Polyploidy and Genome Evolution in Plants

Pamela S. Soltis[1,2]
D. Blaine Marchant[1,3]
Yves Van de Peer[4,5,6]
Douglas E. Soltis[1,2,3]

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

Abstract
Plant genomes vary in size and complexity, fueled in part by processes of whole-genome duplication (WGD; polyploidy) and subsequent genome evolution. Despite repeated episodes of WGD throughout the evolutionary history of angiosperms in particular, the genomes are not uniformly large, and even plants with very small genomes carry the signatures of ancient duplication events. The processes governing the evolution of plant genomes following these ancient events are largely unknown. Here, we consider mechanisms of diploidization, evidence of genome reorganization in recently formed polyploid species, and macroevolutionary patterns of WGD in plant genomes and propose that the ongoing genomic changes observed in recent polyploids may illustrate the diploidization processes that result in ancient signatures of WGD over geological timescales.

Introduction

Plant genomes vary spectacularly in size, ranging from 0.063 Gb to 148.8 Gb, a 2400-fold difference (Dodsworth et al. this issue; www.data.kew.org/cvalues). Much of this diversity results from differential expansion and loss of repeats (reviewed in Leitch and Leitch 2013), but an additional major driver shaping variation in genome size in plants is whole-genome duplication (WGD; polyploidy). Moreover, genome structure and gene content in plants are intimately tied to the history of WGD. We therefore contend that any understanding of plant genome structure, content, and evolution requires consideration of WGD and its consequences.

Polyploidy in Plants
Polyploidy has long been considered an important mechanism of speciation in plants, particularly angiosperms (e.g., see reviews by Stebbins 1947, 1950; Grant 1971, 1981; Soltis and Soltis, 1993, 2000, 2009; Otto and Whitten 2000). The estimated frequency of polyploidy in major clades of green plants (Viridiplantae) has varied dramatically, ranging from very low in bryophytes to as high as 95% for ferns (Grant, 1981). For angiosperms alone, estimates have ranged from 30-35% (Stebbins 1950) to ~70% (Masterson 1994), with most estimates near 50% (e.g., Müntzing 1936; Darlington 1937; Grant 1963, 1971). Genomic data for plants, in contrast, have demonstrated a dazzling history of repeated

1

WGDs throughout evolutionary history. Even the small genome of *Arabidopsis thaliana* (0.157 Gb) – the small genome was one of the keys to the selection of this species as a genomic model – shows signatures of ancient WGD (Vision, 2002; Bowers et al. 2003). All angiosperms share an ancient WGD, as do all seed plants (Jiao et al. 2011). Thus, in recent years, interpretations of plants as 'diploids' or 'polyploids' have blurred, requiring much more nuanced vocabulary to describe plant genomes.

The origin of new species via polyploidy requires a series of seemingly low-probability events of hybridization, unreduced gamete formation, and establishment (see e.g., Levin 1975; Ramsey and Schemske 1998, 2002). Despite these apparent barriers, polyploids are common elements of all floras worldwide and are particularly abundant at high latitudes and high elevations (e.g., Ehrendorfer 1980; Brochmann et al. 2004). The 'success' of polyploids is often attributed to the increased genetic diversity held within single polyploid individuals relative to that of their diploid progenitors (e.g., Roose and Gottlieb 1976; Rieseberg and Soltis 1986; Soltis and Soltis 1989, 1993, 2000, 2009; Brochmann et al. 2004). Moreover, this genetic diversity may be manifested in novelty at the biochemical, physiological, morphological, and ecological levels, giving polyploids an advantage, at least in the short term, over their diploid parents (e.g., Levin 1983; Soltis et al. 2014). However, if polyploidy were a uniformly successful strategy, then plant genomes, such as that of *A. thaliana*, should show more obvious evidence of WGDs, such as high chromosome numbers, large genomes, and routinely duplicate (triplicate, quadruplicate, etc.) gene copies. Instead, we see many apparently recently formed polyploids with these attributes, but other species, such as *A. thaliana*, harbor ancient signatures of WGD within generally 'diploid' genomes but without many of the expectations of WGD. Certainly, processes of diploidization are at play, leading to repeated cycles of polyploidy followed by diploidization followed by polyploidy, etc. (e.g., Haufler 1987; Wendel 2015; Soltis et al. 2016).

