Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

The long and short of doubling down: polyploidy, epigenetics, and the temporal dynamics of genome fractionation

Jonathan F Wendel[1], Damon Lisch[2], Guanjing Hu[1] and Annaliese S Mason[3]

We consider the rapidly advancing discipline of plant evolutionary genomics, with a focus on the evolution of polyploid genomes. In many lineages, polyploidy is followed by 'biased fractionation', the unequal loss of genes from ancestral progenitor genomes. Mechanistically, it has been proposed that biased fractionation results from changes in the epigenetic landscape near genes, likely mediated by transposable elements. These epigenetic changes result in unequal gene expression between duplicates, establishing differential fitness that leads to biased gene loss with respect to ancestral genomes. We propose a unifying conceptual framework and a set of testable hypotheses based on this model, relating genome size, the proximity of transposable elements to genes, epigenetic reprogramming, chromatin accessibility, and gene expression.

Addresses
[1] Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011, United States
[2] Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, United States
[3] Department of Plant Breeding, IFZ Research Centre for Biosystems, Land Use and Nutrition, Justus Liebig University, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

Corresponding author: Wendel, Jonathan F (jfw@iastate.edu)

Current Opinion in Genetics & Development 2018, **49**:1–7

This review comes from a themed issue on **Genome architecture and expression**

Edited by **Kathleen Burns** and **Damon Lisch**

https://doi.org/10.1016/j.gde.2018.01.004

Introduction

The rapidly advancing discipline of plant evolutionary genomics is being propelled by a veritable explosion in genome sequencing efforts and allied technologies. This, in turn, is enabling exquisitely detailed insights into plant genome structure and function, from which a number of key generalizations about plant genome architecture have emerged [1•]. Among the most important are the following: firstly, that *all* plant lineages have evolutionary histories that include several to many whole-genome doubling (polyploidy) events, separated by thousands to millions of years, with each doubling episode superimposed on genomic remnants from earlier rounds of polyploidy; secondly, that most non-genic DNA in plant genomes consists of active, silenced, and dead-and-decaying transposable elements (TEs), varying widely in composition and copy numbers within and among populations, species, and higher levels of classification; thirdly, that plant genomes contain several different categories of small RNAs playing vital roles in genome organization, gene expression, and evolution; and fourthly, that chromatin contains a diverse suite of DNA and histone modifications that interact in myriad ways to generate a remarkably complex epigenomic and chromatin landscape [2], which in turn plays a role in specifying phenotypes and hence evolutionary trajectories.

These four principle generalizations are rooted in overlapping but different disciplines, ranging from taxonomy to plant physiology to molecular genetics, and were motivated by different biological questions and perspectives. Yet it has recently become clear that all four principal realizations are intimately intertwined. In our view, this interconnection offers a conceptually unifying lens through which to view numerous topics relevant to the architecture of modern plant genomes. Emblematic of this are the many questions surrounding the temporal dynamics and mechanistic forces governing the evolution of polyploid genomes, a subject of active interest [3,4•,5,6].

