

## Review Article

# Recent Insights into Mechanisms of Genome Size Change in Plants

**Corrinne E. Grover and Jonathan F. Wendel**

*Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50011, USA*

Correspondence should be addressed to Corrinne E. Grover, corrinne@iastate.edu

Received 4 February 2010; Accepted 19 March 2010

Academic Editor: Ilia Judith Leitch

Copyright © 2010 C. E. Grover and J. F. Wendel. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Genome sizes vary considerably across all eukaryotes and even among closely related species. The genesis and evolutionary dynamics of that variation have generated considerable interest, as have the patterns of variation themselves. Here we review recent developments in our understanding of genome size evolution in plants, drawing attention to the higher order processes that can influence the mechanisms generating changing genome size.

## 1. Introduction

It has long been known that tremendous variation in DNA content exists, even within closely related species, and that organismal complexity is poorly correlated with genome size. In plants, genome size ranges from 63 Mbp in *Genlisea margaretae* to 124,852 Mbp in *Fritillaria assyriaca* [1], a 2000-fold difference. This diversity of genome size has generated considerable interest in the nature of sequence variation among genomes, the mechanisms that operate to add and/or remove DNA, and the suite of internal and external evolutionary forces that collectively shape or control the molecular drivers. Key insights into each of these arenas have emerged in the recent years from the explosion of genomic sequence data. Here we provide a synopsis of this burgeoning field, focusing on the recent developments that have improved our understanding of the processes that underlie genome size change in plants.

## 2. Heterogeneity in Genome Size Fluctuations

After the initial puzzling observation that greater perceived complexity does not equate to a larger genome [2], it was quickly realized that the non-protein-encoding fraction of the genome comprised sequence types that actually correlated with genome size [3]. Repetitive elements, it was

learned, and primarily LTR-retrotransposons in the case of plant genomes, could achieve surprisingly high copy numbers, by themselves accounting for half or more of the genome for species having “large” genomes [4]. Many subsequent studies have demonstrated that the fraction and composition of the genome occupied by these sequences reflects, mechanistically, the antagonistic effects of insertion, due primarily to transposable element (TE) proliferation, and deletion, primarily mediated by unequal intrastrand homologous recombination (i.e., recombination between directly repeated sequences, such as the LTRs of single or adjacent retrotransposons) and illegitimate recombination (i.e., *RecA* independent recombination capable of deleting sequence intervening regions of microhomology) [5–7]. To generate the extraordinary range in extant angiosperm genome sizes, it has been reasoned, the magnitude of these mechanisms must also vary among species.

That this is true has now been confirmed in multiple studies involving taxa distributed widely among angiosperms. Most notably, it has become apparent that differential proliferation of TEs explains the majority of genome size differences among species. In the wild rice species *Oryza australiensis*, for example, amplification of three TEs accounts for a 2-fold increase in genome size [8]. Similarly, evidence from *Gossypium* (cotton) [9] indicates that the majority of the threefold range in diploid genome

sizes may be accounted for by amplification of the *Gorge3* *gypsy* element in the larger genomes (but see below).

One important dimension of the foregoing, and other studies, is the realization that TE proliferation is not a constant feature of plant genomes nor of any specific lineage, but that instead it is an episodic or saltational process throughout the angiosperms. Bursts of transposition have been inferred from TE dating analyses in many species [7, 8, 10, 11]. These and other studies also have revealed that different types of TEs (e.g., *cop* or *gypsy* LTR retrotransposons) or different subfamilies of a single type (e.g., *cop* LTR retrotransposons) may episodically proliferate at different times. The result is that lineages experience periodic quantum gains in genome size that are likely controlled by myriad factors (e.g., epigenetics, recombination rate, etc.), which vary by element type/family and, presumably, in response to genomic and environmental factors (e.g., hybridization or environmental stress).

In addition to heterogeneity in the amount of historical TE proliferation, deletional mechanisms also have been demonstrated to vary in importance among species. The first such survey of the relative importance of unequal intrastrand homologous recombination (UR) versus illegitimate recombination (IR) was conducted in *Arabidopsis thaliana*, where it was concluded that IR has had a larger impact than UR, removing more than fivefold DNA [12]. In rice, however, unequal homologous recombination has been more efficient at purging extraneous DNA (3.3 Mbp for UR versus 2.8 Mbp for IR) [13]. Since those two initial studies, the relative effectiveness of UR versus IR has been evaluated for different species [14–17] and the subject of their effectiveness relative to one another and relative to TE proliferation has been a topic of debate [5, 7]. That is to say, since IR often results in substantially smaller deletions than UR, the relative effectiveness of IR versus UR has been questioned, as has the ability of either to reverse, or even slow, genome size growth in the face of massive transposable element proliferation, both when considered individually or together. While it is clear UR has the ability to more rapidly remove large amounts of DNA, IR has a broader scope of action (i.e., it is not reliant upon sequence homologies, such as LTRs, to operate); therefore, the relative impact of each mechanism will vary as the number of potential sites for UR diminishes. New evidence (discussed below) also elucidates the effect genomic properties can have on both deletional mechanisms, which further underscores the impact that additional data will have in increasing our understanding of these mechanisms.