But what are these processes of diploidization, and how can they be reconciled with observations of gene family diversity, ancient signatures of WGD, and macroevolutionary patterns of polyploidization? In this paper, we will attempt to unify (i) hypothesized mechanisms of of diploidization, (ii) data on genome reorganization shortly after polyploidization, based on the evolutionary model, *Tragopogon*, and other species, and (iii) macroevolutionary patterns in plant genomes. We propose that the ongoing genomic changes observed in recent polyploids, such as *Tragopogon*, may represent the diploidization processes that result in ancient signatures of WGD over geological timescales.

Mechanisms of Diploidization
Repetitive DNA sequences comprise substantial portions of plant genomes (e.g. *A. thaliana*: 15-20%; maize ~85%) and can largely influence genome size (Leitch and Leitch 2008; Schnable et al. 2009). Consequently, the two mechanisms by which these repetitive sequences (i.e. retrotransposons, DNA transposons, simple repeats) are lost (illegitimate recombination and unequal intra-strand homologous recombination) are the principal processes responsible for genome downsizing to a "diploid" state. Illegitimate recombination is hypothesized to remove DNA sequences via double-stranded breaks and/or slippage during replication, whereas unequal homologous recombination occurs between two repeat sequences and results in the loss of the DNA between the repeats, as

well as one of the repeats (Devos et al. 2002). The relative importance of the two primary diploidization mechanisms appears to be species-specific in angiosperms, as the efficiency of DNA loss via either mechanism is highly variable and not phylogenetically related (Devos et al. 2002; Bennetzen et al. 2005; Vitte and Bennetzen 2006). Very few studies have addressed these processes outside of flowering plants; however, recent genomic sequencing of three conifer taxa have suggested that the enormous genomes of these taxa could be due to a very low recombination rate and high homologous chromosome fidelity, removing the opportunity for genome-downsizing (Nystedt et al. 2013), at least via these mechanisms. (Figure 1)
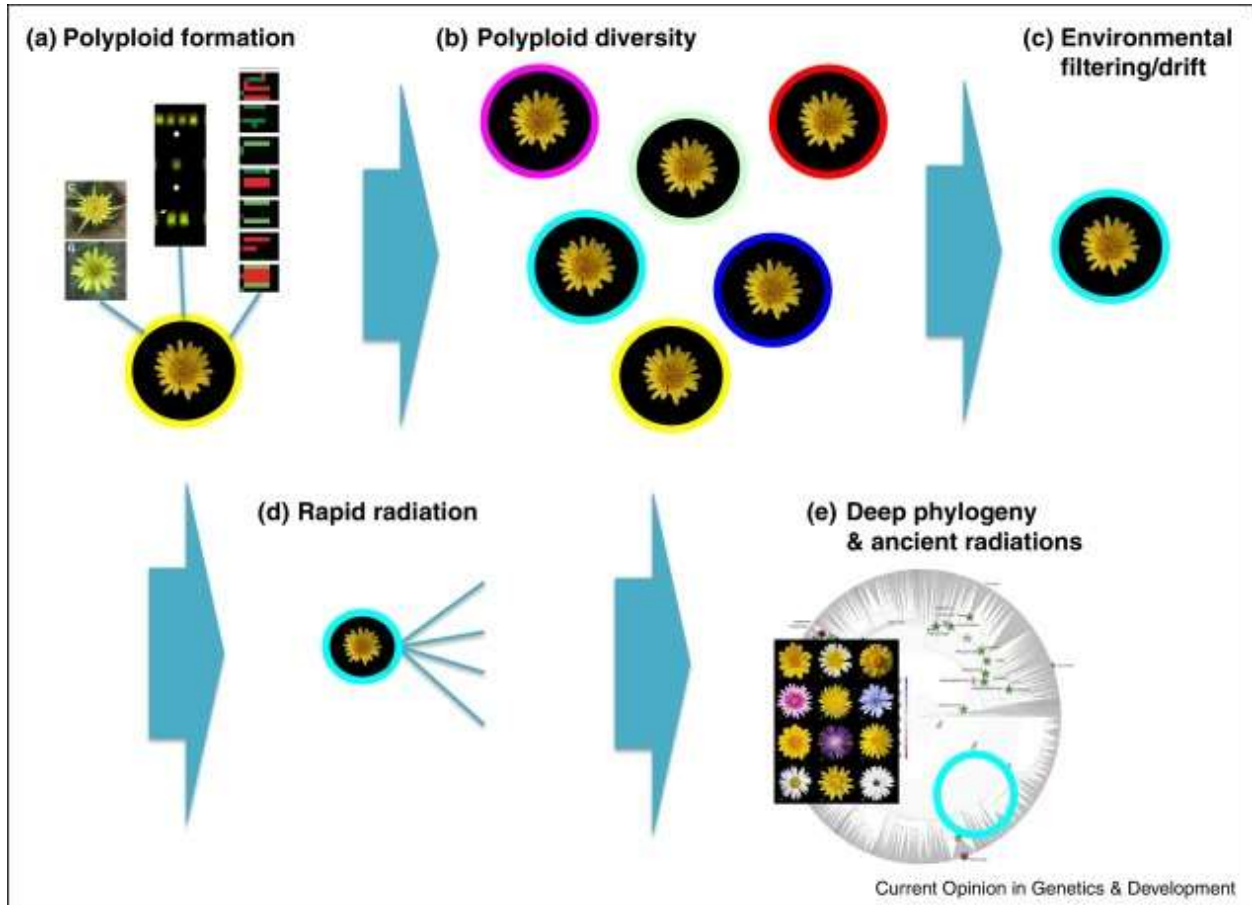