Prevalence of polyploidy in plants

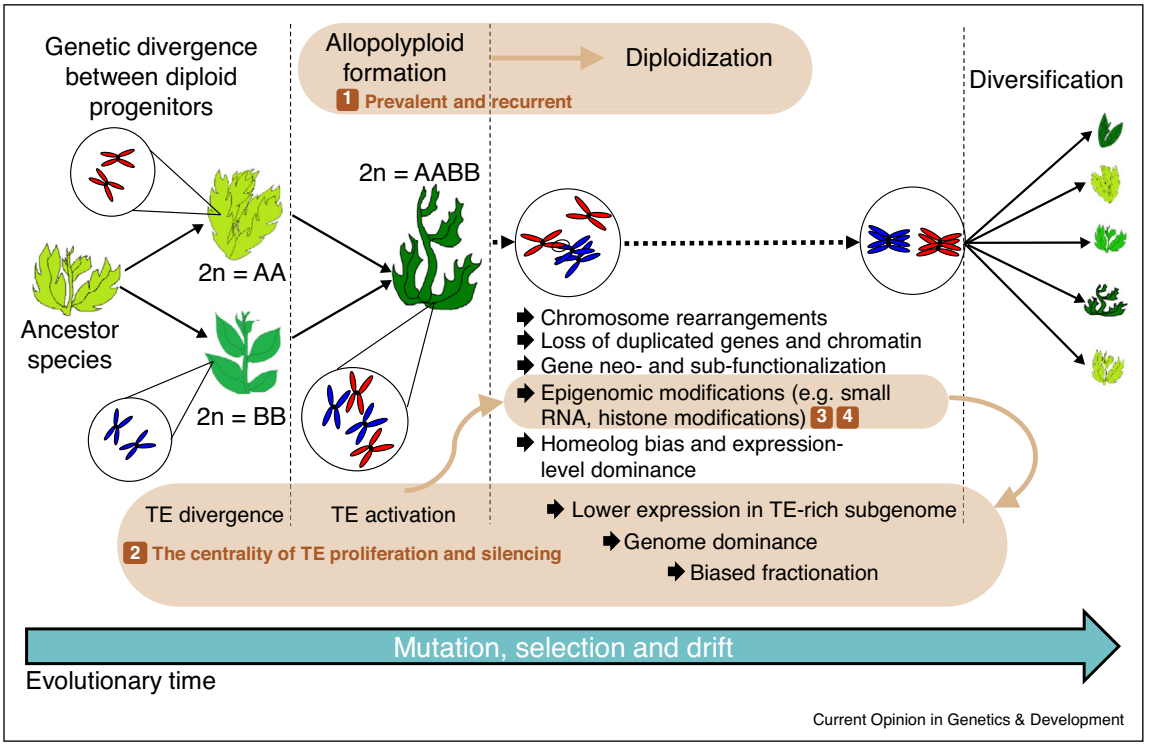
Early genome sequencing efforts revealed unexpectedly high genomic redundancy in many modern 'diploid' genomes that could only be accounted for by postulating one or more cryptic (not evident from chromosome numbers or behavior) polyploidization events. Thus, in addition to commonly observed and classical 'neoallopolyploids' (chromosomally evident polyploids), such as cotton, wheat, and sugarcane, it became clear that all plants are 'paleopolyploid', that is, derived from ancient polyploidy events [3,7]. Genomic investigations of both ancient and modern polyploids revealed

immediate as well as longer-term genomic and transcriptomic consequences resulting from the merger of two genomes (Figure 1).

Immediate and short-term consequences of polyploidy were found to include gene silencing and loss of 'redundant' genes and sequences, chomosome exchanges leading to doubling or loss of chromosome fragments or sequences, and massive, genome-wide transcriptomic rewiring. The latter encompasses a diversity of phenomena, including expression sub-functionalization or neo-functionalization, biased expression of homoeologs (genes duplicated by polyploidy) [8,9], expression level dominance [10], and mobilization of previously dormant transposable elements [11–13]. Simultaneously, it was established that longer term, polyploidy is followed by massive loss of redundant sequences, such that a pseudodiploid-like state is restored. Strikingly, nearly all genes retained in duplicate in paleopolyploids have become subfunctionalized or neofunctionalized [14].

Surprisingly, it turns out that genome downsizing following polyploidy is non-random with respect to both genes [15,16] and genomes. Different classes of genes (e.g. transcription factors) are preferentially retained following polyploidy, whereas other categories of genes (e.g. those involved in organellar processes and meiosis) are more likely to be restored to singleton status [15,16]. Critically, not only is duplicate *gene* loss non-random, but so is loss of sequences from duplicate *genomes*: this biased loss of ancestral genomes following allopolyploidy (resulting from the merger of two dissimilar genomes, typically between species) is termed 'biased fractionation' [17], though the expression is not strictly binary, but rather quantitative. Biased fractionation of genomes may be closely connected to the expression-level observation of genome-wide homeolog expression bias [8,9], the latter frequently referred to as 'genome dominance' [8]. This was first demonstrated in maize [17], but subsequently has been found in many polyploids (Table 1). The model systems studied to date include four eudicot