The picture that has emerged from an increased understanding of TE proliferation as well as deletional processes is one which surmises that extant genome sizes reflect the often oppositional processes of genomic expansion and contraction. Thus, for example, recently inserted TE DNA often is rapidly removed, potentially leading to rapid genomic turnover [18, 19] and even within species variation [20–22]. A comprehensive study of LTR-retrotransposons in rice showed that in addition to several bursts of transposition experienced during the last 5 million years, the rice genome also experienced a flurry of deletions, ultimately leading to

removal of over half of the inserted LTR-retrotransposon DNA [11]. Similarly, Hawkins et al. demonstrated lineage-specific, differential removal of the TE most responsible for the threefold variation in genome size among *Gossypium* diploids, that is, the *gypsy*-like *Gorge3* element [23]. By phylogenetically partitioning *Gorge3* elements into time points representing: (1) pre-*Gossypium* amplification, (2) *Gossypium*-specific amplification, and (3) lineage-specific amplification, and utilizing a novel modeling approach, the authors were able to reconstruct the ancestral copy number for *Gorge3* and infer gains and losses for each lineage. A key conclusion is that the smaller genomes are not only *not* gaining *Gorge3* as quickly as the larger genomes, but they are also more effective in *removing* elements, and at a rate that actually exceeds the rate of gain.

This demonstration that genomic contraction via TE removal can actually exceed the rush toward genomic obesity [24] implied by bursts of TE proliferation is mirrored in an additional study in cotton using a phylogenetically informed approach to polarize small indels. Indels in two genomic regions were catalogued and characterized for five genomes among species whose phylogenetic relationships were well-established, thus providing the opportunity to interpret small indels as losses or gains of DNA sequence [16]. Differences in the *rates* of sequence gain and loss were demonstrated among terminal branches and between ancestor and descendent, demonstrating that temporal heterogeneity characterizes multiple mechanisms of genome size evolution. Overall, the trend for the diploid genomes, both extant and ancestral, was toward growth, albeit slowly in some cases; however, the polyploid experienced growth of one subgenome ( $D_T$ ) and contraction of the other ( $A_T$ ), resulting in a net loss for the polyploid genome. While many deletions contributed to overall contraction, the majority of sequence loss was attributable to the removal of a single *gypsy* element, once again underscoring the potential for rapid removal via UR.

Finally, it has become evident that insertion and deletion operate heterogeneously with respect to genomic location. This is clear from studies using from FISH [25, 26], sequencing of genomic regions [17, 27–31], and whole genome sequencing projects [32–35]. This unevenness is pronounced for the LTR-retrotransposon *gypsy* superfamily, whose members often experience a significant bias toward residing in genomic locations that are considered more heterochromatic in nature and most often are associated with pericentromeric or centromeric regions [32, 33, 35–38]. Other locational biases, however, have also been noted (e.g., euchromatic regions for *cop* LTR-retrotransposons and UTR/exonic regions for SINEs in maize [34, 35], generic regions for maize *Mutator* and Helitron elements, [35, 39, 40], and introns for LINEs in maize and soybean [29, 41]). These observed biases likely reflect myriad factors, including but not limited to insertional preferences, disruptive potential for insertion in a given region, local rates of recombination (discussed below), and lineage-specific effects. The influence of genomic location on deletional mechanisms has been evaluated less; however, as UR acts largely on LTR-retrotransposons and both UR and IR are

recombination based, it is not hard to imagine the relevance of genomic location to these processes.

### 3. Epigenetics and Genome Size

While the principal mechanisms responsible for genome size expansion and contraction appear to be relatively clear, the factors that stimulate or control each mechanism remain enigmatic. Because of the heterogeneity in the operation of insertional and deletional mechanisms with respect to genomic region, lineage, and time, it stands to reason that this heterogeneity reflects multiple interacting external environmental forces as well as intrinsic genomic properties. One that has garnered much attention in recent years is epigenetic regulation of transposable elements.

Epigenetic regulation of transposable elements is considered to be the first line of defense against uncontrolled TE proliferation. Methylation and heterochromatization of TEs as a means to limit proliferation is not a recent idea, with observations in the Mutator system of maize representing some of the earlier research into epigenetic regulation of TEs [42–46]. More recently, the pathways by which TEs are silenced have been illuminated, specifically the dependence upon RNAi to silence transcription and remodel TE-containing chromatin (reviewed in [47–49]). Recent evidence from the maize genome highlights and furthers some of the advances made in understanding these regulatory processes. In the maize collection, Jia et al. evaluated the consequences of a deficiency in RNA-dependent RNA polymerase 2 (AtRDR2; *mop1* in maize), a component of the RNA-directed DNA methylation silencing pathway [50]. Previous gene expression analyses suggested that even though *mop1* is expressed >100-fold higher in maize shoot apical meristems, the expression of some retrotransposons is substantially higher as well (when compared to seedlings) [51]. In this new analysis, the authors surveyed the expression changes in the *mop1* mutant for 797 DNA TE families and 608 retrotransposons to find that while most of the DNA TE families behaved as expected in a plant that is deficient in this silencing pathway (i.e., 78% of expression changes in DNA TE families was toward increased expression in the mutant), the retrotransposons behaved counter to expectations in that the expression changes observed were most often (68%) toward decreased expression [50]. In addition to changes in TE expression, many gene expression changes were also observed, most notably genes involved in chromatin modifications. Several histone deacetyltransferases, which have been implicated in heterochromatin formation in yeast [52], experienced increased expression in the *mop1* mutant indicating that those families experiencing decreased expression may be responding to increased heterochromatin formation in the mutant. The salient conclusion here is that not only are there multiple pathways and processes by which TEs are silenced, but these processes can interact and sometimes in an *antagonistic* fashion, which may permit relaxed control of some TE families and stricter control of other types. This highlights the complexities

involved in TE regulation and impels a need for further exploration.