Figure 1. Timeline showing origin of polyploid species, generation of novelty, environmental filtering, rapid radiation from selected line, and resulting signature of paleopolyploidy. Note that diploidization processes (not pictured) may occur continuously through this timeline. **(a)** Polyploid formation (in yellow circle) and attendant processes that yield variation in, for example, inflorescence morphology, karyotype, and homeolog loss; other novelties, for example, altered gene expression, transposon activity, are not pictured. **(b)** Array of polyploid genotypes/phenotypes, represented by different colors of rings around the photograph. **(c)** A single remaining polyploid 'type' after environmental filtering or drift. **(d)** Rapid radiation from this 'successful' polyploid and resultant formation of multiple lineages that trace back to this single common ancestor. **(e)** Deep phylogeny showing the correspondence (turquoise ring) between ancient WGD and radiation, as derived in d. Other paleopolyploid events are also shown.

Although great strides have been made in determining the processes by which DNA content is reduced following polyploidy, few studies have addressed the second aspect of

diploidization: the mechanisms by which entire chromosomes are lost. Only recently have synthetic polyploid studies demonstrated the high prevalence of chromosomal instability immediately after genome duplication (Mestiri et al. 2010; Xiong et al. 2011; Chester et al. 2012; Zhang et al. 2013). While aneuploidy has been found to negatively correlate with fertility in synthetic *Brassica napus* polyploids (Xiong et al. 2011) and pollen viability in synthetic wheat polyploids (Zhang et al. 2013), the fact that chromosome number is no longer a suitable corollary for polyploidy history (e.g. *A. thaliana* has five chromosomes and five known polyploidy events), there must be some selective force to reduce chromosome number. Polyploid systems in which multiple temporal polyploid samplings (synthetic, nascent, ancient) arise will be critical for evaluating these diploidization processes in the near future.

