Journal of the Linnean Society

Botanical Journal of the Linnean Society, 2015, ••, ••–••. With figures

Is post-polyploidization diploidization the key to the evolutionary success of angiosperms?

STEVEN DODSWORTH[1,2]*, MARK W. CHASE[2,3] and ANDREW R. LEITCH[1]

[1]*School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK*
[2]*Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK*
[3]*School of Plant Biology, The University of Western Australia, Crawley, WA 6009, Australia*

Received 30 September 2015; revised 30 September 2015; accepted for publication 1 October 2015

Advances in recent years have revolutionized our understanding of both the context and occurrence of polyploidy in plants. Molecular phylogenetics has vastly improved our understanding of plant relationships, enabling us to better understand trait and character evolution, including chromosome number changes. This, in turn, has allowed us to appreciate better the frequent occurrence and extent of polyploidy throughout the history of angiosperms, despite the occurrence of low chromosome numbers in some groups, such as in *Arabidopsis* (*A. thaliana* was the first plant genome to be sequenced and assembled). In tandem with an enhanced appreciation of phylogenetic relationships, the accumulation of genomic data has led to the conclusion that all angiosperms are palaeopolyploids, together with better estimates of the frequency and type of polyploidy in different angiosperm lineages. The focus therefore becomes when a lineage last underwent polyploidization, rather than simply whether a plant is 'diploid' or 'polyploid'. This legacy of past polyploidization in plants is masked by large-scale genome reorganization involving repetitive DNA loss, chromosome rearrangements (including fusions and fissions) and complex patterns of gene loss, a set of processes that are collectively termed 'diploidization'. We argue here that it is the diploidization process that is responsible for the 'lag phase' between polyploidization events and lineage diversification. If so, diploidization is important in determining chromosome structure and gene content, and has therefore made a significant contribution to the evolutionary success of flowering plants. © 2015 The Authors. Botanical Journal of the Linnean Society published by John Wiley & Sons Ltd on behalf of The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2015, ••, ••–••.

ADDITIONAL KEYWORDS: chromosome number – flowering plants – genome downsizing – genome rearrangement – genomics – lag phase – polyploidy – WGD radiation lag-time model.

POLYPLOIDY AND DIPLOIDIZATION

Polyploidy, or whole genome duplication (WGD), is a frequent phenomenon in plants, especially in flowering plants. It has been estimated that *c*. 15% of angiosperm speciation events involve a change in ploidy (neopolyploidy; Wood *et al*., 2009) and that all flowering plants have experienced at least one WGD episode in their evolutionary history (palaeopolyploidy; Bowers *et al*., 2003; Blanc & Wolfe, 2004; Van de Peer, Maere & Meyer, 2009; Jiao *et al*., 2011). Ferns contain an even greater number of speciation events involving polyploidy (~31%; Wood *et al*., 2009) and they are also the group of plants with the highest reported chromosome number ($2n = c.$ 1440; Abraham & Ninan, 1954). In ferns, multiple rounds of polyploidy occur apparently without the same diploidization processes that mask ancestral polyploidy in angiosperms (Leitch & Leitch, 2012). What role, if any, diploidization plays in the story of fern evolution is currently unknown. Ferns are relatively homogeneous in terms of developmental flexibility, morphological diversity and ecological specialization, at least when compared with angiosperms. In contrast, angiosperms exhibit a plethora of floral and vegetative forms that are often thought to account for their diversification and abundance relative to that of gymnosperms, ferns, lycopods and bryophytes. In

*Corresponding author. E-mail: s.dodsworth@cantab.net

particular, annual life histories are almost unknown outside the angiosperms (there are a few ferns that have managed this). Might diploidization following polyploidy, particularly following allopolyploidy (hybridization involving polyploidy), be a crucial factor in expanding evolutionary innovation versus relative evolutionary stasis?