The recognition that TE silencing is RNAi based provides a reasonable explanation for the cessation of the transpositional bursts that shape genome size growth; as the copy number of a TE type increases, siRNAs derived from the new copies increase in number, which subsequently enhance silencing of that element type. Bursts of transposition are often considered to provide a temporary release from this or related forms of epigenetic suppression, a release suggested to be initiated by environmental stress [53–56] or an organismal process such as interspecific hybridization. In rice, for example, introgression has been linked to retrotransposon activation which was subsequently shut-down via cytosine methylation [57], while in sunflower, three independent hybridizations between the same two parents led to rapid proliferation of the same *gypsy* element in all three hybrid species [58, 59]. These and other studies lead to the widely held conceptualization that TEs are ever-present, typically “well-behaved” genomic residents being held in check through epigenetic suppression or by flying under the radar as low-copy elements, but which occasionally are “set loose” in different lineages and times in response to internal and external stresses that are not fully understood.

Recent work in *Arabidopsis* hints at a possible explanation for these periodic releases from suppression, while also underscoring the role that epigenetics plays with respect to TE deletion [60]. Based on the realization that methylated sequences often affect the expression of neighboring genes [61–63], the authors hypothesize that genes near to methylated TEs (met-TEs) would experience lower expression. In addition, if there exists a cost for methylating TEs that insert near genes, met-TEs would be subject to purifying selection and be more quickly removed from gene-rich regions. By analyzing the methylation status of TEs in the *Arabidopsis* genome, it was shown that TE methylation affects gene expression in a 1.5–2 kb window surrounding the gene. These results might help explain the differential deletion rates that exist for TE families, by suggesting that the insertion preferences and propensity for methylation characteristic of each family may influence the amount of negative selective pressure experienced by different types of elements. A second implication is that periodic releases from suppression may be a consequence of increased expression of a met-TE suppressed gene, achieved by TE demethylation under conditions of stress. Insight into the influence of epigenetic processes provided by this and similar studies on the various mechanisms that add to or eliminate DNA from genomes represent a key area for future research into genome size evolution.

### 4. Genetic Recombination and Genome Size

Because deletional mechanisms such as unequal intrastrand homologous recombination (UR) and illegitimate recombination (IR) are both recombination based, it stands to reason that local rates of genetic recombination might impact genome size evolution. Previous studies in rice [13],

*Arabidopsis* [12], and other species [15] have evaluated global rates of UR and IR; however, the effects of local rates of genetic recombination have been less frequently evaluated and results have been inconsistent. In *Arabidopsis lyrata*, a global bias exists between regions of differing recombination rates (i.e. TEs are more abundant in the gene-poor pericentromeric regions than in the more genic chromosome arms), but overall recombination rate does not correlate with TE abundance (when excluding pericentromeric-specific TEs) in intergenic space [64]. The authors suggest that the main factor influencing the differential association of TEs with pericentromeric and/or gene-poor regions is due to the short distances between genes in *A. lyrata* and the disruptive factor of TE insertions. A recent study in rice, however, suggests that the rate of genetic recombination can influence the rate of TE removal [65]. The authors evaluated the distribution and structural variation of LTR-retrotransposons (full-length, UR- or IR-deletion types) in the rice genome and related these to genomic features, including local rates of genetic recombination and gene density. Both the local rate of genetic recombination and gene density were negatively correlated with TE density, indicating that more TEs were allowed to accumulate in gene-poor regions of low recombination. In addition, UR recombination had the greatest effect in regions of high genetic recombination, whereas IR was most active in regions of low genetic recombination. Combined with their observation that UR is able to more quickly remove DNA than IR, the genomic balance and proportion of regions experiencing high or low rates of recombination may partly explain the differences each of these mechanisms have had historically in shaping genome size. That is, in species with relatively large areas of high recombination, UR will likely be the more active mechanism and will be responsible for rapidly removing DNA, whereas in species with large areas of low recombination, IR will be relatively more active and will be responsible for removing DNA at a slower rate. Furthermore, as recombination rates are fairly labile in plants [66], some of the differences observed in the rates of UR and IR between genomic regions and closely related species may be due to the average rate of genetic recombination for that region or species [16].