Figure 1



Genome dominance and biased fractionation in polyploids, illustrating the connections between several key realizations about plant genome architecture. Allopolyploidy is both prevalent and recurrent in angiosperms ('1', top), involving progenitors (AA and BB, left) that may vary in genome size, their complements of transposable elements ('2', bottom), and other features. Common genomic responses to allopolyploidy (middle) over the short and long term include massive loss of duplicated genes and chromatin (collectively, 'fractionation', or 'diploidization', which may be biased or unbiased with respect to the progenitor genomes), chromosome rearrangements, genic neo-functionalization and sub-functionalization, and a host of expression level phenomena including homoeolog bias and expression level dominance. Higher TE load or TE adjacency to genes in one of the two progenitors may lead to epigenetically (small RNA silencing machinery, '3'; DNA and histone modifications, '4') mediated repression of expression of one set of homoeologs (known as 'genome dominance', bottom), creating differential fitness between homoeologs. This is envisioned to interact with evolutionary processes (bottom) during allopolyploid diversification (right).

Table 1

Polyploid systems for which biased fractionation has been studied, arranged in order of antiquity of the polyploidization event(s)

Taxon or taxa	Approximate age (MY)	Biased fractionation?	Genome dominance?[a]	Reference
Arabidopsis suecica	0.02	Yes	Yes	[47,48]
Capsella bursa-pastoris	0.2	No	No	[39]
Zea mays	8	Yes	Yes	[17,49]
Glycine max	13	No	No	[18]
Cucurbita spp.	3–26	No	No	[41]
Brassica rapa	15	Yes	Yes	[15,42••]
Arabidopsis thaliana	47	Yes	Yes	[18,50,51]
Medicago sativa	58	Yes	na	[18]
Gossypium spp.	60	Yes	Yes	[43]
Musa acuminata	65	No	No	[18,52]
Populus trichocarpa	65	No	No	[18]
Poaceae	70	Yes	Yes	[18]

[a] Defined as genome-wide homeolog expression bias.

(Brassicaceae, Fabaceae, Malvaceae, Salicaceae) and two monocot (Musaceae, Poaceae), families, so the phenomenon appears widespread. Remarkably, biased fractionation can be detected in very young (*Arabidopsis suecica*) as well as ancient (*Medicago*, *Gossypium*) polyploids.

The centrality of transposable elements

Importantly, biased fractionation does not always appear to follow polyploidy events [18], raising the question as to why it is so striking in some polyploids and absent in others. Hints at answers to this question, as well as the mechanistic underpinnings of biased fractionation, have emerged from recent advances in our understanding of the centrality of TEs to genome size and genic evolution [19•,20–22]. Comparative genomic analyses have demonstrated that flowering plant evolution is punctuated by repeated bouts of proliferation of different families of various kinds of Class I (retrotransposons, or 'copy and paste') and Class II (DNA transposons, or 'cut and paste') TEs (Figure 1). These saltational bursts are followed by silencing and decay of most of the dozens to thousands of the newly inserted elements, as a consequence of the twin mechanisms of small RNA-mediated epigenetic repression and longer-term mutational decay and deletion.

As with episodes of polyploidy, bursts of TE mobilization are temporally erratic and are not predictable, although in some cases hybridization and polyploidy may stimulate bursts of TE activity [11–13], as may stress [23,24]. Plant genome sizes thus largely reflect both historical polyploidy events and TE-related genome expansion and shrinkage (via the slower mutational processes involved in TE decay and deletion). Exactly how these TE-related genome evolution processes are molded by natural selection and drift is unclear. Yet, the net consequence, played out on macroevolutionary scale, is a 2400-fold variation in plant genome size among angiosperms [25], with some lineages exhibiting relative stasis, whereas others exhibit dramatic variation in TE content among closely related species or even populations [1•].

TE insertions also provide a potent mutagenic mechanism for the evolution of new genes or functions [19•,20,26]. Examples abound of the phenotypic consequences and evolutionary significance of individual TE insertions in generating novel phenotypes, including famous ones such the Ac/Ds family insertion into a starch synthase gene that created Mendel's wrinkled peas [27] and the hopscotch insertion (*tb1*) that helped shape the architecture of the modern maize plant [28]. TEs often provide the raw material from which novel regulatory sequences are derived [20,29], with individual bursts potentially seeding the genome with new sequences that may be evolutionarily modified to function as promoters or enhancers, altered transcription start sites, or which lead to local spreading or loss of heterochromatin [30•,31]. TEs in many lineages have been co-opted by host genomes to function as new genes in a diversity of physiological and developmental processes, including hormone signaling, stress responses, disease resistance, response to light, and flowering time [22,26]. Insertion of a TE or a group of TEs that have become epigenetically silenced can also impact expression of nearby genes, with a generally repressive, quantitative effect on expression [31–33]. When integrated across all TEs in the genomes, one can readily imagine that the spatial relationships between TE insertions and genes are thus an important force shaping gene coexpression networks, their downstream metabolic and physiological outputs, and hence phenotypes.

