Genome Size Reduction through Illegitimate Recombination Counteracts Genome Expansion in Arabidopsis

Katrien M. Devos, James K.M. Brown and Jeffrey L. Bennetzen

References	This article cites 27 articles, 16 of which can be accessed free at: **http://www.genome.org/cgi/content/full/12/7/1075#References** Article cited in: **http://www.genome.org/cgi/content/full/12/7/1075#otherarticles**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Notes

To subscribe to *Genome Research* go to:
http://www.genome.org/subscriptions/

Letter

Genome Size Reduction through Illegitimate Recombination Counteracts Genome Expansion in *Arabidopsis*

Katrien M. Devos,[1,2,3] James K.M. Brown,[1] and Jeffrey L. Bennetzen[2]

[1]*John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, United Kingdom;* [2]*Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907-1392, USA*

Genome size varies greatly across angiosperms. It is well documented that, in addition to polyploidization, retrotransposon amplification has been a major cause of genome expansion. The lack of evidence for counterbalancing mechanisms that curtail unlimited genome growth has made many of us wonder whether angiosperms have a "one-way ticket to genomic obesity." We have therefore investigated an angiosperm with a well-characterized and notably small genome, *Arabidopsis thaliana*, for evidence of genomic DNA loss. Our results indicate that illegitimate recombination is the driving force behind genome size decrease in *Arabidopsis*, removing at least fivefold more DNA than unequal homologous recombination. The presence of highly degraded retroelements also suggests that retrotransposon amplification has not been confined to the last 4 million years, as is indicated by the dating of intact retroelements.

Flowering plants (angiosperms) vary enormously in genome size, from <50 Mb in some members of the Cruciferae to >85,000 Mb in some Liliaceae (Bennett and Leitch 1995). The mechanisms that account for dramatic expansion of angiosperm genomes have been documented, primarily polyploidization and retrotransposon amplification (SanMiguel et al. 1996, 1998; Wendel 2000); however, counterbalancing modes of genome contraction have not been convincingly shown. In the absence of an equally comprehensive and aggressive mechanism for genome size decrease, the question remains whether angiosperms have a "one-way ticket to genome obesity" (Bennetzen and Kellogg 1997). We have addressed this fundamental issue in the genome size debate by studying the structure and evolution of long terminal repeat (LTR) retrotransposons in *Arabidopsis*.

LTR retrotransposons constitute a large part of the repetitive DNA fraction in plant species. They are characterized by LTRs that vary in size from a few 100 base pairs (bp) to several kilobases and terminate in short inverted repeats, usually 5′-TG-3′ and 5′-CA-3′ (Kumar and Bennetzen 1999). The well-defined structure of LTR retrotransposons, their prevalence and dispersion in the genome, their acknowledged role in genome size expansion, and the fact that individual elements have little or no selective significance make LTR retrotransposons suitable elements for studying genome evolution (Petrov 2001). The prevalence and distribution of LTR retrotransposons have been the subject of several studies, including in *Arabidopsis* (Marín and Lloréns 2000; Terol et al. 2001). These studies, however, are generally based on the analysis of intact elements of relatively recent origin and provide no information on the long-term fate of these sequences. In our study, LTR-retrotransposon families were established on the basis of homology of the LTRs rather than the open reading frames. An important advantage of this approach is that not only complete elements but also solo LTRs and elements that have undergone a variety of deletions can be identified. It is precisely the structure of this latter group that provides the most important clues regarding plant genome evolution.