Surprisingly, the mechanisms of deletion are not the only ones associated with recombination. In an interesting discovery, Liu et al. uncovered a strong correlation between *Mutator* element genomic markers indicative of open chromatin, which is also associated with increased recombination. When examining *Mu* insertion site preferences, the authors found a nonrandom distribution that was similar to the patterns observed for recombination and gene density; that is, the rate of *Mu* insertions and recombination and gene density tend to be highest near the more euchromatic chromosome ends and then decrease as the distance to the centromere decreases [40]. Neither gene density nor a previously described preference for intragene insertion [67, 68] could adequately describe the pattern observed; however, upon comparison with existing cytosine methylation and histone modification data [69, 70], a strong association began to emerge. Both *Mu* insertions and recombination largely favored regions with

strong signals for H3K4me3 and H3K9ac and with low levels of cytosine methylation, all suggestive of open chromatin structure. The authors suggest that other TEs that display biases toward genic regions (e.g., Ac/Ds or MITEs) may also rely on open chromatin structure for insertion. Thus, while the distribution of *Mu* insertions does not rely on recombination rate, per se, recombination rate itself may be indicative of regions susceptible to open chromatin targeting TEs. As the aforementioned lability in recombination rates may mirror a similar lability in chromatin structure, the success of these types of elements may be influenced by differences among species. Clearly more data are needed to gain a clearer perspective on the effect of recombination and chromatin structure on TE success and persistence.

## 5. Population Genetics and Genome Size

Discussions concerning genome size evolution often center on the mechanisms, events or rates of change that historically have influenced the genome size of a given lineage or a set of taxa, often using one individual per species. Population-level processes such as effective population size and breeding system may also contribute to the shaping of genomes, though at present few studies address this relationship. Because most TE insertions that survive are neutral to slightly deleterious, they are subject to the twin processes of selection and drift, and thus relative levels of TE survival are contingent not just on their internal, genomic ecology but also on external, population level forces. Much of the empirical evidence (from mostly bacterial and animal systems) and relevant theory is presented in Lynch's recent book *The Origins of Genome Architecture* [71]. This relatively new area of research is now attracting interest from plant biologists.

Illustrative of this approach, Hollister and Gaut detail the influence of population dynamics on the retention of a Helitron element, *Basho*, in the genome of the selfing plant, *Arabidopsis thaliana*. Using a subset (278 of 565) of the *Basho* elements predicted in the *A. thaliana* genome, they screened a diverse panel of 47 accessions to determine the frequency of Helitron occupation at each insertion site and related these to genomic factors such as element length, proximity to genes, and recombination rate. A high rate of fixation was detected, as compared to an analogous study in *Drosophila* [72, 73], with nearly 50% of the evaluated elements achieving fixation and 81% existing in over half of the accessions surveyed, consistent with the notion that genetic drift in this inbred plant permitted the fixation of presumably slightly deleterious sequences. The authors also found that the age of an element is strongly and positively correlated with fixation, as expected for neutral alleles, whereas length and proximity to genes is negatively correlated with fixation. Thus, selection against element accumulation is likely weak and dependent upon the size of the element, due to the potentially deleterious effects of a greater potential for ectopic recombination in longer sequences. The authors suggest further that ectopic recombination is important in governing the persistence of *Basho* in the *A. thaliana* genome



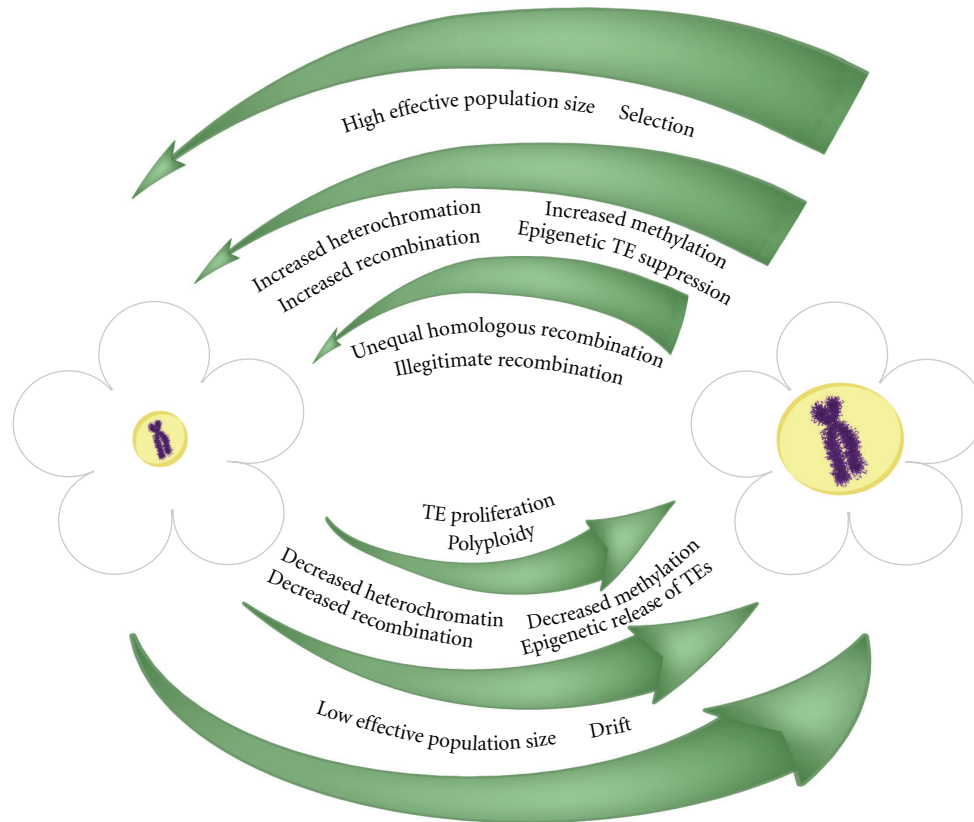


FIGURE 1: The effects that various mechanisms, genomic properties, and population influences exert on genome size. The innermost ring represents proximal mechanisms responsible for generating genome size differences, while the second and third rings represent genomic properties and population size effects, respectively, which affect the magnitude of those mechanisms.