Given the rapidity with which the TE complement among diploid species may diverge, it is unsurprising that the merger of two diploids, via hybridization and/or allopolyploidy, has novel evolutionary consequences with respect to TEs [19•]. First, genome merger may lead to bursts of TE mobilization, as noted above, presumably due to parental mismatches in 'genome surveillance' mechanisms; 'defects' in the small RNA silencing machinery may cause ephemeral episodes of insertional mutagenesis and possibly chromosomal rearrangements

in nascent polyploid lineages. Second, extensive genomic variation may be introduced into nascent polyploid lineages due to the insertional mutagenesis accompanying TE mobilization. It seems likely that, when compared to the diploid condition, polyploid genomes will better tolerate insertions due to their genomic redundancy, with a full spectrum of expected effects ranging from loss of duplicate genes to novel expression domains and/or function.

A third consequence of polyploidy related to TEs is more subtle, and, remarkably, appears to be intimately connected to the observation of biased fractionation in present-day genomes resulting from ancient polyploidy events (Table 1). As recently summarized [34•], two key observations are proposed to be mechanistically interconnected and largely responsible for the differential loss of ancestral genomes. First, for retained duplicate genes, homoeologs in the more highly fractionated genome are often expressed at lower levels than their counterparts in the more intact genome [17,35] ('Genome dominance', in Table 1). Secondly, epigenetically silenced TEs often are physically closer to the homoeolog with lower expression, suggestive of a position-effect repression of gene transcription [32,33,36••]. These two observations suggest a plausible scenario of how we 'get from here to there', that is, from a progenitor situation involving two intact genomes at the time of polyploid formation, to a derivative many millions of years later in which one genome has been preferentially retained (Figure 1). Specifically, mutations, either wholly or partly deleterious, should be less consequential when they arise in the homoeolog with lower rather than higher expression, at least when both homoeologs are functionally equivalent. This sets up a fitness differential, leading to preferential loss of homoeologs with lower expression, which, if played out genome-wide and differentially between the two ancestral genomes, might generate 'genome dominance' at the expression level and biased loss of duplicate genes at the DNA level. Alternatively, differential retention of sequences may be envisioned to arise from expression-level genome dominance, establishing a feedback loop between genome dominance and fractionation.

This hypothesis for the genesis of biased fractionation is appealing in that it accounts for a diversity of otherwise seemingly disconnected observations. Additionally, it focuses attention on the initial conditions at the time of hybridization and polyploid formation. Thus, the 'genomic legacy' of the diploids [34•,37], would seem to figure prominently in predicting the future genomic evolution of an allopolyploid, and indeed, of past polyploidizations. One can imagine, for example, the case where two ancient and highly divergent genomes become reunited as a consequence of polyploidization, in which one genome had a much greater number of TEs and or TEs that were closer to genes; this would seem to be a perfect setup for a future in which there is strong biased fractionation. Perhaps illustrative of these connections, the recent analysis of a 140 year old allohexaploid of *Mimulus peregrinus* [38•] revealed correlations between subgenome homoeolog expression bias and methylation levels, possibly showing future biased fractionation being 'caught in the act'. If, on the other hand, the two progenitors are little diverged in TE number and distribution, we might not expect genome loss to be particularly biased. Garsmeur *et al.* [18] hypothesized that the observation, or not, of biased fractionation in modern descendants of ancient polyploidies is indicative of whether the event entailed allo-polyploidy or auto-polyploidy. A similar explanation has also recently been invoked for the differences observed in fractionation between maize and soybeans [36••]. While reasonable, biased fractionation may also not occur in allopolyploidy events if the two diploid species are similar in their spectrum of TE-mediated position effects. This might be the case, for example, in *Capsella bursa-pastoris* [39,40] and allotetraploid *Cucurbita* [41] species, which display neither biased fractionation nor genome dominance (Table 1), and for which the progenitor diploid genome sizes are similar.