EVOLUTIONARY DYNAMICS OF POLYPLOIDS

As a result of their prevalence in plants, polyploids have been speculated to hold a selective advantage over diploids through the evolution of novel genetic (and indeed genomic) variation (Soltis & Soltis, 2000; Leitch & Leitch, 2008; Flagel & Wendel, 2009). In theory, duplicated genes provide the substrate for mutation-driven evolution of new copies, as a result of freedom from selective constraints. With multiple copies comes the potential for subfunctionalization and/or neofunctionalization, the two often being difficult to distinguish. The extent of neofunctionalization is currently unknown and is difficult to document empirically. However, subfunctionalization is a relatively common phenomenon in angiosperms. The origin of flowers, for instance, ostensibly requires the concerted function of various MADS box transcription factor complexes, and the evolution of such transcription factors has been attributed to ancient (i.e. as a result of palaeopolyploidy) and recent gene-specific duplications, with subsequent subfunctionalization of paralogous gene copies. It has become apparent in orchids, the most species rich and perhaps most florally diverse family of angiosperms, that subfunctionalization of duplicated B genes has led to the development of three petal-like organs: three outer tepals, two inner tepals and a highly modified lip (Mondragón-Palomino & Theißen, 2011; Mondragón-Palomino, 2013). Polyploids also have increased fixed heterozygosity, leading to increased heterosis and a higher tolerance of selfing (perhaps even promoting the evolution of self-compatibility), which leads to a tolerance of habitat fragmentation and population disturbance. In some cases, polyploids may also occupy new ecological niches or a broader range of niches compared with their diploid relatives. Collectively, these factors may contribute to the success of polyploids as invasive species (Pandit, Pocock & Kunin, 2011).

In contrast with the above advantages, polyploidy can also create a barrier to selection, as new mutations are masked by existing alleles, thereby 'diluting the effects of new mutations' (Stebbins, 1971). This depends on the dominance of new beneficial mutations versus the pre-existing allele; therefore, if it is at least partially recessive it will be masked, resulting in inefficient selection (Stebbins, 1971; Otto, 2007), and leading to the idea that polyploids are 'evolutionary dead-ends' (Stebbins, 1950). The loss of some duplicate copies following polyploidy can cause a dosage imbalance and disruptions to gene networks – the gene balance hypothesis (Birchler & Veitia, 2007). This is particularly significant for genes that contribute to macromolecular complexes, which are often retained post-polyploidy in order to maintain a dosage-sensitive relationship (Conant, Birchler & Pires, 2014). Furthermore, recent polyploids typically have reduced fertility as a result of pairing problems at meiosis (Chester *et al*., 2012; Yant *et al*., 2013). They can also, on formation, experience 'genomic shock' from the combination of two disparate subgenomes in one nucleus, resulting in an elevated frequency of (retro)transposition (McClintock, 1984; Petit *et al*., 2010) and chromosomal rearrangements that reduce fitness, potentially leading to extinction (Leitch & Leitch, 2008; Mayrose *et al*., 2011). Newly formed polyploids are at low frequencies in populations, and there is strong selection pressure for self-compatibility to evolve – minority cytotype exclusion (Levin, 1975; Husband, 2000). When they first form, allopolyploids are typically, for many characters and traits, intermediate between their two parents, and they are in instant competition if they occur sympatrically with their parents. They may also lack an ecological niche and/or experience low rates of pollination as a result of no specific adaptations to a pollinator. It is a combination of these problems that often causes neopolyploids to go extinct, but, as soon as a polyploid population forms, there will be selection for particular better adapted genotypes that direct the trajectory of subsequent genome evolution. This includes selection for genotypes with increased fertility, genomic stability and better-balanced gene copies. These and other (see below) directional changes are among the most important pressures that lead to diploidization of the neopolyploid genome.

A comparison of diversification rates across angiosperms has led to the suggestion that (neo)polyploids are more likely to go extinct and less likely to speciate than diploids (Mayrose *et al*., 2011), although there are sampling and analytical issues that make this a topic of much recent debate (see Soltis *et al*., 2014). There are also arguments for polyploids not contributing to adaptive radiations *per se*, but rather polyploids simply arising through their immediate reproductive isolation from parental lineages (purely as a result of differences in chromosome number). Where recurrent polyploidization occurs between the same or different parental species, (non-adaptive) radiations can result (Gorelick & Olson, 2013). Even if neopolyploids do have higher extinction rates and make a lower contribution to recent species diversification, as has been

argued, all angiosperm species nonetheless have (often multiple rounds of) polyploidy in their ancestry; the ramifications of this are significant and an important focus of research.

