and *ERN1* are expressed predominantly in the nodule, whereas *LYR1* and *ERN2* are highly expressed during mycorrhizal colonization. The average transcript levels of three replicates are shown, scaled by dividing each data point by the maximum mean transcript level across all experiments. HPI, hours past inoculation; DPI, days past inoculation; DAP, days after pollination. Figure adapted from Young *et al.* [41].

components and regulators, were retained after the papilionoid WGD and gained specialized roles in nodulation afterwards. However, some other nodule-related genes were found to derive from the core-eudicot-specific hexaploidy. This confirms a more complex model wherein the capacity for primitive interaction with new symbionts evolved quite early, derived from the existing mycorrhizal machinery, explaining the evolution of nodulation in multiple plant lineages [39,47], after which the papilionoid WGD allowed the creation of additional genes that were incorporated into the development of more advanced nodulation characteristics [41]. A recent integrated comparative genomic approach based on the sequenced genomes of four papilionoid species (*M. truncatula*, *Lotus japonicus*, *Glycine max* and *Cajanus cajan*) supports this by demonstrating that many of the approximately 25% of WGD-derived duplicate pairs that have been retained, show high levels of expression divergence and function in different processes required for successful nodulation [48].

(b) Speciation

The previous examples of biological innovations originating through the retention of WGD duplicates suggest that WGDs, through assisting biological innovations and diversifications, might also facilitate speciation. For instance, as stated previously, specialization for interactions between

particular vertebrate frugivores for seed dispersal in fleshy fruits or with specific rhizobial symbionts in nodulation, might aid speciation. However, the question remains whether WGD itself can also actively promote speciation. Some of the WGDs associated with the K–Pg boundary (figure 1) pre-date extremely successful plant lineages characterized by species radiations following the WGD event. These include the Brassicaceae (approx. 3700 species), Poaceae (approx. 10 000 species), Asteraceae (approx. 23 600 species), Solanaceae (approx. 2460 species) and Fabaceae (approx. 19 500 species). Many of these, however, have a species-poor sister group that shared the WGD event, which led to the development of the WGD-radiation lag time model that emphasizes that the success of these plant families should be viewed in the light of their specific evolutionary routes taken [49]. Even the limited set of species in figure 1 demonstrates that many present-day plant families, such as the Cucurbitaceae, represented by *Cucumis melo*, *Cucumis sativus* and *Citrullus lanatus*, did not undergo any WGD in the past approximately 100 Myr. Using the number of species as a simple, albeit admittedly crude, measure for success, this family of about 950–980 species can also be considered fairly successful [50]. Alternatively, some plant families with a palaeopolyploid history, such as the Nymphaeaceae, have arguably not been very successful in terms of species radiation, counting only around 70 species [51]. Such observations emphasize the importance of ecological opportunity for realizing plant evolutionary potential, irrespective of polyploidization [2,49,52,53].

Nevertheless, the success of many plant families that have undergone a WGD suggests that their strong diversification may be ascribed, at least partly, to their polyploid ancestry. In an attempt to gauge the effect of WGD on speciation, Soltis *et al.* [3] tested whether such post-WGD clades displayed higher diversification rates, while accounting for the confounding effects of extinction. Although the results were considered preliminary, owing to the lack of reliable genomic data for palaeopolyploidy in combination with insufficient taxon sampling to place WGDs confidently on plant family phylogenies, a highly statistically significant relationship between diversification and WGD was found for four of the five aforementioned successful plant families. The fifth plant family, the Asteraceae, was not considered, and a statistical relationship hence remains untested. It should however be noted that the latter constitutes the single largest present-day angiosperm family [54].

The molecular mechanisms that might promote speciation after WGD are still not very well understood. One often-quoted mechanism is reciprocal gene loss (RGL), the genetic isolation of separated populations through loss of different gene copies that lead to incompatibilities when the populations encounter each other again [55,56]. Through WGD, a very large pool of loci becomes available simultaneously for divergent resolution between subpopulations, which could quickly result in reproductive isolation if essential genes are involved. Scannell *et al.* [57] demonstrated that the pattern of duplicate gene pair loss differs at 20% of all loci between three different yeast species that shared a WGD. Similarly, about 8% of ancestral *Tetraodon* and zebrafish loci were subjected to RGL after the teleost fish WGD [58]. For plants, the situation is less clear. Schnable *et al.* [59] separated the two subgenomes of modern grasses derived from the WGD shared by the Poaceae. In contrast to the aforementioned studies in yeast and teleost fishes, strong evidence of RGL

between homeologues of the different subgenomes was lacking, suggesting post-WGD RGL was unlikely to be a driving force in the radiation of the grasses, although systematic studies about RGL in plants are still missing.

However, genes do not necessarily need to get lost or silenced, as other neutral scenarios after gene duplication might also promote speciation. Many genes perform multiple functions through differential expression at different developmental stages and/or tissues. Duplication of such genes often leads to subfunctionalization, the division of the subfunctions over the two daughter copies [55,60]. Alternatively, genes can have trace activity for a second function whose optimization is constrained by adaptive conflicts with the primary function, which can be resolved by optimizing the functions separately in different paralogues after duplication [61]. Reproductive isolation of such a population, for instance driven by geological phenomena that lead to geographical barriers, could lead to orthologues of the two isolated populations acquiring different subfunctions. Although F_1 hybrids in contact zones from the two populations would develop correctly because each (sub)function is performed by one of the genes from each population, one-eighth of the F_2 zygotes will lack one of the (sub)functions, which could be lethal if such functions are essential [62,63]. As for RGL, this effect would be exacerbated in the case of WGD, which generates a much larger number of duplicate loci that can be divergently subfunctionalized [2]. Lineage-specific subfunctionalization could therefore in theory accelerate speciation, but this remains untested.

4. Both neutral and adaptive processes most likely contribute towards enhanced polyploid establishment under stressful conditions

Above, we discussed new evidence that seems to provide further support for the association between plant palaeo-polyploidizations and the K–Pg boundary, some of which can be linked to particularly successful biological innovations and increased diversification rates. The K–Pg boundary is especially known for its associated extinction event, which constitutes the last of the five major mass extinctions in the Phanerozoic eon [64]. This cataclysmic event most likely resulted from the combination of several factors such as increased volcanism, greenhouse warming, and in particular the bolide impact near Chicxulub (Mexico) [65], resulting in a challenging unstable environment impairing the survival of most living organisms [66]. The question remains, at a time when an estimated 75% of all species went extinct [67], why did many of the plant species we are all so familiar with probably undergo a WGD? Similar observations have been made for present-day polyploids, which are often encountered in unstable and stressful environments [68]. For instance, there is an overabundance of recently formed polyploids in the Arctic [69]. Below, we discuss two, not mutually exclusive, processes that could help explain this pattern and the implications thereof for plant evolution.