A role for epigenetics

Closely connected to the hypotheses regarding the genesis of biased fractionation are insights from discoveries in small RNA biology and epigenetics. The merger of two diverged suites of TEs and their silencing small RNAs can lead to quantitative and qualitative mismatches between progenitor silencing machineries, as noted above and as described and illustrated in Springer *et al.* [30•] and Wendel *et al.* [1•]. These regulatory mismatches may lead to various perturbations in the small RNA populations produced in hybrid and polyploid tissues, including production of novel small RNAs, with corresponding effects on gene expression. Notably, in both *Brassica* [35,42••] and *Gossypium* [43] paleopolyploids (modern diploids), the density of small RNAs targeting TEs is higher in regions adjacent to homoeologs that exhibit lower expression levels, consistent with their heterochromatizing effects and repression of expression of nearby genes. Similar conclusions were reached in a comparison of small RNA targeting regions in the subgenomes of maize versus soybeans [36••]. These results complement the large body of data demonstrating that polyploidization typically induces epigenetic reprogramming, most often detected as altered distributions of methylated cytosines [19•,30•,31] but also in other epigenetic marks, such as H3K4 trimethylation in wheat [44] and MNase sensitivity in maize [45••]. In the latter study, it was demonstrated that the less fractionated genome is more likely to have an open chromatin configuration upstream of retained homoeologs than the more fractionated genome, indicative of a higher level of transcription.

Concluding remarks

The topic of genome evolution in polyploids has stimulated a convergence of some of the most important realizations of the genomics era. It was entirely unexpected that genome evolution on the scale of millions of years would be traceable, and potentially mechanistically explained, by structural differences between modern coresident genomes combined with an understanding of TE behavior and epigenetics. Importantly though, differences in gene expression between homoeologs are not the only mechanism leading to genome fractionation in polyploids, biased or otherwise. Selective sweeps can lead to preferential loss of genes and surrounding genic regions, with chromosome rearrangements extending selective sweeps to large stretches of genes. Genetic drift can also ameliorate bias by random fixation of fractionation patterns in small populations (or small effective populations, like self-pollinators [34•]). Biased fractionation, presently conceived as being mediated by chromatin-induced differential repression of homoeolog expression, may be constrained or overwhelmed by strong selection operating at other levels, such as gene balance to retain stoichiometric equilibrium in multiprotein complexes [46], or by the selective fixation of advantageous mutations in otherwise 'less fit' genomes. In addition, present intragenomic TE distributions may also be a consequence, rather than a cause, of biased fractionation, as TE insertions are more likely to be tolerated near the less 'fit' homoeologs. It is clear, then, that 'genomic dominance' represents only a partial explanation for biased fractionation.

Our understanding of biased fractionation is far from complete, and still resides in the realm of *ex post facto* description rather than experimental hypothesis testing. Although much in evolutionary biology faces this conundrum, the stage is now set for focused experiments that examine a range of related, synthetic allopolyploids that vary in their degree of parental divergence in key features such as TE quantity and distribution, and heterochromatinizing small RNA populations. Such experiments are feasible in many genera, for example, in *Aegilops*, *Oryza* and *Gossypium*, among others, and would lead to testable predictions about the relationships between genome size and TE adjacency to genes, epigenetic reprogramming, chromatin accessibility, and gene expression. Further, depending on the ease of obtaining mutants in these species pairs, it may be possible to perform hybridizations in the context of mutations in silencing pathways. In addition, natural systems involving young polyploids (e.g. in *Tragopogon*, *Spartina*, *Mimulus*, *Cardamine*, *Senecio*) may prove informative, particular with respect to the interplay between natural selection and genomic processes associated with polyploid genome evolution.

Finally, at present there is virtually no information on the functional and hence potential adaptive significance of biased fractionation. The only study connecting the phenomenon to phenotype is from maize, where it was recently demonstrated that higher-expressing paralogues contribute disproportionately to phenotypic variation and diversity [45••]. Comparable studies in other systems, integrated over many different polyploids of varying ages, will likely enhance our understanding of the evolutionary significance of biased fractionation, particularly if they are combined with 'omics' interrogations of advanced populations, such as experimental lines containing alternatively fractionated states.