THE ECOLOGICAL COST OF POLYPLOIDY

With a larger genome comes the ecological burden of needing more macronutrients to build nucleic acids, particularly nitrogen and phosphorus, the latter limiting in many natural environments (Vitousek *et al*., 2010; Šmarda *et al*., 2013). In addition, it has been shown that interactions between plant genome size and macronutrient availability influence plant distribution in semi-natural field experiments (Šmarda *et al*., 2013; M. S. Guignard *et al*., unpubl. data). Given the general trend towards genome downsizing following polyploidy (Leitch & Bennett, 2004) and the strong skew towards small genome sizes in angiosperms despite recurrent polyploidy and in comparison with other land plant lineages (Leitch & Leitch, 2012), it is probable that there is selection favouring smaller genomes in angiosperms (but see also Oliver *et al*., 2007 for a neutral theory to explain the skew), thereby negating the effects of genome enlargement generated by polyploidy. However, the extent to which nitrogen and phosphorus availability influences genome size in the natural environment is a topic of debate, and further work is needed (Leitch & Leitch, 2008; Greilhuber & Leitch, 2013; Neiman *et al*., 2013).

DIPLOIDIZATION IS NECESSARY FOR EVOLUTIONARY PERSISTENCE AND DIVERSIFICATION

Diploidization of the genome post-polyploidization is associated with neofunctionalization, subfunctionalization and genome downsizing. Removal of extra DNA (often repetitive DNA) and extraneous gene copies occurs through recombination-based deletion and other mechanisms, whilst retaining duplicated genes, some of which may have new or altered functions. Selection can then act on individuals with varying genome sizes, and those with smaller genome sizes may be favoured, particularly perhaps in nutrient-poor environments. In addition, diploidization has been associated with chromosome number reduction, potentially involving complex chromosomal rearrangements (Franzke *et al*., 2011; Mándaková *et al*., 2012). Chromosome reorganization can be so extensive such that, in some taxa, e.g. *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae), which has had multiple polyploidy events in its ancestry (potentially the species is 48- or 96-ploid), the chromosome number has been reduced to a mere $n = 5$ pairs. Recent work in *Veronica* L. (Plantaginaceae) has suggested a link between increased diversification and genome downsizing following polyploidy (Meudt *et al*., 2015).

The WGD radiation lag-time model (Schranz, Mohammadin & Edger, 2012) postulates that species diversification often follows WGD events, but only after a 'lag phase' that can last up to several million years. This model explains the often observed pattern of a depauperate clade sister to a highly diverse one, with an observable time lag between the formation of polyploids and their subsequent diversification. Significant statistical support for this model has been garnered by Tank *et al*. (2015), who analysed nine well-documented ancient WGD events and demonstrated a non-random association between WGDs and a delayed increase in rates of diversification. It is likely that similar approaches will reveal more recent examples of a lag phase below family level (Tank *et al*., 2015). Schranz *et al*. (2012) tied the context for the lag to 'later migration events, changing environmental conditions and/or differential extinction rates'. Tank *et al*. (2015) suggested it could represent unsampled extinct lineages or the evolution of complex key traits/innovations; however, they emphasized that there is a real need to study the causes and nature of the lag phase in greater detail, in terms of both genomics and ecology. Our hypothesis here is that this lag phase is the time required for diploidization to take effect and provide a polyploid clade with the potential to radiate.

Polyploidy is important for the generation of genetic and genomic novelty, but it also requires extensive genome reorganization in order for this evolutionary potential to be fully realized (i.e. 'diploidization'). Over intermediate timescales, up to tens of millions of years, selection may favour smaller genomes that have an ecological advantage, at the same time favouring genotypes that retain advantageous alleles in enlarged gene families. Genomic rearrangements that occur after polyploidy may also enable novel *cis*-acting gene responses and the accumulation of locally adaptive genes in linkage groups (Yeaman, 2013). In addition, De Smet *et al*. (2013) documented consistent patterns of gene deletion in neopolyploid genomes, indicating that genes controlling expression and those in balanced macromolecular complexes were preferentially retained. This process of turnover takes time, and almost certainly leads to novel patterns of expression during the removal of extraneous gene copies.

CONCLUSIONS

It is clear from global analyses of chromosome number and genome size in a phylogenetic context that, despite the current frequency (and the important legacy) of polyploidization in angiosperms, there is also an irrefutable role for diploidization after polyploidization