and suggest that the weak selection observed in *A. thaliana* may be higher in an outcrossing relative (e.g., *A. lyrata*) where the heterozygous state of many TEs would provide more potential for ectopic pairings.

In a similar investigation, Lockton et al. used transposon display for six diverse families of TEs (*Gypsy*-like LTR-retrotransposon; SINE-like and LINE-like non-LTR-retroelements; Ac-like, CACTA-like, and *Tourist*-like DNA elements) to generate polymorphism datasets for five populations of *Arabidopsis lyrata* that had previously been described demographically [74, 75]. Interestingly, individual TE bands were found in intermediate to high population frequencies, suggesting that the selection has not strongly been operated to remove these TEs. Because mean TE “allele” frequencies are lower in *A. lyrata* than in *A. thaliana* (24% versus 60% across TE families), the data are consistent with the expectation that TEs in an outcrossing species experience stronger selection and are less subject to drift. An interesting added dimension of this analysis was the calculation of selection coefficients for the individual populations (located in Sweden, Iceland, Russia, the United States, and Canada, and a larger refugial population located in Germany). Most of the populations, save the refugial German population, had positive estimates of selection, counter to the expectation of purifying (negative) selection against TEs. They inferred that in these populations, all of which had significantly

lower effective population sizes than the larger German population (by 7–18 fold), drift is able to overcome the weak selection against TE insertions experienced in *A. lyrata* (as estimated using the larger German population). One important implication is their suggestion, based on these data, that the rate of genomic flux in TEs is influenced not only by current effective population size and breeding system, but also by demographic history. Thus, small population sizes, inbreeding, and population bottlenecks are all conditions that lead to a less effective environment for purging TEs, and hence genome sizes might be predicted to expand more rapidly than comparable plant populations not experiencing these conditions (all else being equal).

## 6. The Future of Genome Size

Our understanding of the mechanisms responsible for genome size evolution has vastly improved over the past decade, with a number of reviews devoted to the patterns exhibited by these mechanisms among a variety of species [5, 6, 76, 77]. Less discussed and only more recently addressed are the multiple factors that influence insertional and deletional processes as well as their context-dependent interactions. Thus, for example, we have only begun to explore how genomic properties, such as recombination and

epigenetic context, and population level processes, including effective size, history, and breeding system, affect short-term and longer-term genome size evolution (Figure 1). As next-generation sequencing technologies become more accessible and increasingly applied, it is now possible to design studies that will enhance our understanding of these many interactions. Incorporating analyses across several phylogenetic scales will yield insights into the forces that shape modern plant genomes and help explain their current diversity and distinctions.