Conflict of interest statement

Nothing declared.

Acknowledgements

The authors acknowledge support from the National Science Foundation (MCB1118646, DBI160466, DBI0817707 to JFW, and DBI-0820828 to DL), Cotton Incorporated (JFW), and the Deutsche Forschungsgemeinschaft (to ASM, MA6473/1-1; MA6473/2-1).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. • Wendel JF, Jackson SA, Meyers BC, Wing RA: **Evolution of plant genome architecture**. *Genome Biol* 2016, **17**:37.
A concise introduction to several of the key realizations about plant genomes that emerged during the genomics era.

2. Vergara Z, Gutierrez C: **Emerging roles of chromatin in the maintenance of genome organization and function in plants**. *Genome Biol* 2017, **18**:96.

3. Wendel JF: **The wondrous cycles of polyploidy in plants**. *Am J Bot* 2015, **102**:1753-1756.

4. • Van de Peer Y, Mizrachi E, Marchal K: **The evolutionary significance of polyploidy**. *Nat Rev Genet* 2017, **18**:411-424.
A highly readable introduction to many different aspects of polyploid evolution.

5. Soltis PS, Soltis DE: **Ancient WGD events as drivers of key innovations in angiosperms**. *Curr Opin Plant Biol* 2016, **30**: 159-165.

6. Soltis DE, Visger CJ, Marchant DB, Soltis PS: **Polyploidy: pitfalls and paths to a paradigm**. *Am J Bot* 2016, **103**:1146-1166.

7. Ruprecht C, Lohaus R, Vanneste K, Mutwil M, Nikoloski Z, Van de Peer Y, Persson S: **Revisiting ancestral polyploidy in plants**. *Sci Adv* 2017, **3**:e1603195.

8. Yoo M-J, Liu X, Pires JC, Soltis PS, Soltis DE: **Nonadditive gene expression in polyploids**. *Annu Rev Genet* 2014, **48**:485-517.

9. Flagel L, Udall J, Nettleton D, Wendel J: **Duplicate gene expression in allopolyploid *Gossypium* reveals two temporally distinct phases of expression evolution**. *BMC Biol* 2008, **6**:16.

10. Grover C, Gallagher J, Szadkowski E, Yoo M, Flagel L, Wendel J: **Homoeolog expression bias and expression level dominance in allopolyploids**. *New Phytol* 2012, **196**:966-971.

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

12. Kawakami T, Strakosh S, Zhen Y, Ungerer M: **Different scales of *Ty1/copia*-like retrotransposon proliferation in the genomes of three diploid hybrid sunflower species**. *Heredity* 2010, **104**:341.

13. Senerchia N, Felber F, Parisod C: **Genome reorganization in F1 hybrids uncovers the role of retrotransposons in reproductive isolation**. *Proc R Soc Lond B: Biol Sci* 2015, **282**:20142874.

14. Renny-Byfield S, Gallagher JP, Grover CE, Szadkowski E, Page JT, Udall JA, Wang X, Paterson AH, Wendel JF: **Ancient gene duplicates in *Gossypium* (cotton) exhibit near-complete expression divergence**. *Genome Biol Evol* 2014, **6**:559-571.

15. Mandáková T, Li Z, Barker MS, Lysak MA: **Diverse genome organization following 13 independent mesopolyploid events in Brassicaceae contrasts with convergent patterns of gene retention**. *Plant J* 2017, **91**:3-21.

16. De Smet R, Adams KL, Vandepoele K, Van Montagu MC, Maere S, Van de Peer Y: **Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants**. *Proc Natl Acad Sci U S A* 2013, **110**:2898-2903.

17. Schnable JC, Springer NM, Freeling M: **Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss**. *Proc Natl Acad Sci U S A* 2011, **108**:4069-4074.

18. Garsmeur O, Schnable JC, Almeida A, Jourda C, D'Hont A, Freeling M: **Two evolutionarily distinct classes of paleopolyploidy**. *Mol Biol Evol* 2013, **31**:448-454.