RESULTS AND DISCUSSION

We have analyzed a total of 291 LTR-retrotransposon elements belonging to 12 families (four *copia,* six *gypsy,* two unknown). The retroelements are distributed over the five *Arabidopsis* chromosomes and show the typical pericentromeric clustering previously observed for LTR retrotransposons (Lin et al. 1999; Mayer et al. 1999), indicating that the 291 elements form a representative sample. The 12 families were originally identified in two bacterial artificial chromosome (BAC) clones that were randomly chosen from a selection of annotated *Arabidopsis* BACs that contained putative LTR retroelements. The LTRs of these elements were then used as query sequences in BLAST searches against the *Arabidopsis* genomic sequence (http://www.arabidopsis.org). Incomplete elements were taken into account only if they retained at least one of the LTR-retrotransposon characteristics such as a primer-binding site (PBS), a polypurine tract (PPT), or a target duplication site (Kumar and Bennetzen 1999). Thus, many severely deleted LTR retrotransposons that we detected were not further studied because their highly fragmentary structure made it impossible to determine the nature of specific rearrangements that they had undergone. Of the 291 studied elements, 87 (29.9%) were found to be "complete"; that is, they contain two LTRs flanked by a 5-bp target-site duplication and separated by an internal region containing a PBS and PPT (Fig. 1A). By use of the dating strategy described by SanMiguel et al. (1998), but applying the synonymous substitution rate of 1.5×10^{-8} mutations per site per year determined for the *Chs* and *Adh* genes in the Brassicaceae (Koch et al. 2000), we estimated that these retrotransposons all inserted in the *Arabidopsis* genome during the last 4 million years, most within the last 2 million years (data not shown). These estimates are based on the assumption that LTRs evolve at approximately

[3]**Corresponding author.**
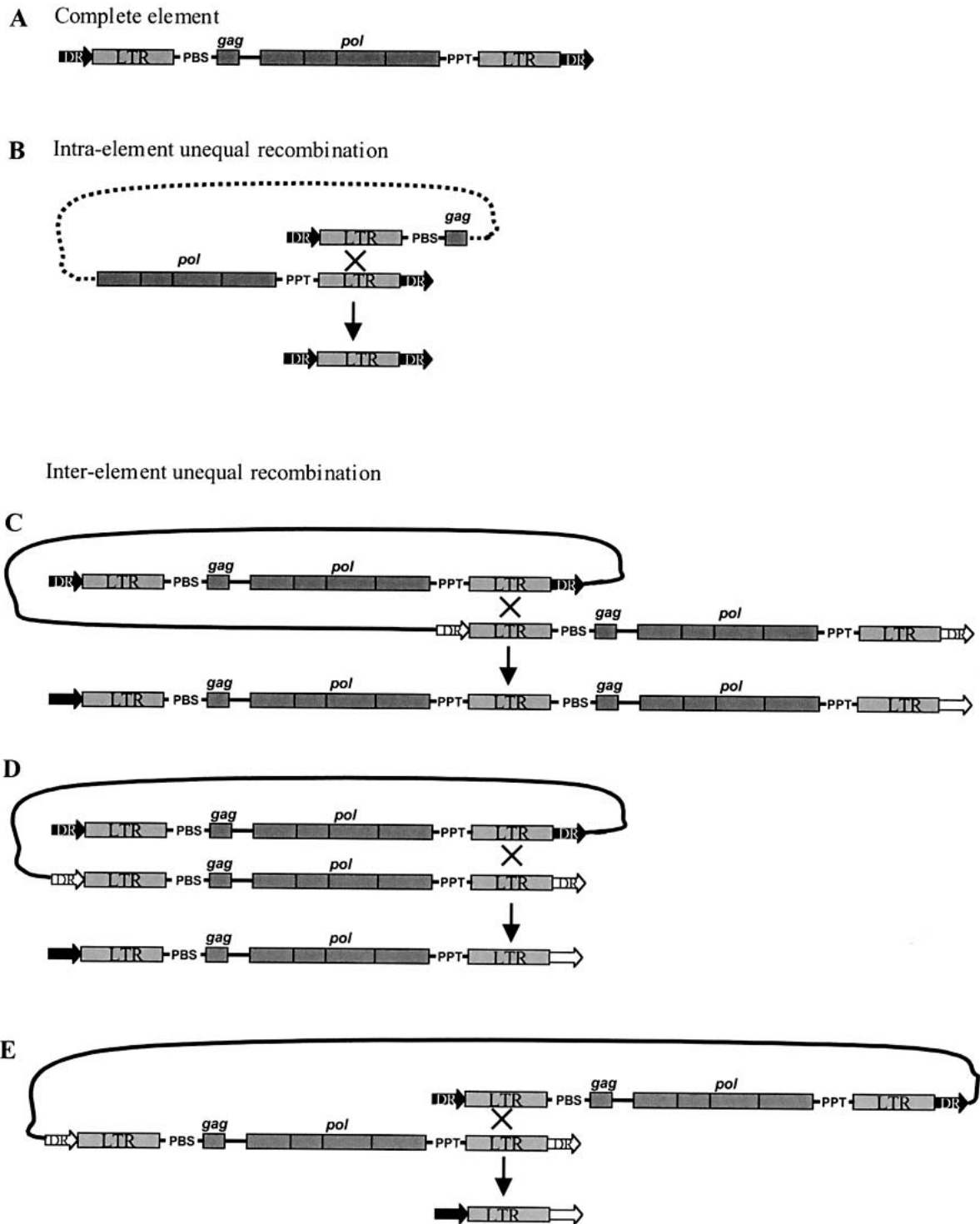**E-MAIL Katrien.devos@bbsrc.ac.uk; FAX 44 1603 450 023/24.**