(a) The adaptive scenario

The adaptive scenario explaining polyploid success has been explored extensively in the past decade [2,3,70–73], and will therefore only be covered concisely here. This scenario is mostly based on a characteristic often displayed by newly

formed polyploids, namely transgressive segregation, i.e. the formation of more extreme phenotypes in the resulting hybrid populations compared with their diploid parents [70]. This becomes more pronounced as the two parental genomes contributing to the polyploid become more diverged, especially so in allopolyploids that result from the merger of two different species, which may display strong hybrid vigour (heterosis) by virtue of possessing novel allelic combinations not found in either parent [74]. However, the exact molecular mechanisms behind hybrid vigour are still largely unknown [75], although it has been suggested recently that cells might distinguish between parental alleles based on their relative protein and mRNA stability, which therefore conserves energy otherwise required for removal of such unstable products that can be used to promote growth and expression of new favourable traits [76].

Irrespective of the exact molecular mechanisms, genomic instability and gene expression changes soon after polyploid formation may result in increased phenotypic variability of the polyploids with respect to their diploid progenitors [2]. Genomic instability refers to the extensive structural changes of the chromosomal DNA that typically take place in the first few generations after polyploidization, such as fusions, fissions, duplications, inversions, translocations and eliminations [77], often coupled to mitotic and meiotic abnormalities [78,79]. Gene expression typically changes markedly [80], in conjunction with widespread epigenetic repatterning [81], in the first few generations after polyploidization. These structural and expression changes have collectively been described as genomic shock, and in the case of allopolyploids seem to be attributable to both the hybridization process [82] and the genome doubling itself, with the latter possibly having a calming effect [83]. Although these extensive changes often result in decreased polyploid fitness and increased offspring sterility, in the light of increased phenotypic variability, they can also confer plasticity to the polyploid genome to allow quick adaptation to new environments and changing conditions [70,71,73,84,85].

Other potential adaptive advantages of newly formed polyploids include the masking of deleterious recessive alleles leading to increased genetic redundancy [86], network redundancy on a larger scale [87] and possibly even an increased capacity for phenotypic plasticity itself [88,89]. Polyploids also often exhibit traits that promote their establishment through mitigating the minority cytotype disadvantage, which is a strong negative frequency-dependent selection on the polyploid through a large proportion of ineffective matings with the diploid progenitor majority cytotype [90]. Such traits include the loss of self-incompatibility, which enables selfing, and the gain of apomixis, which enables asexual reproduction. Polyploidization is also sometimes associated with a shift from annual to perennial habit, which opens up a longer time window for successful mating. Lastly, their fast morphological and/or physiological differentiation can enhance the number of successful matings through sympatric niche separation from the diploid progenitor population [73,91,92].

(b) The neutral scenario

A series of recent findings point to the possibility of a more neutral scenario to explain the apparent association between palaeopolyploidizations and the K–Pg boundary [21]. It has been acknowledged for a long time that the formation of

unreduced gametes is the main mode of polyploid formation in plants, but the low estimates of unreduced gamete production in natural populations typically seemed too restrictive for the establishment of polyploids [93,94]. Although the chance of two unreduced $2n$ gametes meeting is very low, tetraploid occurrence is most likely facilitated by a triploid bridge, the creation of an intermediate triploid stage through the combination of an unreduced $2n$ and reduced n gamete [95]. Such triploids often display large fertility and fitness defects; however, they also produce enhanced levels of unreduced $3n$ gametes that can form tetraploids through backcrosses with reduced n gametes from the diploid progenitor population, and hence alleviate the minority cytotype disadvantage [96,97]. Accordingly, a recent general gametic modelling approach for diploid–polyploid systems that predicts equilibrium ploidy frequencies based on empirical estimates of unreduced gamete formation, demonstrated that these low levels can be adequate to explain a drift towards higher ploidy [98].

Another well-documented observation is that levels of unreduced gamete formation can be increased by external stimuli such as stress and a fluctuating environment [94,99–104]. Temperature in particular has a pronounced effect on unreduced gamete formation. Increasing temperatures to extreme levels in *Rosa* species resulted in more unreduced gametes being produced through alterations in spindle formation during meiosis II [105]. Similarly, inducing cold stress increased unreduced gamete formation in *A. thaliana* through alterations in post-meiotic cell plate formation and cell wall establishment [106]. Although hybridization itself typically also increases the levels of unreduced gamete formation in plants [107], temperature levels can potentially also enhance this hybrid trait, as witnessed in some *Brassica* interspecific hybrids after cold treatment [97]. Moreover, it became clear recently that the effect of the environment on unreduced gamete formation is most likely not limited to present-day plants. Increased levels of fossil unreduced pollen were observed in the now extinct conifer family Cheirolepidiaceae at the Triassic–Jurassic transition, which corresponds to the fourth of the five major extinction events [108]. Abnormal gymnosperm pollen [109] and lycophyte spores [110] have also been reported during the Permian–Triassic transition, corresponding to the third of the five major extinction events.

Increased unreduced gamete production during times of environmental stress and/or fluctuation could thus be an important factor in explaining the apparent clustering of palaeopolyploidizations at the K–Pg boundary [21]. It could also explain why many present-day polyploids often are more abundant in stressful environments, such as the Arctic [69] or habitats created by anthropogenic disturbance [111]. For both the K–Pg boundary and present-day examples, the association between increased polyploid establishment and environmental stress and/or fluctuation would not require any explicit adaptive advantage, but could be explained by a neutral mechanism [99] such as increased unreduced gamete formation. This is in agreement with modelling approaches that predict increased replacement of diploids by polyploids under a changing environment, without assuming any *a priori* adaptive advantage of the polyploids [112]. The effect of increased unreduced gamete production during environmental stress and/or fluctuation is even expected to be intensified through higher background extinction levels of the diploid populations [34], increasing the overall relative frequency of unreduced gametes to the total gamete pool,
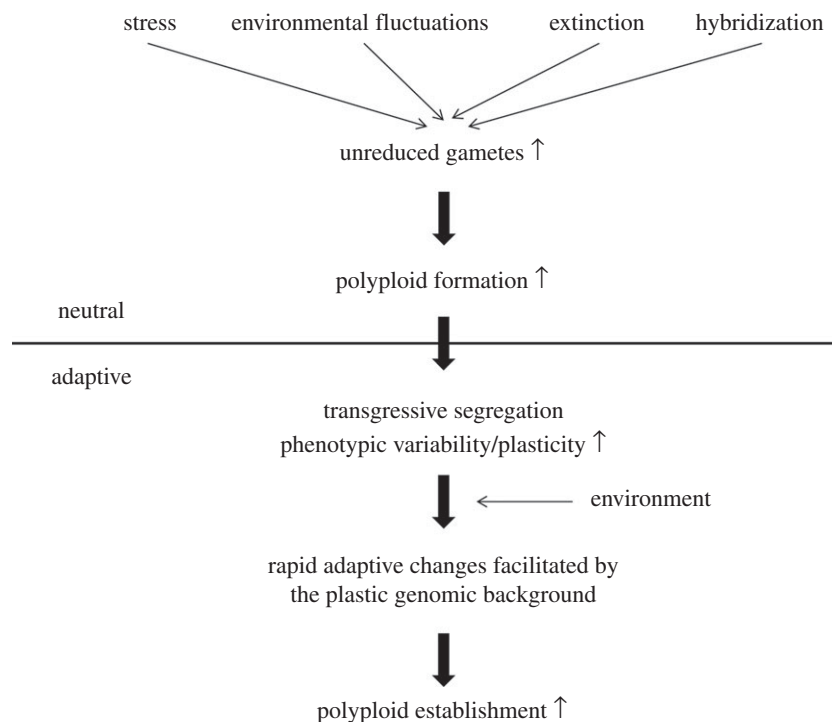
Figure 4. Both neutral and adaptive processes probably contribute to enhanced polyploid establishment under environmental stress and/or fluctuations. The latter likely increase the formation of unreduced gametes, whereas other processes such as hybridization and extinction of the background diploid population can also contribute to an overall increase of unreduced gametes in the total gamete pool. This is expected to lead to more polyploids being formed even in the absence of any active adaptive advantage. Transgressive segregation and genomic instability of polyploids on the other hand may lead to heterotic phenotypes, increased phenotypic variability and plasticity that, if beneficial under the changing environment, can be rapidly selected for, which is expected to lead to more polyploids being established even in the absence of increased polyploid formation. Note that irrespective of which scenario carries more weight, the environment plays an important role in polyploid establishment.