Genome Reorganization in Recent Polyploids: Generating Novelty
The genomes of newly formed natural polyploids, as well as those of synthetic polyploids, may experience rapid changes in homeolog loss, as well as genome restructuring post-polyploidization, and altered patterns of gene expression may set the stage for subsequent loss of duplicate gene copies (e.g., Ainouche et al. 2012; Hegarty et al. 2012; Soltis et al. 2012, in prep.; Madlung et al., 2005; Akama et al., 2014; Gaeta et al., 2007; Xiong et al., 2011; Shen et al. 2015; Schnable et al. Freeling maize paper). Both the extent and speed with which these diverse changes occur may vary considerably across diverse polyploidy systems (reviewed in Soltis et al. in prep).

In the recent and repeatedly formed allotetraploids *Tragopogon mirus* and *T. miscellus*, species in the sunflower family that originated in the early 20th century (Ownbey 1950), frequent homeolog loss, subfunctionalization, and major chromosomal changes, including translocations and compensated and non-compensated aneuploidy, were detected in natural populations, as well as in synthetic lines. Transcriptomic shock was observed; hybridization and polyploidy *per se* both play important roles in these young polyploids (Buggs et al., 2010; 2011a,b, 2012; reviewed in Soltis et al. 2012). Investigations of an older allotetraploid (*T. castellanus*) and its parents indicate that gene loss/expression changes and chromosomal alterations mirror what is seen in the recently formed *T. mirus* and *T. miscellus* and demonstrate that some of the alterations that occur immediately post-polyploidization may be retained over long evolutionary timeframes (Mavrodiev et al., 2015; Soltis, Soltis, Barbazuk, Schable et al. unpubl.). In *Senecio cambrensis* (also in the sunflower family and estimated to have originated in the 1700s; see review by Hegarty et al. 2012), transcriptome shock was also detected; hybridization altered gene expression and DNA methylation, and genome duplication resulted in an additional burst of transcriptional and epigenetic change (Hegarty et al., 2012). In the grass *Spartina anglica*, which originated in the 1800s (see review by Ainouche et al. 2012), rapid changes in gene expression were observed; hybridization played a larger role in methylation changes than polyploidy *per se* (Ainouche et al. 2012). Transcriptomic shock was detected; at the transcriptomic level, both hybridization and polyploidy are important. No chromosomal changes were noted.

Synthetic lines of older, established polyploids, including *A. suecica* (Madlung et al., 2005; Chen 2007; Wright et al., 2009; Akama et al., 2014), *Brassica napus* (Gaeta et al., 2007;

Xiong et al., 2011), and a synthetic *Brassica* hexaploid (Shen et al., 2015), have also been used to assess genomic and expression changes that arise shortly after polyploid formation. Transcriptome shock and rapid changes in expression and methylation are also observed, and the relative importance of hybridization and genome doubling seems to vary among species. The extent of homeolog loss versus expression changes also varies. Homeolog loss and chromosomal changes are frequent in *Brassica* (as in the natural polyploids of *Tragopogon*), but expression changes predominate in other systems (e.g., *Arabidopsi*s, as in the natural polyploid, *Spartina anglica*, which has a stable karyotype).

Comparisons of synthetic and natural polyploids also indicate variation in the repeatability of evolution across independently formed polyploid lines, whether natural or synthetic. In *Tragopogon*, *Senecio*, and *A. suecica,* the consequences of polyploidy are repeated—that is, the evolutionary tape of life is replayed. However, in *Brassica,* a synthetic *Arabidopsis* hexaploid (Shen et al. 2015), and *Spartina* polyploids, independent origins respond differently. However, many aspects of polyploidy cannot be compared across all of these systems because of large gaps in the overall data set. For example, other features of polyploid genomes that may contribute to both genomic novelty in the short term and lead to genome downsizing over the longer term – e.g., transposon activity, methylation, subfunctionalization, and proteomic diversity – are only available for a few systems (Soltis et al. in prep).