Figure 1 Unequal intrastrand recombination between LTR retrotransposons. (A) Structure of a complete element, with a direct repeat (DR) of flanking target-site DNA, two long terminal repeats (LTRs), a primer-binding site (PBS), and polypurine tract (PPT) needed for element replication and encoded gene products (gag, pol). (B) Solo LTR resulting from intra-element recombination. The dotted line is presented to facilitate depiction of the folding needed to accomplish this recombination and does not represent any significant stretch of DNA. (C) Recombinant element resulting from recombination between 5′ and 3′ LTRs of two adjacent elements. The solid black line represents sequences between the two LTR retro-transposons that are deleted via this intrastrand event. Note that the two flanking direct repeats are now different, because one is the 3′ target site of one element (black arrow) and the other is the 5′ target site from the adjacent element (white arrow). (D) "Complete" element without target-site duplication resulting from unequal recombination between two 5′ LTRs, two 3′ LTRs, or the internal regions of two elements. Once again, DNA between the two LTR retrotransposons is lost, along with one LTR-retrotransposon copy. (E) "Solo" LTR without target-site duplication resulting from recombination between 3′ and 5′ LTRs of two elements. As before, DNA between the two LTR retrotransposons (solid black line) is lost. In all cases shown, unequal intrastrand recombination only creates a deletion if the two participant retroelements are in direct orientation on the same chromosome arm.

the same rate as coding regions and on our observations that conversion does not frequently occur in these elements, as evidenced by the even distribution of sequence variation (SanMiguel et al. 1998). It would thus appear that, similar to maize (SanMiguel et al. 1998), the *Arabidopsis* genome has undergone a surge of retrotransposon amplification in recent times.

In contrast with maize, which contains mainly intact retroelements and rare solo LTRs (SanMiguel et al. 1996; W. Ramakrishna and J.L. Bennetzen, unpubl.), the ratio of solo LTRs to intact elements in *Arabidopsis* is ~1 : 1. Solo LTRs can be derived from unequal intrastrand recombination between the 5′ and 3′ LTRs of a single element (Fig. 1B). Barley, on the other hand, which has a genome size twice that of maize, contains 16-fold more LTRs than internal retroelement domains for the BARE-1 element, and this excess of LTRs has been ascribed to an abundance of solo LTRs (Vicient et al. 1999; Shirasu et al. 2000). Although intraelement recombination can never neutralize the genome expansion driven by LTR-retrotransposon amplification because a solo LTR is retained, it can play a role in attenuating genome growth (Bennetzen and Kellogg 1997; Vicient et al. 1999). Unequal intrastrand homologous recombination between LTRs of different elements belonging to the same family can result in a net loss of DNA (Fig. 1, C–E). Six examples of this were found in our study (Table 1), four of which resulted in clearly recognizable recombinant products in which an LTR was flanked by both a PBS and PPT (Fig. 1C). An apparently intact element lacked the 5-bp target-site duplication and was therefore expected to be the product of homologous recombination between two 5′ LTRs, two 3′ LTRs, or the internal regions of two family members (Fig. 1D). Similarly, a solo LTR that lacked the target-site duplication was assigned as a recombinant element (Fig. 1E). As observed in numerous studies, including our own, LTR retrotransposons in *Arabidopsis* are particularly abundant in pericentromeric regions, which are largely devoid of genes (Lin et al. 1999; Mayer et al. 1999). Inter-element deletions therefore are unlikely to have a negative effect on the overall fitness of an individual.