## References

- [1] J. Greilhuber, T. Borsch, K. Müller, A. Worberg, S. Porembski, and W. Barthlott, "Smallest angiosperm genomes found in Lentibulariaceae, with chromosomes of bacterial size," *Plant Biology*, vol. 8, no. 6, pp. 770–777, 2006.
- [2] A. E. Mirsky and H. Ris, "The desoxyribonucleic acid content of animal cells and its evolutionary significance," *The Journal of General Physiology*, vol. 34, no. 4, pp. 451–462, 1951.
- [3] W. Y. Chooi, "Comparison of the DNA of six vicia species by the method of DNA-DNA hybridization," *Genetics*, vol. 68, no. 2, pp. 213–230, 1971.
- [4] J. L. Bennetzen, "Transposable element contributions to plant gene and genome evolution," *Plant Molecular Biology*, vol. 42, no. 1, pp. 251–269, 2000.
- [5] J. L. Bennetzen, J. Ma, and K. M. Devos, "Mechanisms of recent genome size variation in flowering plants," *Annals of Botany*, vol. 95, no. 1, pp. 127–132, 2005.
- [6] C. E. Grover, J. S. Hawkins, and J. F. Wendel, "Phylogenetic insights into the pace and pattern of plant genome size evolution," in *Plant Genomes*, J.-N. Volff, Ed., Karger, Basel, Switzerland, 2008.
- [7] J. S. Hawkins, G. Hu, R. A. Rapp, J. L. Grafenberg, and J. F. Wendel, "Phylogenetic determination of the pace of transposable element proliferation in plants: copia and LINE-like elements in *Gossypium*," *Genome*, vol. 51, no. 1, pp. 11–18, 2008.
- [8] B. Piegu, R. Guyot, N. Picault, et al., "Doubling genome size without polyploidization: dynamics of retrotransposon-driven genomic expansions in *Oryza australiensis*, a wild relative of rice," *Genome Research*, vol. 16, no. 10, pp. 1262–1269, 2006.
- [9] J. S. Hawkins, H. Kim, J. D. Nason, R. A. Wing, and J. F. Wendel, "Differential lineage-specific amplification of transposable elements is responsible for genome size variation in *Gossypium*," *Genome Research*, vol. 16, no. 10, pp. 1252–1261, 2006.
- [10] P. Neumann, A. Koblov, A. Navrátilová, and J. Macas, "Significant expansion of *Vicia pannonica* genome size mediated by amplification of a single type of giant retroelement," *Genetics*, vol. 173, no. 2, pp. 1047–1056, 2006.
- [11] C. Vitte, O. Panaud, and H. Quesneville, "LTR retrotransposons in rice (*Oryza sativa*, L.): recent burst amplifications followed by rapid DNA loss," *BMC Genomics*, vol. 8, article 218, 2007.
- [12] K. M. Devos, J. K. M. Brown, and J. L. Bennetzen, "Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*," *Genome Research*, vol. 12, no. 7, pp. 1075–1079, 2002.
- [13] J. Ma, K. M. Devos, and J. L. Bennetzen, "Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice," *Genome Research*, vol. 14, no. 5, pp. 860–869, 2004.
- [14] N. Chantret, J. Salse, F. Sabot, et al., "Molecular basis of evolutionary events that shaped the *hardness* locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*)," *Plant Cell*, vol. 17, no. 4, pp. 1033–1045, 2005.
- [15] C. Vitte and J. L. Bennetzen, "Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 47, pp. 17638–17643, 2006.
- [16] C. E. Grover, Y. Yu, R. A. Wing, A. H. Paterson, and J. F. Wendel, "A phylogenetic analysis of indel dynamics in the cotton genus," *Molecular Biology and Evolution*, vol. 25, no. 7, pp. 1415–1428, 2008.
- [17] T. Wicker, N. Yahiaoui, R. Guyot, et al., "Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat," *Plant Cell*, vol. 15, no. 5, pp. 1186–1197, 2003.
- [18] C. E. Grover, J. S. Hawkins, and J. F. Wendel, "Tobacco genomes quickly go up in smoke," *New Phytologist*, vol. 175, no. 4, pp. 599–602, 2007.
- [19] M. Petit, K. Y. Lim, E. Julio, et al., "Differential impact of retrotransposon populations on the genome of allotetraploid tobacco (*Nicotiana tabacum*)," *Molecular Genetics and Genomics*, vol. 278, no. 1, pp. 1–15, 2007.
- [20] S. Brunner, K. Fengler, M. Morgante, S. Tingey, and A. Rafalski, "Evolution of DNA sequence nonhomologies among maize inbreds," *Plant Cell*, vol. 17, no. 2, pp. 343–360, 2005.
- [21] H. Fu and H. K. Dooner, "Intraspecific violation of genetic colinearity and its implications in maize," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 14, pp. 9573–9578, 2002.
- [22] Q. Wang and H. K. Dooner, "Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 47, pp. 17644–17649, 2006.
- [23] J. S. Hawkins, S. R. Proulx, R. A. Rapp, and J. F. Wendel, "Rapid DNA loss as a counterbalance to genome expansion through retrotransposon proliferation in plants," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 42, pp. 17811–17816, 2009.
- [24] J. L. Bennetzen and E. A. Kellogg, "Do plants have a one-way ticket to genomic obesity?" *Plant Cell*, vol. 9, no. 9, pp. 1509–1514, 1997.
- [25] S. E. Staton, M. C. Ungerer, and R. C. Moore, "The genomic organization of *Ty3/gypsy*-like retrotransposons in *Helianthus* (Asteraceae) homoploid hybrid species," *American Journal of Botany*, vol. 96, no. 9, pp. 1646–1655, 2009.
- [26] R. E. Hanson, M. N. Islam-Faridi, C. F. Crane, et al., "Ty1-copia-retrotransposon behavior in a polyploid cotton," *Chromosome Research*, vol. 8, no. 1, pp. 73–76, 2000.
- [27] C. E. Grover, H. R. Kim, R. A. Wing, A. H. Paterson, and J. F. Wendel, "Incongruent patterns of local and global genome size evolution in cotton," *Genome Research*, vol. 14, no. 8, pp. 1474–1482, 2004.
- [28] C. E. Grover, H. Kim, R. A. Wing, A. H. Paterson, and J. F. Wendel, "Microcolinearity and genome evolution in the *AdhA* region of diploid and polyploid cotton (*Gossypium*)," *Plant Journal*, vol. 50, no. 6, pp. 995–1006, 2007.