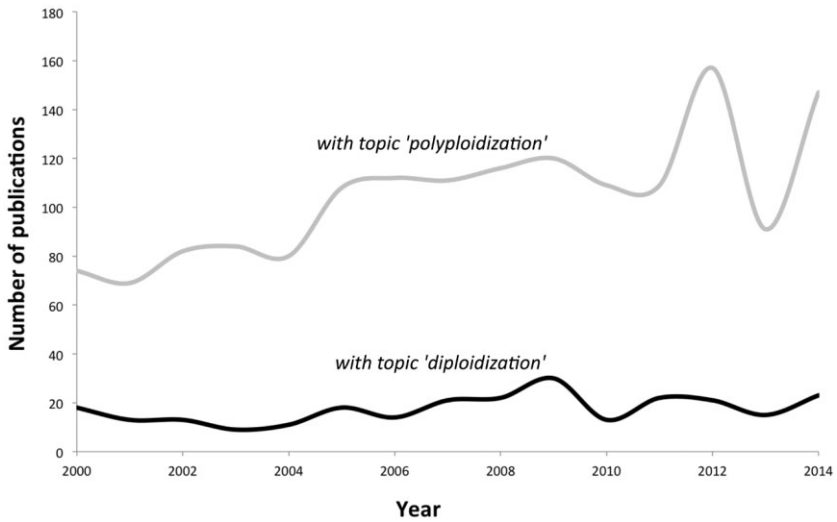

Figure 1. The number of publications with diploidization (black line) or polyploidization (grey line) in the title or keywords over the last 14 years. Source data: Web of Knowledge (Thomson Reuters).

has occurred. Our argument is that diploidization negates the disadvantages of polyploidy, rearranges genomes in novel ways and generates a higher level of genomic and transcriptomic variation upon which selection can act. Many genes return to their original copy number, thereby negating the effects of inefficient selection and the idea that polyploids are 'evolutionary dead-ends' (Stebbins, 1950). More sophisticated fine-tuning of expression and subfunctionalization can then enable novel phenotypic changes. The combination of high-throughput sequencing, cytogenetics and evolutionary developmental genetics with our best estimates of phylogenetic relationships will undoubtedly start to uncover the processes that have led to both ecological persistence and diversification of diploidized angiosperms. To test these hypotheses, we suggest physiological (stress) experiments on polyploids of different ages to examine potential 'genomic plasticity' enabled by the retention of increased numbers of transcription factors and genes controlling expression. Diploidization subsequent to polyploidization is an under-studied topic (see Fig. 1 for the number of publications on 'diploidization' versus 'polyploidization' in the last decade). Although the importance of diploidization is often acknowledged, botanists have never been in a better position to begin to answer exactly how post-polyploidization diploidization has contributed to the evolutionary success of the angiosperms.

ACKNOWLEDGEMENTS

This work was supported by a Natural Environmental Research Council (NERC) studentship to SD.

REFERENCES

Abraham A, Ninan CA. 1954. The chromosomes of *Ophioglossum reticulatum* L. *Current Science* **23:** 213–214.

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

Blanc G, Wolfe KH. 2004. Widespread palaeopolyploidy in model plant species inferred from age distributions of duplicate genes. *The Plant Cell* **16:** 1667–1678.

Bowers JE, Chapman BA, Rong JK, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422:** 433–438.

Chester M, Gallagher JP, Vaughan Symonds V, Veruska Cruz da Silva A, Mavrodiev EV, Leitch AR, Soltis PS, Soltis DE. 2012. Extensive chromosomal variation in a recently formed natural allopolyploid species, *Tragopogon miscellus*. *Proceedings of the National Academy of Sciences* **109:** 1176–1181.

Conant GC, Birchler JA, Pires CP. 2014. Dosage, duplication, and diploidization: clarifying the interplay of multiple models for duplicate gene evolution over time. *Current Opinion in Plant Biology* **19:** 91–98.

De Smet R, Adams KL, Vandepoele K, Van Montagu MCE, Maere S, Van de Peer Y. 2013. Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants. *Proceedings of the National Academy of Sciences* **110:** 2898–2903.

Flagel LE, Wendel JF. 2009. Gene duplication and evolutionary novelty in plants. *New Phytologist* **183:** 557–564.

Franzke A, Lysak MA, Al-Shehbaz IA, Koch MA, Mummenhoff K. 2011. Cabbage family affairs: the evolutionary history of Brassicaceae. *Trends in Plant Science* **16:** 108–116.

Gorelick R, Olson K. 2013. Polyploidy is genetic hence may cause non-adaptive radiations, whereas pseudopolyploidy is genomic hence may cause adaptive non-radiations. *Journal of Experimental Zoology. Part B. Molecular and Developmental Evolution* **5:** 286–294.

Greilhuber J, Leitch IJ. 2013. Genome size and the phenotype. In: Leitch IJ, Greilhuber J, Doležel J, Wendel J, eds. *Plant genome diversity, Vol. 2. Physical structure, behaviour and evolution of plant genomes*. London: Springer.

Husband BC. 2000. Constraints on polyploid evolution: a test of the minority cytotype exclusion principle. *Proceedings of the Royal Society B* **267:** 217–233.