which would enhance the chance of successful unreduced gamete matings.

Accumulating evidence for a more prominent role of the neutral scenario does not however preclude a role for the adaptive one. Figure 4 summarizes an intertwined situation wherein environmental stress and/or fluctuation drive polyploid formation through increased unreduced gamete production, after which adaptive processes act to ensure polyploid establishment. Dependent upon specific circumstances, either the neutral or adaptive component could carry more weight. The apparent association of palaeo-polyploidizations with the K–Pg boundary [21], and present-day polyploids with stressful habitats [69,111], in combination with evidence that unreduced gamete formation is a major route towards polyploidization [98] that may be intensified through environmental stress and/or fluctuations as witnessed at several large-scale extinction events [108], hints at a strong role for the neutral component. There are however many observations that also argue in favour of the adaptive component [73]. Although one has to remain cautious with generalizations about the distribution and prevalence of recent polyploids, because many exceptions can be found [113], some trends are apparent. For instance, recent polyploids appear to have larger habitat distributions, suggesting they can tolerate a wider range of ecological conditions [114–116]. Most strikingly, they are less likely to be endangered and more likely to be invasive on a worldwide scale compared with diploids [117]. Such observations would be difficult to explain purely through neutral mechanisms.

The genetic component of unreduced gamete production merits some more attention. Traditional breeding studies established that diploid gamete production is a highly heritable trait that can be enhanced in as few as two to three cycles of recurrent selection in species such as alfalfa [118] and red clover [119]. In *Arabidopsis*, a surprisingly strong tolerance of gametes to both trisomy and several other complex karyotypes exists [120], whereas several genetic players that can influence unreduced gamete production through their effect on the orientation of the spindle apparatus in male meiosis have recently been identified [121], such as *AFH14* [122], *JAS* [123] and *AtPS1* [124]. Stress-induced altered functionality of these genetic components may explain the effect of the environment on unreduced gamete production [104]. These observations open up the possibility that polyploidization might even constitute an inducible evolutionary mechanism by which plants cope with ecological disasters, much akin to the stress-inducible mutator systems such as the SOS response in bacteria [125]. The latter is a transient response to stress and changing environments by means of a set of 'evolution genes' that decrease replication fidelity and increase mutation rates to generate genetic diversity upon which natural selection can act [126,127]. Such evolution genes are thought to undergo biological evolution themselves through indirect selection, and their presence in higher organisms has been hypothesized [128]. Because all extant angiosperms shared at least two rounds of WGD [5], with an extra shared WGD at the base of the core eudicots [6] and possibly also the monocots [9], recurring WGD events [2–4] could have maintained residual heritable genetic variation in diploid plants for the ability to produce unreduced gametes and form polyploids in times of ecological upheaval. Despite a genetic component, this does not need to be necessarily under the direct control of any adaptive programme, as it could just as well primarily be an 'evolutionary spandrel' that received secondary functionality [129]. In any

5. Enhanced establishment of polyploids at the Cretaceous–Palaeogene boundary may have paved the way for angiosperm success in the Cenozoic

The neutral and adaptive processes described above offer a framework for the apparent clustering of WGDs at the K–Pg boundary, but fail to explain their long-term success in terms of speciation and biological novelty. For all examples we considered, it was apparent that the duplication of the whole genome provided an increase in raw genetic material on which evolution could work. In accordance with Ohno's classical models [130,131], the newly created gene copies could undergo neofunctionalization (the creation of a new function), subfunctionalization (the division of an ancestral function or functions over the daughter copies), or be kept for dosage amplification (the production of more of a beneficial gene product) or any combination thereof as explained by more complex population genetic models [132]. Although the fate of most duplicated genes is in fact loss through pseudogenization [1], WGDs provide a massive number of contemporarily created gene duplicates, of which only a small fraction seems to have contributed to some major biological innovations and/or elaborations.

It has become increasingly clear that rather than just the functional divergence of the coding regions and/or regulatory sequences of individual genes, the rewiring of the regulatory network containing these individual components following WGD is of major importance [133,134]. A body of literature exists demonstrating that regulatory and developmental genes in particular are retained in excess after WGDs. This is most likely due to dosage-balance constraints, i.e. selection against loss of individual components of completely duplicated macromolecular complexes and/or pathways, because this would disrupt their overall stoichiometry [135–139]. Retention of balance-sensitive duplicates thus does not provide an immediate evolutionary advantage, but results from the fact that their loss would lead to an immediate disadvantage. In this respect, the retained regulators may be considered an evolutionary spandrel [129,135], which might later on have facilitated the evolutionary innovations and/or diversifications observed in many of these post-WGD lineages [2,3,140]. Selection to maintain dosage balance eventually relaxes over time allowing functional divergence in the context of the environment [138,141], so that part of the duplicated network can be rewired to execute novel functions [133]. However, the underlying mechanisms are currently unclear. Gene duplication has been shown to contribute to innovations even after prolonged periods between the original duplication event and the origin of novelty [142], suggesting that individual components of these duplicated networks can undergo neo- and subfunctionalization in accordance with Ohno's classical models [131,132] even long after the duplication event itself. Some of these processes could have caused network-rewiring events that could help explain the vast post-WGD success

observed in some of the plant families that experienced a WGD at the K–Pg boundary.