- [29] B. Joseph, et al., "Retrotransposons within syntenic regions between soybean and medicago truncatula and their contribution to local genome evolution," *The Plant Genome*, vol. 2, no. 3, pp. 211–223, 2009.
- [30] T. Wicker, N. Stein, L. Albar, C. Feuillet, E. Schlagenhauf, and B. Keller, "Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution," *Plant Journal*, vol. 26, no. 3, pp. 307–316, 2001.
- [31] R. Bruggmann, A. K. Bharti, H. Gundlach, et al., "Uneven chromosome contraction and expansion in the maize genome," *Genome Research*, vol. 16, no. 10, pp. 1241–1251, 2006.
- [32] T. Sasaki, "The map-based sequence of the rice genome," *Nature*, vol. 436, no. 7052, pp. 793–800, 2005.
- [33] S. Kaul, H. L. Koo, J. Jenkins, et al., "Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*," *Nature*, vol. 408, no. 6814, pp. 796–815, 2000.
- [34] R. S. Baucom, J. C. Estill, C. Chaparro, et al., "Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome," *PLoS Genetics*, vol. 5, no. 11, Article ID e1000732, 2009.
- [35] P. S. Schnable, D. Ware, R. S. Fulton, et al., "The B73 maize genome: complexity, diversity, and dynamics," *Science*, vol. 326, no. 5956, pp. 1112–1115, 2009.
- [36] J. Jiang, J. A. Birchler, W. A. Parrott, and R. K. Dawe, "A molecular view of plant centromeres," *Trends in Plant Science*, vol. 8, no. 12, pp. 570–575, 2003.
- [37] J.-Y. Lin, B. H. Jacobus, P. SanMiguel, et al., "Pericentromeric regions of soybean (*Glycine max* L. Merr.) chromosomes consist of retroelements and tandemly repeated DNA and are structurally and evolutionarily labile," *Genetics*, vol. 170, no. 3, pp. 1221–1230, 2005.
- [38] B. D. Peterson-Burch, D. Nettleton, and D. F. Voytas, "Genomic neighborhoods for *Arabidopsis* retrotransposons: a role for targeted integration in the distribution of the Metaviridae," *Genome Biology*, vol. 5, no. 10, article R78, 2004.
- [39] L. Yang and J. L. Bennetzen, "Distribution, diversity, evolution, and survival of Helitrons in the maize genome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 47, pp. 19922–19927, 2010.
- [40] S. Liu, C.-T. Yeh, T. Ji, et al., "Mu transposon insertion sites and meiotic recombination events co-localize with epigenetic marks for open chromatin across the maize genome," *PLoS Genetics*, vol. 5, no. 11, Article ID e1000733, 2009.
- [41] A. P. Tikhonov, P. J. SanMiguel, Y. Nakajima, N. M. Gorenstein, J. L. Bennetzen, and Z. Avramova, "Colinearity and its exceptions in orthologous adh regions of maize and sorghum," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 13, pp. 7409–7414, 1999.
- [42] J. L. Bennetzen, "The regulation of Mutator function and Mu1 transposition," in *Plant Genetics*, M. Freeling, Ed., pp. 343–353, Alan R. Liss, New York, NY, USA, 1985.
- [43] J. L. Bennetzen, "The mutator transposable element system of maize," *Current Topics in Microbiology and Immunology*, vol. 204, pp. 195–229, 1996.
- [44] J. L. Bennetzen, K. Schrick, P. S. Springer, W. E. Brown, and P. SanMiguel, "Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA," *Genome*, vol. 37, no. 4, pp. 565–576, 1994.
- [45] V. Walbot, "Alterations in the Mutator transposable element family of *Zea mays*," in *Plant Genetics*, M. Freeling, Ed., pp. 333–342, Alan R. Liss, New York, NY, USA, 1985.
- [46] V. Walbot, "Reactivation of Mutator transposable elements of maize by ultraviolet light," *Molecular and General Genetics*, vol. 234, no. 3, pp. 353–360, 1992.
- [47] R. K. Slotkin and R. Martienssen, "Transposable elements and the epigenetic regulation of the genome," *Nature Reviews Genetics*, vol. 8, no. 4, pp. 272–285, 2007.
- [48] C. Weil and R. Martienssen, "Epigenetic interactions between transposons and genes: lessons from plants," *Current Opinion in Genetics and Development*, vol. 18, no. 2, pp. 188–192, 2008.
- [49] D. Lisch, "Epigenetic regulation of transposable elements in plants," *Annual Review of Plant Biology*, vol. 60, pp. 43–66, 2009.
- [50] Y. Jia, D. R. Lisch, K. Ohtsu, M. J. Scanlon, D. Nettleton, and P. S. Schnable, "Loss of RNA-dependent RNA polymerase 2 (RDR2) function causes widespread and unexpected changes in the expression of transposons, genes, and 24-nt small RNAs," *PLoS Genetics*, vol. 5, no. 11, Article ID e1000737, 2009.
- [51] K. Ohtsu, M. B. Smith, S. J. Emrich, et al., "Global gene expression analysis of the shoot apical meristem of maize (*Zea mays* L.)," *Plant Journal*, vol. 52, no. 3, pp. 391–404, 2007.
- [52] T. Yamada, W. Fischle, T. Sugiyama, C. D. Allis, and S. I. S. Grewal, "The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast," *Molecular Cell*, vol. 20, no. 2, pp. 173–185, 2005.
- [53] S. Pouteau, E. Huttner, M. A. Grandbastien, and M. Caboche, "Specific expression of the tobacco Tnt1 retrotransposon in protoplasts," *The EMBO Journal*, vol. 10, no. 7, pp. 1911–1918, 1991.
- [54] S. Pouteau, M. A. Grandbastien, and M. Boccar, "Microbial elicitors of plant defense responses activate transcription of a retrotransposon," *Plant Journal*, vol. 5, no. 4, pp. 535–542, 1994.
- [55] M.-A. Grandbastien, "Activation of plant retrotransposons under stress conditions," *Trends in Plant Science*, vol. 3, no. 5, pp. 181–187, 1998.
- [56] R. Kalendar, J. Tanskanen, S. Immonen, E. Nevo, and A. H. Schulman, "Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 12, pp. 6603–6607, 2000.
- [57] B. Liu and J. F. Wendel, "Retrotransposon activation followed by rapid repression in introgressed rice plants," *Genome*, vol. 43, no. 5, pp. 874–880, 2000.
- [58] M. C. Ungerer, S. C. Strakosh, and Y. Zhen, "Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation," *Current Biology*, vol. 16, no. 20, pp. R872–R873, 2006.
- [59] M. C. Ungerer, S. C. Strakosh, and K. M. Stimpson, "Proliferation of Ty3/gypsy-like retrotransposons in hybrid sunflower taxa inferred from phylogenetic data," *BMC Biology*, vol. 7, article 40, 2009.
- [60] J. D. Hollister and B. S. Gaut, "Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression," *Genome Research*, vol. 19, no. 8, pp. 1419–1428, 2009.
- [61] D. Jahner and R. Jaenisch, "Retrovirus-induced de novo methylation of flanking host sequences correlates with gene inactivity," *Nature*, vol. 315, no. 6020, pp. 594–597, 1985.
- [62] Z. Lippman, A.-V. Gendrel, M. Black, et al., "Role of transposable elements in heterochromatin and epigenetic control," *Nature*, vol. 430, no. 6998, pp. 471–476, 2004.