In addition to intact elements and solo LTRs, 98 truncated elements (33.7% of the total) were identified (Table 1). They include (1) elements in which the two LTRs are still recognizable but have undergone deletions at either their 3′ or 5′ end (8.3%), (2) elements in which the 5′ LTR together with part of the internal sequence has been deleted (16.5%), and (3) elements in which the 3′ LTR together with part of the

internal sequence has been deleted (8.9%). The remaining LTRs of elements belonging to the latter two groups may have undergone further deletions and have been included in the analysis only if their identity could be established unequivocally. Therefore, the percentage of LTR-retrotransposon remnants is much more than 33.7%. The discovery of small deletions as a major mode for genome size determination in *Arabidopsis* parallels results obtained by Petrov and coworkers in *Drosophila*, a species with a DNA content similar to that of *Arabidopsis*. On the basis of the rate of insertions and deletions in a non-LTR retrotransposon, Petrov and Hartl (1998) calculated that pseudogenes lose ~50% of their DNA in 14 million years through spontaneous deletions. Deletions have also been shown to be a frequent event in transposable elements in maize (Masson et al. 1987; Marillonnet and Wessler 1998) and to feature in LTR retrotransposons in wheat (Wicker et al. 2001), which have genome sizes that are 20 and 120 times larger than the *Arabidopsis* genome, respectively. The results in *Arabidopsis*, maize, and wheat indicate that deletions that are independent of homologous recombination (equal or unequal) represent a key mechanism for DNA elimination in plants.

In an attempt to shed light on the molecular mechanism(s) that gave rise to the deletions, we compared the internal regions of retroelements belonging to 3 of the 12 families. The three families contained 37% of the retroelements for which internal regions could be analyzed and were assumed to be a representative sample. The comparisons included 33, 5, and 5 elements of the three families. We analyzed the breakpoints of a total of 59, 8, and 6 deletions, respectively, ranging in size from 10 to 3766 bp. Deletions that were shared between elements within a family were assumed to have a common descent and were considered only once. It should be noted that although some of the deletions encompassed >500 bp, only four affected the structural characteristics such as LTRs, PBS, and PPT used in our assessment of the intactness of the LTR retroelements and thus led to the classification of the corresponding elements as "incomplete." Of the 59, 8, and 6 deletions, 46 (78%), 4 (50%), and 6 (100%), respectively, were flanked by short repeats of 2 to 13 bp, some of which were imperfect. Taking into account the base composition in the internal regions of the LTR retroelements and the distribution of sequences homologous to the short flanking repeats, the association of the repeats with the deletions was highly significant for each of the families (Table 2). Similarly, an analysis of six tandem duplications present in the region under investigation also showed a highly significant association of the duplications with short repeats (Table 2). The importance of short repeats in deletion and duplication formation is well documented in bacteria and yeast. Classical homologous recombination requires homologous sequences of at least 20 bp in bacteria (Ehrlich 1989) and 50 to 100 bp in yeast (Sugawara and Haber 1992), whereas shorter repeats engage solely in illegitimate recombination. The high frequency of short repeats associated with the deletions in *Arabidopsis* retroelements indicates that genome expansion

Table 1. Structure of LTR Retroelements Identified in *Arabidopsis*

Structure	No. of elements	Mechanism
Intact elements	87	
Solo LTRs	101	Intra-element unequal recombination
Intact elements without DR	1	Intra-element unequal recombination
Solo LTRs without DR	1	Intra-element unequal recombination
Recombined elements contain an LTR flanked by both PBS and PPT	3a	Intra-element unequal recombination
3′ LTR deleted	26	Illegitimate recombination
5′ LTR deleted	48	Illegitimate recombination
5′ and/or 3′ LTR partially deleted	24	Illegitimate recombination

aOne element appears to be the result of two inter-element unequal homologous recombination events in which the 5′ and 3′ LTRs of an element recombined with the 3′ LTR of an upstream element, and the 5′ LTR of a downstream element, respectively.
LTR, long terminal repeat; PBS, primer-binding site; PPT, polypurine tract; DR, direct repeat.