There are many examples that support the role of network rewiring over time. The ability for anaerobic fermentation in yeast has been associated with global rewiring of its transcriptional network after genome duplication, involving changes in the promoter regions of several genes such as the loss of specific regulatory motifs [143,144]. Similarly, the abundance of teleost fish pigmentation synthesis pathways has been attributed to the teleost WGD through rewiring in combination with subfunctionalization of existing pathways [145]. In plants, the *gamma* hexaploidy at the base of the core eudicots resulted in expansion of MADS-box gene families, key regulators of reproductive development, which through rewiring of their interaction network in combination with neo- and subfunctionalization, acquired roles in several major plant developmental processes [8,146].

6. Conclusion

Advances in plant genomics, molecular sequence divergence estimation and high-performance computational solutions allow us to address questions about the role of genome duplication that were previously impossible to investigate. It should be emphasized that the fate of most newly formed polyploids appears an evolutionary dead end through outcompetition by their diploid specialized progenitors [147–149] because of a whole range of associated negative effects such as minority cytotype exclusion [90], severe meiotic and mitotic abnormalities [150] and ploidy-associated genomic instability [79]. Nevertheless, it appears that there exists a strong link between environmental stress and/or fluctuation and genome duplication, as currently supported by both present-day polyploids and palaeopolyploids at the K–Pg boundary. Could unreduced plant gamete production have increased polyploid formation at the K–Pg boundary? Alternatively, can the apparent prevalence of polyploids at the K–Pg boundary be explained by their increased adaptability? Or do we observe the signature of another mechanism and/or pattern that currently remains elusive, perhaps because both dating of such ancient events and making generalizations about current polyploids remain particularly problematic? In any case, this polyploid heritage may afterwards have fuelled evolution of biological innovations and speciation in the context of newly encountered conditions during the Cenozoic, through extensive network rewiring and functional diversification of regulatory and developmental genes that were originally guarded against loss through mechanistic dosage-balance constraints. Polyploids in some sense thus seem reminiscent of the 'hopeful monsters' advocated by Goldschmidt [151] (M. Freeling 2009, personal communication), at least at the genomic level, whereas their full potential at the phenotypic level can only be realized given time and the right conditions [52]. It thus appears that the role of the environment in both polyploid establishment and their evolutionary success constitutes an important aspect that merits further investigation.

10

rstb.royalsocietypublishing.org Phil. Trans. R. Soc. B 369: 20130353

References

1. Lynch M, Conery JS. 2000 The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155. (doi:10.1126/science.290.5494.1151)

2. Van de Peer Y, Maere S, Meyer A. 2009 The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* **10**, 725–732. (doi:10.1038/nrg2600)

3. Soltis DE *et al.* 2009 Polyploidy and angiosperm diversification. *Am. J. Bot.* **96**, 336–348. (doi:10.3732/ajb.0800079)

4. Cui L *et al.* 2006 Widespread genome duplications throughout the history of flowering plants. *Genome Res.* **16**, 738–749. (doi:10.1101/gr.4825606)

5. Jiao Y *et al.* 2011 Ancestral polyploidy in seed plants and angiosperms. *Nature* **473**, 97–100. (doi:10.1038/nature09916)

6. Jaillon O *et al.* 2007 The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**, 463–467. (doi:10.1038/nature06148)

7. Jiao Y *et al.* 2012 A genome triplication associated with early diversification of the core eudicots. *Genome Biol.* **13**, R3. (doi:10.1186/gb-2012-13-1-r3)

8. Vekemans D, Proost S, Vanneste K, Coenen H, Viaene T, Ruelens P, Maere S, Van de Peer Y, Geuten K. 2012 Gamma paleohexaploidy in the stem lineage of core eudicots: significance for MADS-box gene and species diversification. *Mol. Biol. Evol.* **29**, 3793–3806. (doi:10.1093/molbev/mss183)

9. Paterson AH, Bowers JE, Chapman BA. 2004 Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl Acad. Sci. USA* **101**, 9903–9908. (doi:10.1073/pnas.0307901101)

10. Van de Peer Y, Fawcett JA, Proost S, Sterck L, Vandepoele K. 2009 The flowering world: a tale of duplications. *Trends Plant Sci.* **14**, 680–688. (doi:10.1016/j.tplants.2009.09.001)

11. Fawcett JA, Maere S, Van de Peer Y. 2009 Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. *Proc. Natl Acad. Sci. USA* **106**, 5737–5742. (doi:10.1073/pnas.0900906106)

12. Soltis DE, Burleigh JG. 2009 Surviving the K–T mass extinction: new perspectives of polyploidization in angiosperms. *Proc. Natl Acad. Sci. USA* **106**, 5455–5456. (doi:10.1073/pnas.0901994106)

13. Sanderson MJ. 2002 Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* **19**, 101–109. (doi:10.1093/oxfordjournals.molbev.a003974)

14. Ho SY. 2009 An examination of phylogenetic models of substitution rate variation among lineages. *Biol. Lett.* **5**, 421–424. (doi:10.1098/rsbl.2008.0729)

15. Mulcahy DG, Noonan BP, Moss T, Townsend TM, Reeder TW, Sites Jr JW, Wiens JJ. 2012 Estimating divergence dates and evaluating dating methods using phylogenomic and mitochondrial data in squamate reptiles. *Mol. Phylogenet. Evol.* **65**, 974–991. (doi:10.1016/j.ympev.2012.08.018)

16. Ho SY, Phillips MJ. 2009 Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* **58**, 367–380. (doi:10.1093/sysbio/syp035)

17. Fritz MHY, Leinonen R, Cochrane G, Birney E. 2011 Efficient storage of high throughput DNA sequencing data using reference-based compression. *Genome Res.* **21**, 734–740. (doi:10.1101/gr.114819.110)

18. Anisimova M, Liberles DA, Philippe H, Provan J, Pupko T, von Haeseler A. 2013 State-of-the-art methodologies dictate new standards for phylogenetic analysis. *BMC Evol. Biol.* **13**, 161. (doi:10.1186/1471-2148-13-161)

19. Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006 Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88. (doi:10.1371/journal.pbio.0040088)

20. Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012 Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973. (doi:10.1093/molbev/mss075)

21. Vanneste K, Baele G, Maere S, Van de Peer Y. In press. Analysis of 41 plant genomes supports a wave of successful genome duplications at the Cretaceous–Tertiary boundary. *Genome Res.* (doi:10.1101/gr.168997.113)

22. Botev ZI, Grotowski JF, Kroese DP. 2010 Kernel density estimation via diffusion. *Ann. Stat.* **38**, 2916–2957. (doi:10.1214/10-AOS799)

23. Hall P, Kang KH. 2001 Bootstrapping nonparametric density estimators with empirically chosen bandwidths. *Ann. Stat.* **29**, 1443–1468. (doi:10.1214/aos/1013203461)

24. Smith SA, Beaulieu JM, Donoghue MJ. 2010 An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proc. Natl Acad. Sci. USA* **107**, 5897–5902. (doi:10.1073/pnas.1001225107)