- [63] X. Zhang, S. Shiu, A. Cal, and J. O. Borevitz, "Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays," *PLoS Genetics*, vol. 4, no. 3, Article ID e1000032, 2008.
- [64] S. I. Wright, B. Lauga, and D. Charlesworth, "Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata*," *Molecular Ecology*, vol. 12, no. 5, pp. 1247–1263, 2003.
- [65] Z. Tian, C. Rizzon, J. Du, et al., "Do genetic recombination and gene density shape the pattern of DNA elimination in rice long terminal repeat retrotransposons?" *Genome Research*, vol. 19, no. 12, pp. 2221–2230, 2009.
- [66] J. Ross-Ibarra, "Genome size and recombination in angiosperms: a second look," *Journal of Evolutionary Biology*, vol. 20, no. 2, pp. 800–806, 2007.
- [67] A. D. Cresse, S. H. Hulbert, W. E. Brown, J. R. Lucas, and J. L. Bennetzen, "Mu1-related transposable elements of maize preferentially insert into low copy number DNA," *Genetics*, vol. 140, no. 1, pp. 315–324, 1995.
- [68] A. M. Settles, S. Latshaw, and D. R. McCarty, "Molecular analysis of high-copy insertion sites in maize," *Nucleic Acids Research*, vol. 32, no. 6, article e54, 2004.
- [69] L. E. Palmer, P. D. Rabinowicz, A. L. O'Shaughnessy, et al., "Maize genome sequencing by methylation filtration," *Science*, vol. 302, no. 5653, pp. 2115–2117, 2003.
- [70] X. Wang, A. A. Elling, X. Li, et al., "Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize," *Plant Cell*, vol. 21, no. 4, pp. 1053–1069, 2009.
- [71] M. Lynch, *The Origins of Genome Architecture*, vol. 16, Sinauer Associates, Sunderland, Mass, USA, 2007.
- [72] X. Maside, S. Assimacopoulos, and B. Charlesworth, "Fixation of transposable elements in the *Drosophila melanogaster* genome," *Genetical Research*, vol. 85, no. 3, pp. 195–203, 2005.
- [73] D. A. Petrov, Y. T. Aminetzach, J. C. Davis, D. Bensasson, and A. E. Hirsh, "Size matters: non-LTR retrotransposable elements and ectopic recombination in *drosophila*," *Molecular Biology and Evolution*, vol. 20, no. 6, pp. 880–892, 2003.
- [74] J. Ross-Ibarra, S. I. Wright, J. P. Foxe, et al., "Patterns of polymorphism and demographic history in natural populations of *Arabidopsis lyrata*," *PLoS ONE*, vol. 3, no. 6, article e2411, 2008.
- [75] S. Lockton, J. Ross-Ibarra, and B. S. Gaut, "Demography and weak selection drives patterns of transposable element diversity in natural populations of *Arabidopsis lyrata*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 37, pp. 13965–13970, 2008.
- [76] T. R. Gregory, "The C-value enigma in plants and animals: a review of parallels and an appeal for partnership," *Annals of Botany*, vol. 95, no. 1, pp. 133–146, 2005.
- [77] J. S. Hawkins, C. E. Grover, and J. F. Wendel, "Repeated big bangs and the expanding universe: directionality in plant genome size evolution," *Plant Science*, vol. 174, no. 6, pp. 557–562, 2008.