Table 2. Statistical Significance of the Association of Short Repeats with Deletions and Tandem Duplications

LTR retrotransposon family	Elements	Deletions/ duplica- tions	No. flanked by short DRs[a]	No. of repeats in randomization test (Test 1)		Bases in repeats	Bases matching in 5′ and 3′ sequences	No. of matching bases in randomization test (Test 2)	
				Mean ± SD (max.)	P			Mean ± SD (max.)	P
(a) Deletions									
Family 1	5	6	6	0.18 ± 0.43 (3)	3×10^{-11}	40	37	9.5 ± 2.5 (19)	2×10^{-7}
Family 2	5	8	4	0.17 ± 0.38 (2)	6×10^{-9}	18	18	4.8 ± 1.8 (10)	4×10^{-4}
Family 3	33	59	46	0.64 ± 0.79 (4)	6×10^{-15}	281	251	62.6 ± 6.9 (87)	4×10^{-34}
All three families	43	73	56	0.85 ± 0.90 (5)	3×10^{-14}	339	306	76.6 ± 7.4 (101)	2×10^{-41}
(b) Tandem duplications									
Family 3	6	6	6	0.20 ± 0.44 (2)	2×10^{-10}	26	25	6.2 ± 2.2 (15)	8×10^{-5}

[a]A repeat was considered as starting from the deletion or duplication breakpoint and was extended if the bases in the 5′ and 3′ flanking sequences matched or (if there was a mismatching base) that mismatch was followed by at least 2 matching bases. The repeat was considered as stopping where a mismatch was followed by only 1 matching base.

through retrotransposon amplification can be counterbalanced by a gradual removal of the elements through illegitimate recombination. Unfortunately, our data set does not allow us to ascertain whether illegitimate recombination takes place by errors in DNA replication, by double-strand break repair (Gorbunova and Levy 1999), or by some unknown mechanism.

The formation of deletions during double-strand break repair was recently investigated in *Arabidopsis* and tobacco, two species that vary 20-fold in their DNA content (Gorbunova and Levy 1997; Kirik et al. 2000). It was shown that strand rejoining after a break frequently occurs at short repeats and results in the deletion of a few base pairs to several kilobases of DNA. The average deletion size was significantly smaller in tobacco than *Arabidopsis* (920 bp vs. 1341 bp) (Kirik et al. 2000). A negative correlation between genome size and rate of DNA loss was also postulated by Petrov and coworkers. They estimated that the rate of DNA loss in *Drosophila* was ~40-fold higher than in the 11-fold larger genome of *Laupala* and 60-fold higher than in, on average, 18-fold larger mammalian genomes (Petrov and Hartl 1998; Petrov et al. 2000). The observation that short repeats are often associated with deletions in *Drosophila* (Petrov and Hartl 1998) indicates to us that illegitimate recombination is also a major determinant of DNA loss in *Drosophila*. Although we can conclude that DNA is effectively removed from small genome organisms through illegitimate recombination, no information is available on the driving force behind the differential loss of DNA in small and large genomes. Genome size is clearly the result of a balance between amplification and loss of DNA. However, it remains to be seen whether organisms have an active role in determining the ratio of DNA gain to loss or whether this ratio is the result of evolutionary forces acting on the nongenic DNA.

The insight that elimination of LTR retrotransposons takes place through illegitimate recombination forces us to reassess our earlier suggestion that the genome size of *Arabidopsis* has increased considerably over the past 4 million years through retrotransposon insertion. Considering the large number of retrotransposon remnants, it now seems likely that the apparent absence of elements older than a few million years is simply a reflection of their gradual degradation over time. As our data set does not allow conclusions to be drawn on the relative rate of DNA removal and amplification, it is

now an open question whether the *Arabidopsis* genome size has increased, decreased, or remained constant over recent times.