25. Inoue J, Donoghue PC, Yang Z. 2010 The impact of the representation of fossil calibrations on Bayesian estimation of species divergence times. *Syst. Biol.* **59**, 74–89. (doi:10.1093/sysbio/syp078)

26. Hug LA, Roger AJ. 2007 The impact of fossils and taxon sampling on ancient molecular dating analyses. *Mol. Biol. Evol.* **24**, 1889–1897. (doi:10.1093/molbev/msm115)

27. Magallon S, Hilu KW, Quandt D. 2013 Land plant evolutionary timeline: gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *Am. J. Bot.* **100**, 556–573. (doi:10.3732/Ajb.1200416)

28. Yang Z, Rannala B. 2006 Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* **23**, 212–226. (doi:10.1093/molbev/msj024)

29. Forest F. 2009 Calibrating the tree of life: fossils, molecules and evolutionary timescales. *Ann. Bot.* **104**, 789–794. (doi:10.1093/aob/mcp192)

30. Hedges SB, Kumar S. 2004 Precision of molecular time estimates. *Trends Genet.* **20**, 242–247. (doi:10.1016/j.tig.2004.03.004)

31. Graur D, Martin W. 2004 Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends Genet.* **20**, 80–86. (doi:10.1016/j.tig.2003.12.003)

32. Meyer A, Van de Peer Y. 2005 From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *Bioessays* **27**, 937–945. (doi:10.1002/bies.20293)

33. Donoghue PC, Purnell MA. 2005 Genome duplication, extinction and vertebrate evolution. *Trends Ecol. Evol.* **20**, 312–319. (doi:10.1016/j.tree.2005.04.008)

34. Crow KD, Wagner GP. 2006 What is the role of genome duplication in the evolution of complexity and diversity? *Mol. Biol. Evol.* **23**, 887–892. (doi:10.1093/molbev/msj083)

35. Macqueen DJ, Johnston IA. 2014 A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proc. R. Soc. B* **281**, 20132881. (doi:10.1098/rspb.2013.2881)

36. Howe HF, Smallwood J. 1982 Ecology of seed dispersal. *Annu. Rev. Ecol. Syst.* **13**, 201–228. (doi:10.1146/annurev.es.13.110182.001221)

37. Givnish TJ. 2010 Ecology of plant speciation. *Taxon* **59**, 1326–1366.

38. Sato S *et al.* 2012 The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**, 635–641. (doi:10.1038/nature11119)

39. Kiers ET, Rousseau RA, West SA, Denison RF. 2003 Host sanctions and the legume–rhizobium mutualism. *Nature* **425**, 78–81. (doi:10.1038/nature01931)

40. Tian CF *et al.* 2012 Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. *Proc. Natl Acad. Sci. USA* **109**, 8629–8634. (doi:10.1073/pnas.1120436109)

41. Young ND et al. 2011 The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480, 520–524. (doi:10.1038/nature10625)

42. Oldroyd GED, Downie JM. 2008 Coordinating nodule morphogenesis with rhizobial infection in legumes. Annu. Rev. Plant Biol. 59, 519–546. (doi:10.1146/annurev.arplant.59.032607.092839)

43. Middleton PH et al. 2007 An ERF transcription factor in Medicago truncatula that is essential for nod factor signal transduction. Plant Cell. 19, 1221–1234. (doi:10.1105/tpc.106.048264)

44. Op den Camp R et al. 2011 LysM-type mycorrhizal receptor recruited for rhizobium symbiosis in nonlegume Parasponia. Science 331, 909–912. (doi:10.1126/science.1198181)

45. Cannon SB, Ilut D, Farmer AD, Maki SL, May GD, Singer SR, Doyle JJ. 2010 Polyploidy did not predate the evolution of nodulation in all legumes. PLoS ONE 5, e11630. (doi:10.1371/journal.pone.0011630)

46. Sauquet H et al. 2012 Testing the impact of calibration on molecular divergence times using a fossil-rich group: the case of Nothofagus (Fagales). Syst. Biol. 61, 289–313. (doi:10.1093/sysbio/syr116)

47. Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG. 1995 Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen-fixation in angiosperms. Proc. Natl Acad. Sci. USA 92, 2647–2651. (doi:10.1073/pnas.92.7.2647)

48. Li QG, Zhang L, Li C, Dunwell JM, Zhang YM. 2013 Comparative genomics suggests that an ancestral polyploidy event leads to enhanced root nodule symbiosis in the Papilionoideae. Mol. Biol. Evol. 30, 2602–2611. (doi:10.1093/molbev/mst152)

49. Schranz ME, Mohammadin S, Edger PP. 2012 Ancient whole genome duplications, novelty and diversification: the WGD radiation lag-time model. Curr. Opin. Plant Biol. 15, 147–153. (doi:10.1016/j.pbi.2012.03.011)

50. Schaefer H, Renner SS. 2011 Phylogenetic relationships in the order Cucurbitales and a new classification of the gourd family (Cucurbitaceae). Taxon 60, 122–138.

51. Borsch T, Hilu KW, Wiersema JH, Lohne C, Barthlott W, Wilde V. 2007 Phylogeny of Nymphaea (Nymphaeaceae): evidence from substitutions and microstructural changes in the chloroplast trnT-trnF region. Int. J. Plant Sci. 168, 639–671. (doi:10.1086/513476)

52. Fawcett J, Van de Peer Y, Maere S. 2013 Significance and biological consequences of polyploidization in land plant evolution. In Plant genome diversity, vol. 2 (ed. IJ Leitch), pp. 277–294. Wien, Germany: Springer.

53. Dobzhansky T. 1973 Nothing in biology makes sense except in the light of evolution. Am. Biol. Teach. 35, 125–129. (doi:10.2307/4444260)

54. Funk VA, Susanna A, Stuessy TF, Bayer RJ. 2009 Systematics, evolution, and biogeography of Compositae. Vienna, Austria: International Association for Plant Taxonomy.

55. Lynch M, Force A. 2000 The probability of duplicate gene preservation by subfunctionalization. Genetics 154, 459–473.

56. Werth CR, Windham MD. 1991 A model for divergent, allopatric speciation of polyploid pteridophytes resulting from silencing of duplicate-gene expression. Am. Nat. 137, 515–526. (doi:10.1086/285180)

57. Scannell DR, Frank AC, Conant GC, Byrne KP, Woolfit M, Wolfe KH. 2007 Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. Proc. Natl Acad. Sci. USA 104, 8397–8402. (doi:10.1073/pnas.0608218104)

58. Semon M, Wolfe KH. 2007 Reciprocal gene loss between Tetraodon and zebrafish after whole genome duplication in their ancestor. Trends Genet. 23, 108–112. (doi:10.1016/j.tig.2007.01.003)

59. Schnable JC, Freeling M, Lyons E. 2012 Genome-wide analysis of syntenic gene deletion in the grasses. Genome Biol. Evol. 4, 265–277. (doi:10.1093/gbe/evs009)

60. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999 Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151, 1531–1545.