Macroevolutionary Patterns of Genome Evolution in Plants
The fact that there are many recognizable polyploids of fairly recent origin, but relatively little evidence for many ancient WGD events (paleopolyploidy; at least within the same evolutionary lineage), provides an interesting paradox. Although methods and data for detecting ancient WGDs are still limited, the inferred number of such events is increasing rapidly; cotton (*Gossypium hirsutum*), for example, is a polyploid with $2n = 52$; the polyploid cottons originated in the last one million years (Wendel and Cronn 2003), and have an estimated xx WGDs in its evolutionary history back to the origin of flowering plants (ref). However, even as the picture of ancient WGDs is clarified, the number of such events will likely continue to underestimate the frequency of extant polyploids. This relative paucity of paleopolyploidy suggests that polyploidy may be an evolutionary dead end, except perhaps in specific cases. Indeed, at some time in evolution, organisms that underwent and survived WGDs must have had an adaptive advantage. Examples of ancient WGD events that have been established on the longer term are one or two WGDs early in the evolution of seed and flowering plants (Jiao et al. 2011), one WGD that is ancestral to most or all of the eudicots (Vekemans et al., 2012; Jiao et al. 2012), and one or two that occurred early in the monocot lineage (Bowers et al., Paterson et al. 2012). Therefore, a question that has received much attention of late is whether these key ancient WGDs, which in many cases characterize major lineages of flowering plants, have survived by coincidence, or whether they may have originated in concert, at very specific geological times, for instance during times of major ecological or environmental upheaval, and/or periods of extinction. In this respect, one of the most striking cases is a wave of WGDs in different flowering plant lineages that seem to coincide with the Cretaceous/Paleogene (K/Pg) boundary (Fawcett et al., 2009; Vanneste et al., 2014a, b; Vanneste et al., 2015; Ciao et al., 2015; Olsen et al., 2015). Furthermore, many of the WGDs clustered around the K/Pg extinction event are at the base

of some of the largest and most successful extant plant families. Polyploidy thus somehow appears to be correlated with plant survival through the K/Pg boundary (Vanneste et al., 2014a) and with species diversification in angiosperms (Tank et al. 2015).

Once polyploids are formed, they must become locally established, reproduce, and survive while adapting to different environments. These processes might ultimately lead to their long-term evolutionary success, where their descendant lineages survive for tens of millions of years. Most likely, both neutral and adaptive processes contribute to polyploid establishment under stressful conditions in the short term. The adaptive scenario is mostly based on characteristics often displayed by newly formed polyploids, such as the formation of more extreme phenotypes in the resulting hybrid populations compared with their diploid parents. Moreover, genomic instability and gene expression changes soon after polyploid formation (shown to occur in both recent natural polyploids and synthetics; above) may result in increased phenotypic variability, which might be advantageous and allow rapid adaptation to changed environments and conditions (Rieseberg et al., 2003; Comai, 2005; Madlung, 2013; Te Beest et al. 2012). Other potential adaptive advantages of newly formed polyploids include the masking of deleterious recessive alleles leading to increased mutational robustness. The neutral scenario gets support from the fact that levels of unreduced gamete formation can be increased by external stimuli such as stress and a fluctuating environment (De Storme and Geelen, 2013). Temperature in particular has a pronounced effect on unreduced gamete formation. Moreover, increased levels of unreduced pollen in the fossil record 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 (Kurschner et al., 2013), while abnormal gymnosperm pollen (Foster and Afonin, 2005) and lycophyte spores (Visscher et al., 2004) have also been reported from 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 paleopolyploidizations at the K/Pg boundary. It could also explain why many present-day polyploids often are more abundant in stressful environments, such as the Arctic (Brochmann et al., 2004) or disturbed habitats (Mraz et al., 2012).

Although many genes undergo homeolog loss after WGD, others, particularly regulatory and developmental genes, are retained in excess after WGD. This pattern of gene loss and retention is most likely due to dosage-balance constraints and selection against loss of individual components of completely duplicated macromolecular complexes and/or pathways, because this would disrupt their overall stoichiometry (Maere et al., 2005; see also Birchler and Veitia 2007, 2012; Conant et al. 2014). Retention of dosage-sensitive duplicates thus does not provide an immediate evolutionary advantage and adaptation, 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 (Freeling and Thomas, 2006), which might later on facilitate evolutionary innovations and/or diversifications. Selection to maintain dosage balance eventually relaxes over time allowing functional divergence and duplicated networks to be rewired to evolve novel functionality and increase biological complexity (De Smet and Van de Peer, 2012), which

could help explain the vast post-WGD success observed in some of the plant families that experienced a WGD at the K/Pg boundary.