19. • Vicient CM, Casacuberta JM: **Impact of transposable elements on polyploid plant genomes**. *Ann Bot* 2017, **120**:195-207.
A current review of TEs in plant genomes, with a focus on the novel possibilities for TEs to effect the evolution of polyploids.

20. Bennetzen JL, Wang H: **The contributions of transposable elements to the structure, function, and evolution of plant genomes**. *Annu Rev Plant Biol* 2014, **65**:505-530.

21. Lisch D: **How important are transposons for plant evolution?** *Nat Rev Genet* 2013, **14**:49-61.

22. Kim N-S: **The genomes and transposable elements in plants: are they friends or foes?** *Genes Genomics* 2017:1-12.

23. Ito H, Kim J-M, Matsunaga W, Saze H, Matsui A, Endo TA, Harukawa Y, Takagi H, Yaegashi H, Masuta Y: **A stress-activated transposon in Arabidopsis induces transgenerational abscisic acid insensitivity**. *Sci Rep* 2016, **6** srep23181.

24. Grandbastien M-A: **LTR retrotransposons, handy hitchhikers of plant regulation and stress response**. *Biochim Biophys Acta (BBA)* 2015, **1849**:403-416.

25. Pellicer J, Fay MF, Leitch IJ: **The largest eukaryotic genome of them all?** *Bot J Linnean Soc* 2010, **164**:10-15.

26. Zhao D, Ferguson AA, Jiang N: **What makes up plant genomes: the vanishing line between transposable elements and genes**. *Biochim Biophys Acta (BBA)* 2016, **1859**:366-380.

27. Bhattacharyya MK, Smith AM, Ellis TN, Hedley C, Martin C: **The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme**. *Cell* 1990, **60**:115-122.

28. Studer A, Zhao Q, Ross-Ibarra J, Doebley J: **Identification of a functional transposon insertion in the maize domestication gene tb1**. *Nat Genet* 2011, **43**:1160-1163.

29. Lisch DR: **Transposons in plant gene regulation**. *Plant Transpos Genome Dyn Evol* 2013:93-116.

30. • Springer NM, Lisch D, Li Q: **Creating order from chaos: epigenome dynamics in plants with complex genomes**. *Plant Cell* 2016, **28**:314-325.
A highly readable synthesis of the state of our knowledge of plant epigenomics, in both diploids and polyploids.

31. Diez CM, Roessler K, Gaut BS: **Epigenetics and plant genome evolution**. *Curr Opin Plant Biol* 2014, **18**:1-8.

32. Hollister JD, Gaut BS: **Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression**. *Genome Res* 2009, **19**:1419-1428.

33. Hollister JD, Smith LM, Guo Y-L, Ott F, Weigel D, Gaut BS: **Transposable elements and small RNAs contribute to gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata***. *Proc Natl Acad Sci U S A* 2011, **108**:2322-2327.

34. • Steige KA, Slotte T: **Genomic legacies of the progenitors and the evolutionary consequences of allopolyploidy**. *Curr Opin Plant Biol* 2016, **30**:88-93.
A concise review of biased fractionation in polyploids, with an emphasis on the significance of initial conditions, that is, the progenitor genomes.

35. Woodhouse MR, Cheng F, Pires JC, Lisch D, Freeling M, Wang X: **Origin, inheritance, and gene regulatory consequences of genome dominance in polyploids**. *Proc Natl Acad Sci U S A* 2014, **111**:5283-5288.

36. •• Zhao M, Zhang B, Lisch D, Ma J: **Patterns and consequences of subgenome differentiation provide insights into the nature of paleopolyploidy in plants**. *Plant Cell* 2017 http://dx.doi.org/10.1105/tpc.17.00595.
This paper demonstrates the importance of initial conditions in influencing subsequent biased fractionation and genomic features such as TEs, siRNAs, and DNA methylation, using maize and soybeans as examples.

37. Buggs RJ, Wendel JF, Doyle JJ, Soltis DE, Soltis PS, Coate JE: **The legacy of diploid progenitors in allopolyploid gene expression patterns**. *Philos Trans R Soc B* 2014, **369**:20130354.