61. Voordeckers K, Brown CA, Vanneste K, van der Zande E, Voet A, Maere S, Verstrepen KJ. 2012 Reconstruction of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovation through gene duplication. PLoS Biol. 10, e1001446. (doi:10.1371/journal.pbio.1001446)

62. Postlethwait J, Amores A, Cresko W, Singer A, Yan YL. 2004 Subfunction partitioning, the teleost radiation and the annotation of the human genome. Trends Genet. 20, 481–490. (doi:10.1016/j.tig.2004.08.001)

63. Volff JN. 2005 Genome evolution and biodiversity in teleost fish. Heredity 94, 280–294. (doi:10.1038/sj.hdy.6800635)

64. Rohde RA, Muller RA. 2005 Cycles in fossil diversity. Nature 434, 208–210. (doi:10.1038/nature03339)

65. Renne PR, Deino AL, Hilgen FJ, Kuiper KF, Mark DF, Mitchell WS, Morgan LE, Mundil R, Smit J. 2013 Time scales of critical events around the Cretaceous–Paleogene boundary. Science 339, 684–687. (doi:10.1126/science.1230492)

66. Robertson DS, Lewis WM, Sheehan PM, Toon OB. 2013 K–Pg extinction: reevaluation of the heat-fire hypothesis. J. Geophys. Res. Biogeosci. 118, 329–336. (doi:10.1002/jgrg.20018)

67. Raup DM. 1994 The role of extinction in evolution. Proc. Natl Acad. Sci. USA 91, 6758–6763. (doi:10.1073/pnas.91.15.6758)

68. Ehrendorfer F. 1980 Polyploidy and distribution. In Polyploidy – biological relevance (ed. WH Lewis), pp. 45–60. New York, NY: Plenum Press.

69. Brochmann C, Brysting AK, Alsos IG, Borgen L, Grundt HH, Scheen AC, Elven R. 2004 Polyploidy in arctic plants. Biol. J. Linn. Soc. 82, 521–536. (doi:10.1111/j.1095-8312.2004.00337.x)

70. Comai L. 2005 The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 6, 836–846. (doi:10.1038/nrg1711)

71. Madlung A. 2013 Polyploidy and its effect on evolutionary success: old questions revisited with new tools. Heredity 110, 99–104. (doi:10.1038/hdy.2012.79)

72. Parisod C, Holderegger R, Brochmann C. 2010 Evolutionary consequences of autopolyploidy. New Phytol. 186, 5–17. (doi:10.1111/j.1469-8137.2009.03142.x)

73. te Beest M, Le Roux JJ, Richardson DM, Brysting AK, Suda J, Kubesova M, Pysek P. 2012 The more the better? The role of polyploidy in facilitating plant invasions. Ann. Bot. 109, 19–45. (doi:10.1093/aob/mcr277)

74. Chen ZJ. 2013 Genomic and epigenetic insights into the molecular bases of heterosis. Nat. Rev. Genet. 14, 471–482. (doi:10.1038/Nrg3503)

75. Freeling M, Woodhouse MR, Subramaniam S, Turco G, Lisch D, Schnable JC. 2012 Fractionation mutagenesis and similar consequences of mechanisms removing dispensable or less-expressed DNA in plants. Curr. Opin. Plant Biol. 15, 131–139. (doi:10.1016/j.pbi.2012.01.015)

76. Goff SA. 2011 A unifying theory for general multigenic heterosis: energy efficiency, protein metabolism, and implications for molecular breeding. New Phytol. 189, 923–937. (doi:10.1111/j.1469-8137.2010.03574.x)

77. Chen ZJ. 2007 Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. Annu. Rev. Plant Biol. 58, 377–406. (doi:10.1146/annurev.arplant.58.032806.103835)

78. Wright KM, Pires JC, Madlung A. 2009 Mitotic instability in resynthesized and natural polyploids of the genus Arabidopsis (Brassicaceae). Am. J. Bot. 96, 1656–1664. (doi:10.3732/Ajb.0800270)

79. Wang YX, Jha AK, Chen RJ, Doonan JH, Yang M. 2010 Polyploidy-associated genomic instability in Arabidopsis thaliana. Genesis 48, 254–263. (doi:10.1002/dvg.20610)

80. Osborn TC et al. 2003 Understanding mechanisms of novel gene expression in polyploids. Trends Genet. 19, 141–147. (doi:10.1016/S0168-9525(03)00015-5)

81. Adams KL. 2007 Evolution of duplicate gene expression in polyploid and hybrid plants. J. Hered. 98, 136–141. (doi:10.1093/jhered/esl061)

82. Wang JL et al. 2006 Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. Genetics 172, 507–517. (doi:10.1534/genetics.105.047894)

83. Hegarty MJ, Barker GL, Wilson ID, Abbott RJ, Edwards KJ, Hiscock SJ. 2006 Transcriptome shock after interspecific hybridization in Senecio is ameliorated by genome duplication. Curr. Biol. 16, 1652–1659. (doi:10.1016/j.cub.2006.06.071)

84. Buggs RJA, Pannell JR. 2007 Ecological differentiation and diploid superiority across a moving ploidy contact zone. Evolution 61, 125–140. (doi:10.1111/j.1558-5646.2007.00010.x)

85. Rieseberg LH et al. 2003 Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**, 1211–1216. (doi:10.1126/science.1086949)

86. Husband BC, Ozimec B, Martin SL, Pollock L. 2008 Mating consequences of polyploid evolution in flowering plants: current trends and insights from synthetic polyploids. *Int. J. Plant Sci.* **169**, 195–206. (doi:10.1086/523367)

87. Hanada K, Kuromori T, Myouga F, Toyoda T, Li WH, Shinozaki K. 2009 Evolutionary persistence of functional compensation by duplicate genes in *Arabidopsis*. *Genome Biol. Evol.* **1**, 409–414. (doi:10.1093/gbe/evp043)

88. Hahn MA, van Kleunen M, Muller-Scharer H. 2012 Increased phenotypic plasticity to climate may have boosted the invasion success of polyploid *Centaurea stoebe*. *PLoS ONE* **7**, e50284. (doi:10.1371/journal.pone.0050284)

89. Paun O, Bateman RM, Fay MF, Luna JA, Moat J, Hedren M, Chase MW. 2011 Altered gene expression and ecological divergence in sibling allopolyploids of *Dactylorhiza* (Orchidaceae). *BMC Evol. Biol.* **11**, 113. (doi:10.1186/1471-2148-11-113)

90. Levin DA. 1975 Minority cytotype exclusion in local plant populations. *Taxon* **24**, 35–43. (doi:10.2307/1218997)

91. Rodriguez DJ. 1996 A model for the establishment of polyploidy in plants. *Am. Nat.* **147**, 33–46. (doi:10.1086/285838)

92. Arrigo N, Barker MS. 2012 Rarely successful polyploids and their legacy in plant genomes. *Curr. Opin. Plant Biol.* **15**, 140–146. (doi:10.1016/j.pbi.2012.03.010)