Synthesis: Linking Microevolutionary Processes with Macroevolutionary Patterns
High levels of unreduced gamete formation in natural populations of angiosperms (e.g. Ramsey and Schemske 1998, 2002) provide a mechanism for polyploid formation and ultimately in the case of successful polyploids, speciation, via allopolyploidy (involving interspecific hybridization) and autopolyploidy (formed within a species). Both modes of formation contribute substantially to angiosperm species diversity (Barker et al. 2015). Nascent polyploids undergo an array of genomic and expression-level processes that may result simply from the presence of multiple genomes within the same nucleus. Although resolution of this multi-genome challenge may take multiple forms (chromosomal restructuring, homeolog loss, alterations in gene expression, etc.), it is clear that polyploids are not merely the additive products of their diploid progenitors. Instead, they are mosaics of parental, additive, and novel features, and even young polyploid species appear to be composed of arrays of genetically unique individuals. This novelty and range of phenotypic diversity may provide polyploid species with unusual adaptive capacity, particularly in times of high environmental stress. In fact, WGD events in angiosperms are non-randomly associated with bursts in diversification (Tank et al. 2015), and these radiations tend to be marked by novelty in morphology and/or chemistry.
It is intriguing indeed to consider that the processes we observe in recent polyploids may explain patterns of WGD and key innovations across macroevolutionary timescales.

Acknowledgments

Literature Cited

Dodsworth et al.
Leitch and Leitch
Stebbins 1947
Stebbins 1950
Grant 1971
Grant 1981
DSoltis and Soltis 1993
PSoltis and Soltis 2000
PSoltis and Soltis 2009
Otto and Whitten 2000
Masterson 1994
Müntzing 1936;
Darlington 1937;
Grant 1963,

Vision et al 2002
Bowers et al 2003
Jiao et al 2011
Levin1975
Ramsey and Schemske 1998
Ramsey and Schemske 2002
Ehrendorfer 1980?
Brochmann et al 2004
Roose and Gottlieb 1976
Rieseberg and Soltis 1986
DSoltis and Soltis 1989
Levin 1983
PSoltis et al 2014
Haufler 1987
Wendel 2015
DSoltis et al 2016

Bennetzen JL, Ma J, Devos KM: **Mechanisms of recent genome size variation in flowering plants** *Annals of Botany* 2005, **95**: 127-132.

Chester M, Gallagher JP, Symonds VV, da Silva AVC, Mavrodiev EV, Leitch AR, Soltis PS, Soltis DE. **Extensive chromosomal variation in a recently formed natural allopolyploid species, Tragopogon miscellus (Asteraceae)** *Proceedings of the National Academy of Sciences* 2012, **109**: 1176-1181.

Devos KM, Brown JKM, Bennetzen JL: **Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis** *Genome Research* 2002, **12**: 1075-1079.

Leitch AR, Leitch IJ: **Genomic plasticity and the diversity of polyploid plants** *Science* 2008, **320:** 481-483.

Mestiri I, Chagué V, Tanguy AM, Huneau C, Huteau V, Belcram H, Coriton O, Chalhoub B, Jahier J: **Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity** *New Phytologist* 2010, **186**: 86-101.

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, et al: **The Norway spruce genome sequence and conifer genome evolution** *Nature* 2013, **497**: 579-584.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, et al: **The B73 maize genome: complexity, diversity, and dynamics** S*cience* 2009, **326**: 1112-1115.

Vitte C, Bennetzen JL: **Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution** *Proceedings of the National Academy of Sciences* 2006, **103**: 17638-17643.

Xiong Z, Gaeta RT, Pires JC: **Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid Brassica napus.** *Proceedings of the National Academy of Sciences* 2011, **108**: 7908-7913.