38. • Edgar PP, Smith RD, McKain MR, Cooley AM, Vallejo-Marin M, Yuan Y-W, Bewick AJ, Ji L, Platts AE, Bowman MJ *et al.*: **Subgenome dominance in an interspecific hybrid, synthetic allopolyploid, and a 140-year-old naturally established neo-allopolyploid monkeyflower**. *Plant Cell* 2017.
An excellent example of the use of recently evolved allopolyploids to study the mechanisms of biased homoeolog expression, here demonstrating a correlation between TE methylation and expression bias.

39. Douglas GM, Gos G, Steige KA, Salcedo A, Holm K, Josephs EB, Arunkumar R, Ågren JA, Hazzouri KM, Wang W: **Hybrid origins and the earliest stages of diploidization in the highly successful recent polyploid *Capsella bursa*-pastoris**. *Proc Natl Acad Sci U S A* 2015, **112**:2806-2811.

40. Hurka H, Friesen N, German DA, Franzke A, Neuffer B: **'Missing link' species *Capsella orientalis* and *Capsella thracica* elucidate evolution of model plant genus Capsella (Brassicaceae)**. *Mol Ecol* 2012, **21**:1223-1238.

41. Sun H, Wu S, Zhang G, Jiao C, Guo S, Ren Y, Zhang J, Zhang H, Gong G, Jia Z: **Karyotype stability and unbiased fractionation in the paleo-allotetraploid Cucurbita genomes**. *bioRxiv* 2017:150755.

42. •• Cheng F, Sun C, Wu J, Schnable J, Woodhouse MR, Liang J, Cai C, Freeling M, Wang X: **Epigenetic regulation of subgenome dominance following whole genome triplication in *Brassica rapa***. *New Phytol* 2016, **211**:288-299.
The authors demonstrate a connection between gene expression, transposable element (TE) distribution and small RNA targeting, illustrating the synergy between genome history and epigenetics.

43. Renny-Byfield S, Gong L, Gallagher JP, Wendel JF: **Persistence of subgenomes in paleopolyploid cotton after 60 my of evolution**. *Mol Biol Evol* 2015, **32**:1063-1071.

44. Zhang W, Fan X, Gao Y, Liu L, Sun L, Su Q, Han J, Zhang N, Cui F, Ji J: **Chromatin modification contributes to the expression divergence of three *TaGS2* homoeologs in hexaploid wheat**. *Sci Rep* 2017:7.

45. •• Renny-Byfield S, Rodgers-Melnick E, Ross-Ibarra J: **Gene fractionation and function in the ancient subgenomes of maize**. *Mol Biol Evol* 2017. msx121.
This study is the first to demonstrate that biased fractionation may be connected to phenotype, here showing that higher-expressing paralogues in maize contribute disproportionately to phenotypic variation and diversity.

46. Birchler JA, Veitia RA: **Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines**. *Proc Natl Acad Sci U S A* 2012, **109**:14746-14753.

47. Chang PL, Dilkes BP, McMahon M, Comai L, Nuzhdin SV: **Homoeolog-specific retention and use in allotetraploid *Arabidopsis suecica* depends on parent of origin and network partners**. *Genome Biol* 2010, **11**:R125.

48. Novikova PY, Tsuchimatsu T, Simon S, Nizhynska V, Voronin V, Burns R, Fedorenko OM, Holm S, Säll T, Prat E: **Genome sequencing reveals the origin of the allotetraploid *Arabidopsis suecica***. *Mol Biol Evol* 2017, **34**:957-968.

49. Swigoňová Z, Lai J, Ma J, Ramakrishna W, Llaca V, Bennetzen JL, Messing J: **Close split of sorghum and maize genome progenitors**. *Genome Res* 2004, **14**:1916-1923.

50. Thomas BC, Pedersen B, Freeling M: **Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes**. *Genome Res* 2006, **16**:934-946.

51. Hohmann N, Wolf EM, Lysak MA, Koch MA: **A time-calibrated road map of Brassicaceae species radiation and evolutionary history**. *Plant Cell* 2015, **27**:2770-2784.

52. D'hont A, Denoeud F, Aury J-M, Baurens F-C, Carreel F, Garsmeur O, Noel B, Bocs S, Droc G, Rouard M: **The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants**. *Nature* 2012, **488**:213.