93. Harlan JR, De Wet JMJ. 1975 On Ö. Winge and a prayer: the origins of polyploidy. *Bot. Rev.* **41**, 361–390.

94. Ramsey J, Schemske DW. 1998 Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annu. Rev. Ecol. Syst.* **29**, 467–501. (doi:10.1146/annurev.ecolsys.29.1.467)

95. Felber F, Bever JD. 1997 Effect of triploid fitness on the coexistence of diploids and tetraploids. *Biol. J. Linn. Soc.* **60**, 95–106. (doi:10.1111/j.1095-8312.1997.tb01485.x)

96. Husband BC. 2004 The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. *Biol. J. Linn. Soc.* **82**, 537–546. (doi:10.1111/j.1095-8312.2004.00339.x)

97. Mason AS, Nelson MN, Yan GJ, Cowling WA. 2011 Production of viable male unreduced gametes in *Brassica* interspecific hybrids is genotype specific and stimulated by cold temperatures. *BMC Plant Biol.* **11**, 103. (doi:10.1186/1471-2229-11-103)

98. Suda J, Herben T. 2013 Ploidy frequencies in plants with ploidy heterogeneity: fitting a general gametic model to empirical population data. *Proc. R. Soc. B* **280**, 20122387. (doi:10.1098/rspb.2012.2387)

99. Mable BK. 2004 'Why polyploidy is rarer in animals than in plants': myths and mechanisms. *Biol. J. Linn. Soc.* **82**, 453–466. (doi:10.1111/j.1095-8312.2004.00332.x)

100. Otto SP, Whitton J. 2000 Polyploid incidence and evolution. *Annu. Rev. Genet.* **34**, 401–437. (doi:10.1146/annurev.genet.34.1.401)

101. Felber F. 1991 Establishment of a tetraploid cytotype in a diploid population: effect of relative fitness of the cytotypes. *J. Evol. Biol.* **4**, 195–207. (doi:10.1046/j.1420-9101.1991.4020195.x)

102. McHale NA. 1983 Environmental induction of high-frequency 2n pollen formation in diploid *Solanum*. *Can. J. Genet. Cytol.* **25**, 609–615. (doi:10.1139/g83-091)

103. Belling J. 1925 The origin of chromosomal mutations in *Uvularia*. *J. Genet.* **15**, 245–266. (doi:10.1007/BF02983121)

104. De Storme N, Geelen D. 2013 The impact of environmental stress on male reproductive development in plants: biological processes and molecular mechanisms. *Plant Cell Environ.* **37**, 1–18. (doi:10.1111/pce.12142)

105. Pecrix Y, Rallo G, Folzer H, Cigna M, Gudin S, Le Bris M. 2011 Polyploidization mechanisms: temperature environment can induce diploid gamete formation in *Rosa* sp. *J. Exp. Bot.* **62**, 3587–3597. (doi:10.1093/jxb/err052)

106. De Storme N, Copenhaver GP, Geelen D. 2012 Production of diploid male gametes in arabidopsis by cold-induced destabilization of postmeiotic radial microtubule arrays. *Plant Physiol.* **160**, 1808–1826. (doi:10.1104/pp.112.208611)

107. Song C, Liu SJ, Xiao J, He WG, Zhou Y, Qin Q, Zhang C, Liu Y. 2012 Polyploid organisms. *Sci. China Life Sci.* **55**, 301–311. (doi:10.1007/s11427-012-4310-2)

108. Kurschner WM, Batenburg SJ, Mander L. 2013 Aberrant *Classopollis* pollen reveals evidence for unreduced (2n) pollen in the conifer family Cheirolepidiaceae during the Triassic–Jurassic transition. *Proc. R. Soc. B* **280**, 20131708. (doi:10.1098/rspb.2013.1708)

109. Foster CB, Afonin SA. 2005 Abnormal pollen grains: an outcome of deteriorating atmospheric conditions around the Permian–Triassic boundary. *J. Geol. Soc.* **162**, 653–659. (doi:10.1144/0016-764904-047)

110. Visscher H, Looy CV, Collinson ME, Brinkhuis H, van Konijnenburg-van Cittert JH, Kurschner WM, Sephton MA. 2004 Environmental mutagenesis during the end-Permian ecological crisis. *Proc. Natl Acad. Sci. USA* **101**, 12 952–12 956. (doi:10.1073/pnas.0404472101)

111. Mraz P, Spaniel S, Keller A, Bowmann G, Farkas A, Singliarova B, Rohr RP, Broennimann O, Muller-Scharer H. 2012 Anthropogenic disturbance as a driver of microspatial and microhabitat segregation of cytotypes of *Centaurea stoebe* and cytotype interactions in secondary contact zones. *Ann. Bot.* **110**, 615–627. (doi:10.1093/Aob/Mcs120)

112. Oswald BP, Nuismer SL. 2011 A unified model of autopolyploid establishment and evolution. *Am. Nat.* **178**, 687–700. (doi:10.1086/662673)

113. Soltis DE, Buggs RJA, Doyle JJ, Soltis PS. 2010 What we still don't know about polyploidy. *Taxon* **59**, 1387–1403. (doi:10.2307/20774036)

114. Pysek P et al. 2009 The global invasion success of Central European plants is related to distribution characteristics in their native range and species traits. *Divers. Distrib.* **15**, 891–903. (doi:10.1111/j.1472-4642.2009.00602.x)

115. Thompson JD, Lumaret R. 1992 The evolutionary dynamics of polyploid plants—origins, establishment and persistence. *Trends Ecol. Evol.* **7**, 302–307. (doi:10.1016/0169-5347(92)90228-4)

116. Chao DY, Dilkes B, Luo H, Douglas A, Yakubova E, Lahner B, Salt DE. 2013 Polyploids exhibit higher potassium uptake and salinity tolerance in *Arabidopsis*. *Science* **341**, 658–659. (doi:10.1126/science.1240561)

117. Pandit MK, Pocock MJO, Kunin WE. 2011 Ploidy influences rarity and invasiveness in plants. *J. Ecol.* **99**, 1108–1115. (doi:10.1111/j.1365-2745.2011.01838.x)

118. Tavoletti S, Mariani A, Veronesi F. 1991 Phenotypic recurrent selection for 2n-pollen and 2n-egg production in diploid alfalfa. *Euphytica* **57**, 97–102. (doi:10.1007/Bf00023066)

119. Parrott WA, Smith RR. 1986 Recurrent selection for 2n pollen formation in red clover. *Crop Sci.* **26**, 1132–1135. (doi:10.2135/cropsci1986.0011183X002600060009x)

120. Henry IM, Dilkes BP, Tyagi AP, Lin HY, Comai L. 2009 Dosage and parent-of-origin effects shaping aneuploid swarms in *A. thaliana*. *Heredity* (*Edinb.*) **103**, 458–468. (doi:10.1038/hdy.2009.81)