Zhang H, Bian Y, Gou X, Zhu B, Xu C, Qi B, Li N, et al: "**Persistent whole-chromosome aneuploidy is generally associated with nascent allohexaploid wheat** *Proceedings of the National Academy of Sciences* 2013, **110**: 3447-3452.

Ainouche et al. 2012;
Hegarty et al. 2012;
Soltis et al. 2012;
Madlung et al., 2005;
Akama et al., 2014;
Gaeta et al., 2007;

Xiong et al., 2011;
Shen et al. 2015
Buggs et al., 2010;
Buggs et al., 2011
Buggs et al., 2012a
Buggs et al., 2012b
Mavrodiev et al., 2015
Madlung et al., 2005;
Chen 2007;
Wright et al., 2009;

Rieseberg LH et al. 2003 Major ecological transitions in wild sunflowers facilitated by hybridization. Science 301, 1211–1216.

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. 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.

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.

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

Comai L. 2005 The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 6, 836–846.

Freeling M, Thomas BC. 2006 Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. Genome Res. 16, 805–814.

Fawcett, J., 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 106, 5737-42.

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.

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.

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

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.

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.

Madlung A. 2013 Polyploidy and its effect on evolutionary success: old questions revisited with new tools. Heredity 110, 99–104.

Vanneste, K., Baele, G., Maere, S., Van de Peer, Y. (2014) Analysis of 41 plant genomes supports a wave of successful genome duplications at the Cretaceous-Tertiary boundary. Genome Res. 24,1334-1347 .

Vanneste, K., Maere, S., Van de Peer, Y. (2014) Tangled up in two: A burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution. Philosophical Transactions B 369, 20130353 .

Vanneste, K., Sterck, L., Myburg, A.A., * Van de Peer, Y., * Mizrachi, E. (2015) Horsetails are ancient polyploids: Evidence from *Equisetum giganteum*. The Plant Cell 27: 1567–1578

Cai, H., Liu, X., Vanneste, K., Proost, S., Tsai, C., Liu, K.-W., Chen, L.-J., He, Y., Xu, SX., Bian, C., Zheng, Z., Sun, F., Liu, W., , ., Pan, Z.-Y., Hsu, C.-C., Yang, Y.-P., Hsu, Y.-C., Chuang, Y.-C., Dievart, A., Dufayard, J.-F., Xu, SX., Wang, X., Wang, X., Xiao, X.-J., Zhao, X.-M., Du, Y., Zhang, G.-Q., Wang, X., Su, Y.-Y., Xie, G.-C., Liu, G.-H., Li, L.-Q., Huang, L.-Q., Luo, Y., Chen, H.-H., Van de Peer, Y., Liu, Z.-J. (2014) The genome sequence of the orchid *Phalaenopsis equestris*. Nature Genetics 47, 65–72.
Tank et al. 2015

Glossary

- Allopolyploidy : polyploidy formed through the combined processes of interspecific hybridization and genome doubling.
- Autopolyploidy: genome doubling that arises within a species; it may involve a single individual or crossing between individuals from genetically distinct lineages within the species.
- Diploidization: the processes that return a polyploid genome to a diploid-like genome; these may include loss of duplicate genes and chromosomes, loss of repetitive DNA, gene silencing, altered chromosome pairing.
- Diploidy: the state of being diploid; that is, containing two complete sets of chromosomes (or genomes).
- Fractionation: the loss of one copy of a gene pair duplicated by polyploidy; losses may be random with respect to the parental genome or biased, with most/all losses from a single parental genome.
- Homeolog (also homoeolog): chromosomes (and the genes they carry) that are duplicated by polyploidy.
- Polyploidy: the state of having more than two complete sets of chromosomes.
- Polyploidization: the process(es) of polyploid formation; this can be duplication, triplication, or higher-order multiplication of a genome.
- Whole-genome duplication: the duplication of a complete genome, for example, of a diploid genome (with two copies of each chromosome) to form a tetraploid (with four copies of each chromosome); this term is sometimes used to refer to the process of duplication (i.e., polyploidization) and sometimes in reference to the state of having multiple, duplicate genomes (i.e., polyploidy).