We only included clearly recognizable elements in our study. Sequences of a few tens to hundreds of base pairs with homology with retroelement LTRs, however, were identified and bear further witness to the fact that genomes use mechanisms other than unequal homologous recombination to remove repetitive elements. Moreover, our data indicate that deletion through illegitimate recombination is more important than unequal homologous recombination events in eliminating DNA in *Arabidopsis*. We only analyzed deletions in relatively intact retrotransposons because they could be precisely defined; however, clearly identifiable retroelements make up <10% of the *Arabidopsis* genome (The Arabidopsis Genome Initiative 2000). Although retrotransposons served as a particularly clear indicator for genome size contraction, we have no reason to doubt that illegitimate recombination will also remove DNA from the other 90% of the *Arabidopsis* nuclear genome. Selection against gene loss will attenuate deletions in the 44% of the genome that is genic (25,000 genes with an average length of 2 kb) (The Arabidopsis Genome Initiative 2000), but illegitimate recombination is likely to proceed unimpeded in the noncoding DNA. Furthermore, in contrast with unequal homologous recombination, which requires the presence of closely linked direct repeats and ends when only one LTR unit remains, multiple independent illegitimate recombination events can and do occur in any region, eventually removing all unselected sequence. We predict that illegitimate recombination removes at least fivefold more DNA than unequal homologous recombination because illegitimate recombination can act on at least 5 times more of the genome than can unequal recombination between LTRs, and because we saw many more severely deleted LTR retrotransposons than we did solo LTRs. Observations that the *Arabidopsis* genome is composed of numerous duplicated segments with subsequent genic deletions (Blanc et al. 2000; Ku et al. 2000) is totally compatible with our model of genome contraction via illegitimate recombination.

METHODS

Identification and Alignment of LTR Retroelements

The programs `Repeat` and `Gap` from the Wisconsin Package Version 10.1, Genetics Computer Group were used for the

initial identification and alignment of LTRs belonging to the same retroelement on *Arabidopsis* BAC clones K11J14 (AP000411) and T24G23 (AC006268). Putative LTR retroelements were scrutinized manually for the presence of a TG/CA inverted repeat in the LTRs, a PBS, a PPT, and a target-site duplication. LTRs of confirmed elements were used as query sequences in BLASTN (NCBI BLAST 2.0) searches against the *Arabidopsis thaliana* database (http://www.arabidopsis.org) to identify additional family members. LTRs and internal regions of elements belonging to the same family were aligned using CLUSTALX (Thompson et al. 1997). If needed, sequence alignments were edited manually using JalView (M. Clamp, EBI).

Statistical Analysis

Randomization tests involving matching of flanking sequences (Test 1) and matching of individual bases in the flanking sequences (Test 2) were performed to determine the statistical significance of the association of short repeats with deletions and tandem duplications. In each run of Randomization Test 1, one flanking sequence of each deletion or duplication was held fixed while sequences of the same length were sampled at random from all available sequences of the appropriate LTR-retrotransposon family. Deletions or duplications for which the match between the randomly sampled sequence and the fixed sequence was at least as close as that of the two actual flanking sequences were counted. One thousand separate randomizations were run for deletions in each LTR-retrotransposon family for the complete set of LTR retrotransposons and for tandem duplications involving Family 3. The probability (P) is that of obtaining by chance at least the observed number of matches under the null hypothesis that sequences flanking each deletion or duplication are unrelated (one-tailed t test using a logarithmic transformation). Randomization Test 2 was essentially similar but involved matching of individual bases in the flanking sequences rather than the complete sequences.

ACKNOWLEDGMENTS

REFERENCES

Bennett, M.D. and Leitch, I.J. 1995. Nuclear DNA amounts in Angiosperms. *Ann. Bot.* **76:** 113–176.

Bennetzen, J.L. and Kellogg, E.A. 1997. Do plants have a one-way ticket to genomic obesity? *Plant Cell* **9:** 1509–1514.

Blanc, G., Barakat, A., Guyot, R., Cooke, R., and Delseny, M. 2000. Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* **12:** 1093–1101.

Ehrlich, S.D. 1989. Illegitimate recombination in bacteria. In *Mobile DNA* (eds. D.E. Berg and M.M. Howe), pp. 799–832. American Society for Microbiology, Washington, D.C.

Gorbunova, V. and Levy, A.A. 1997. Non-homologous DNA end-joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Res.* **25:** 4650–4657.