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderball AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, de Pamphilis CW. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* **473:** 97–100.

Leitch IJ, Bennett MD. 2004. Genome downsizing in polyploid plants. *Biological Journal of the Linnean Society* **82:** 651–663.

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

Leitch AR, Leitch IJ. 2012. Ecological and genetic factors linked to contrasting genome dynamics in seed plants. *New Phytologist* **194:** 629–646.

Levin DA. 1975. Minority cytotype exclusion in local plant populations. *Taxon* **24:** 35–43.

Mándaková T, Mummenhoff K, Al-Shehbaz IA, Mucina L, Mühlhausen A, Lysak MA. 2012. Whole-genome triplication and species radiation in the southern African tribe Heliophileae (Brassicaceae). *Taxon* **61:** 989–1000.

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.

McClintock B. 1984. The significance of responses of the genome to challenge. *Science* **226:** 792–801.

Meudt HM, Rojas-Andrés BM, Prebble JM, Low E, Garnock-Jones PJ, Albach DC. 2015. Is genome downsizing associated with diversification in polyploid lineages of *Veronica*? *Botanical Journal of the Linnean Society* **178:** 243–266.

Mondragón-Palomino M. 2013. Perspectives on MADS-box expression during orchid flower evolution and development. *Frontiers in Plant Science* **4:** article 377.

Mondragón-Palomino M, Theißen G. 2011. Conserved differential expression of paralogous *DEFICIENS*- and *GLOBOSA*-like MADS-box genes in the flowers of Orchidaceae: refining the 'orchid code'. *Plant Journal* **66:** 1008–1019.

Neiman N, Theisen KM, Mayry ME, Kay AD. 2013. Can resource costs of polyploidy provide an advantage to sex? *Heredity* **110:** 152–159.

Oliver M, Petrov D, Ackerly D, Falkowski P, Schofield OM. 2007. The mode and tempo of genome size evolution in eukaryotes. *Genome Research* **17:** 594–601.

Otto S. 2007. The evolutionary consequences of polyploidy. *Cell* **131:** 452–462.

Pandit MK, Pocock MJO, Kunin WE. 2011. Ploidy influences rarity and invasiveness in plants. *Journal of Ecology* **99:** 1108–1115.

Petit M, Guidat C, Daniel J, Denis E, Montoriol E, Bui QT, Lim KY, Kovarik A, Leitch AR, Grandbastien M-A, Mhiri C. 2010. Mobilization of retrotransposons in synthetic allotetraploid tobacco. *New Phytologist* **186:** 135–147.

Schranz ME, Mohammadin S, Edger PE. 2012. Ancient whole genome duplications, novelty and diversification: the WGD radiation lag-time model. *Current Opinion in Plant Biology* **15:** 147–153.

Šmarda P, Hejcman M, Březinová A, Horová L, Steigerová H, Zedek F, Bureš P, Hejcmanová P, Schellberg J. 2013. Effect of phosphorus availability on the selection of species with different ploidy levels and genome sizes in a long-term grassland fertilization experiment. *New Phytologist* **200:** 911–921.

Soltis DE, Segovia-Salcedo MC, Jordon-Thaden I, Majure L, Miles NM, Mavrodiev EV, Mei W, Cortez MB, Soltis PS, Gitzendanner MA. 2014. Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose *et al*. (2011). *New Phytologist* **202:** 1105–1117.

Soltis PS, Soltis DE. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences* **97:** 7051–7057.

Stebbins GL. 1971. *Chromosomal evolution in higher plants*. London: Edward Arnold.

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

Tank DC, Eastman JM, Pennell MW, Soltis PS, Soltis DE, Hinchliff CE, Brown JW, Sessa EB, Harmon LJ. 2015. Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist* **207:** 454–467.

Van de Peer Y, Maere S, Meyer A. 2009. The evolutionary significance of ancient genome duplications. *Nature Reviews Genetics* **10:** 725–732.

Vitousek PM, Porder S, Houlton BZ, Chadwick OA. 2010. Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen–phosphorus interactions. *Ecological Applications* **20:** 5–15.

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences* **106:** 13 875–13 879.

Yant L, Hollister JD, Wright KM, Arnold BJ, Higgins JD. 2013. Meiotic adaptation to genome duplication in *Arabidopsis arenosa*. *Current Biology* **23:** 2151–2156.

Yeaman S. 2013. Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proceedings of the National Academy of Sciences* **110:** 1743–1751.