121. De Storme N, Geelen D. 2013 Sexual polyploidization in plants: cytological mechanisms and molecular regulation. *New Phytol.* **198**, 670–684. (doi:10.1111/nph.12184)

122. Li Y, Shen Y, Cai C, Zhong C, Zhu L, Yuan M, Ren H. 2010 The type II *Arabidopsis* formin14 interacts with microtubules and microfilaments to regulate cell division. *Plant Cell.* **22**, 2710–2726. (doi:10.1105/tpc.110.075507)

123. De Storme N, Geelen D. 2011 The *Arabidopsis* mutant *jason* produces unreduced first division restitution male gametes through a parallel/fused spindle mechanism in meiosis II. *Plant Physiol.* **155**, 1403–1415. (doi:10.1104/pp.110.170415)

124. d'Erfurth I, Jolivet S, Froger N, Catrice O, Novatchkova M, Simon M, Jenczewski E, Mercier R. 2008 Mutations in AtPS1 (*Arabidopsis thaliana* parallel spindle 1) lead to the production of diploid pollen grains. *PLoS Genet.* **4**, e1000274. (doi:10.1371/journal.pgen.1000274)

125. Humayun MZ. 1998 SOS and mayday: multiple inducible mutagenic pathways in *Escherichia coli*. *Mol. Microbiol.* **30**, 905–910. (doi:10.1046/j.1365-2958.1998.01120.x)

126. Aertsen A, Michiels CW. 2005 Diversify or die: generation of diversity in response to stress. *Crit. Rev. Microbiol.* **31**, 69–78. (doi:10.1080/10408410590921718)

127. Radman M, Taddei F, Matic I. 2000 Evolution-driving genes. *Res. Microbiol.* **151**, 91–95. (doi:10.1016/S0923-2508(00)00122-4)

128. Arber W. 2000 Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS Microbiol. Rev.* **24**, 1–7. (doi:10.1111/j.1574-6976.2000.tb00529.x)

129. Gould SJ, Lewontin RC. 1979 The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B.* **205**, 581–598. (doi:10.1098/rspb.1979.0086)

130. Taylor JS, Raes J. 2004 Duplication and divergence: the evolution of new genes and old ideas. *Annu. Rev. Genet.* **38**, 615–643. (doi:10.1146/annurev.genet.38.072902.092831)

131. Ohno S. 1970 *Evolution by gene duplication*. Berlin, Germany: Springer.

132. Hahn MW. 2009 Distinguishing among evolutionary models for the maintenance of gene duplicates. *J. Hered.* **100**, 605–617. (doi:10.1093/jhered/esp047)

133. De Smet R, Van de Peer Y. 2012 Redundancy and rewiring of genetic networks following genome-wide duplication events. *Curr. Opin. Plant Biol.* **15**, 168–176. (doi:10.1016/j.pbi.2012.01.003)

134. Presser A, Elowitz MB, Kellis M, Kishony R. 2008 The evolutionary dynamics of the *Saccharomyces cerevisiae* protein interaction network after duplication. *Proc. Natl Acad. Sci. USA* **105**, 950–954. (doi:10.1073/pnas.0707293105)

135. Freeling M, Thomas BC. 2006 Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. *Genome Res.* **16**, 805–814. (doi:10.1101/gr.3681406)

136. Hakes L, Pinney JW, Lovell SC, Oliver SG, Robertson DL. 2007 All duplicates are not equal: the difference between small-scale and genome duplication. *Genome Biol.* **8**, R209. (doi:10.1186/gb-2007-8-10-r209)

137. Rodgers-Melnick E, Mane SP, Dharmawardhana P, Slavov GT, Crasta OR, Strauss SH, Brunner AM, DiFazio SP. 2012 Contrasting patterns of evolution following whole genome versus tandem duplication events in *Populus*. *Genome Res.* **22**, 95–105. (doi:10.1101/gr.125146.111)

138. Bekaert M, Edger PP, Pires JC, Conant GC. 2011 Two-phase resolution of polyploidy in the *Arabidopsis* metabolic network gives rise to relative and absolute dosage constraints. *Plant Cell.* **23**, 1719–1728. (doi:10.1105/tpc.110.081281)

139. Maere S, De Bodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, Van de Peer Y. 2005 Modeling gene and genome duplications in eukaryotes. *Proc. Natl Acad. Sci. USA* **102**, 5454–5459. (doi:10.1073/pnas.0501102102)

140. Birchler JA, Veitia RA. 2010 The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution. *New Phytol.* **186**, 54–62. (doi:10.1111/j.1469-8137.2009.03087.x)

141. Evangelisti AM, Conant GC. 2010 Nonrandom survival of gene conversions among yeast ribosomal proteins duplicated through genome doubling. *Genome Biol. Evol.* **2**, 826–834. (doi:10.1093/gbe/evq067)

142. Arnegard ME, Zwickl DJ, Lu Y, Zakon HH. 2010 Old gene duplication facilitates origin and diversification of an innovative communication system: twice. *Proc. Natl Acad. Sci. USA* **107**, 22 172–22 177. (doi:10.1073/pnas.1011803107)

143. Ihmels J, Bergmann S, Gerami-Nejad M, Yanai I, McClellan M, Berman J, Barkai N. 2005 Rewiring of the yeast transcriptional network through the evolution of motif usage. *Science* **309**, 938–940. (doi:10.1126/science.1113833)

144. Conant GC, Wolfe KH. 2007 Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Mol. Syst. Biol.* **3**, 129. (doi:10.1038/msb4100170)

145. Braasch I, Schartl M, Volff JN. 2007 Evolution of pigment synthesis pathways by gene and genome duplication in fish. *BMC Evol. Biol.* **7**, 74. (doi:10.1186/1471-2148-7-74)

146. Veron AS, Kaufmann K, Bornberg-Bauer E. 2007 Evidence of interaction network evolution by whole-genome duplications: a case study in MADS-box proteins. *Mol. Biol. Evol.* **24**, 670–678. (doi:10.1093/molbev/ms1197)

147. Stebbins GL. 1950 *Variation and evolution in plants*. New York, NY: Columbia University Press.

148. Wagner WH. 1970 Biosystematics and evolutionary noise. *Taxon* **19**, 146–151. (doi:10.2307/1217945)

149. Mayrose I, Zhan SH, Rothfels CJ, Magnuson-Ford K, Barker MS, Rieseberg LH, Otto SP. 2011 Recently formed polyploid plants diversify at lower rates. *Science* **333**, 1257. (doi:10.1126/science.1207205)

150. Madlung A, Tyagi AP, Watson B, Jiang H, Kagochi T, Doerge RW, Martienssen R, Comai L. 2005 Genomic changes in synthetic Arabidopsis polyploids. *Plant J.* **41**, 221–230. (doi:10.1111/j.1365-313X.2004.02297.x)

151. Goldschmidt R. 1940 *The material basis of evolution*. New Haven, CT: Yale University Press.