———. 1999. How plants make ends meet: DNA double-strand break repair. *Trends Plant Sci.* **4:** 263–269.

Kirik, A., Salomon, S., and Puchta, H. 2000. Species-specific double-strand break repair and genome evolution in plants. *EMBO J.* **19:** 5562–5566.

Koch, M.A., Haubold, B., and Mitchell-Olds, T. 2000. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). *Mol. Biol. Evol.* **17:** 1483–1498.

Ku, H.-M., Vision, T., Liu, J., and Tanksley, S.D. 2000. Comparing sequenced segments of the tomato and *Arabidopsis* genomes: Large-scale duplication followed by selective gene loss creates a network of synteny. *Proc. Natl. Acad. Sci.* **97:** 9121–9126.

Kumar, A. and Bennetzen, J.L. 1999. Plant retrotransposons. *Annu. Rev. Genet.* **33:** 479–532.

Lin, X.Y., Kaul, S.S., Rounsley, S., Shea, T.P., Benito, M.I., Town, C.D., Fujii, C.Y., Mason, T., Bowman, C.L., Barnstead, M., et al. 1999. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402:** 761–768.

Marillonnet, S. and Wessler, S.R. 1998. Extreme structural heterogeneity among the members of a maize retrotransposon family. *Genetics* **150:** 1245–1256.

Marín, I. and Lloréns, C. 2000. *Ty2/Gypsy* retrotransposons: Description of new *Arabidopsis thaliana* elements and evolutionary perspectives derived from comparative genomic data. *Mol. Biol. Evol.* **17:** 1040–1049.

Masson, P., Surosky, R., Kingsbury, J.A., and Fedoroff, N.V. 1987. Genetic and molecular analysis of the *Spm-dependent a-m2* alleles of the maize *a* locus. *Genetics* **117:** 117–137.

Mayer, K., Schuller, C., Wambutt, R., Murphy, G., Volkaert, G., Pohl, T., Düsterhölt, A., Stickema, W., Entian, K.-D., Terryn, N. et al. 1999. Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. *Nature* **402:** 769–777.

Petrov, D.A. 2001. Evolution of genome size: New approaches to an old problem. *Trends Genet.* **17:** 23–28.

Petrov, D.A. and Hartl, D.L. 1998. High rate of DNA loss in the *Drosophila melanogaster* and *Drosophila virilis* species group. *Mol. Biol. Evol.* **15:** 293–302.

Petrov, D.A., Sangster, T.A., Johnston, J.S., Hartl, D.L., and Shaw, K.L. 2000. Evidence for DNA loss as a determinant of genome size. *Science* **287:** 1060–1062.

SanMiguel, P., Tikhonov, A., Jin, Y.-K., Motchoulskaia, N., Zakharov, D., Melake Berhan, A., Springer, P.S., Edwards, K.J., Avramova, Z., et al. 1996. Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274:** 765–768.

SanMiguel, P., Gaut, B.S., Tikhonov, A., Nakajima, Y., and Bennetzen, J.L. 1998. The paleontology of intergene retrotransposons of maize: Dating the strata. *Nat. Genet.* **20:** 43–45.

Shirasu, K., Schulman, A.H., Lahaye, T., and Schulze-Lefert, P. 2000. A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. *Genome Res.* **10:** 908–915.

Sugawara, N. and Haber, J.E. 1992. Characterization of double-strand break-induced recombination: Homology requirements and single-stranded DNA formation. *Mol. Cell. Biol.* **12:** 563–575.

Terol, J., Castillo, M.C., Bargues, M., Pérez-Alonso, M., and de Frutos, R. 2001. Structural and evolutionary analysis of the *copia*-like elements in the *Arabidopsis thaliana* genome. *Mol. Biol. Evol.* **18:** 882–892.

The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408:** 796–815.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. 1997. The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24:** 4876–4882.

Vicient, C.M., Suoniemi, A., Anamthawat-Jónsson, K., Tanskanen, J., Beharav, A., Nevo, E., and Schulman, A.H. 1999. Retrotransposon *BARE-1* and its role in genome evolution in the genus *Hordeum*. *Plant Cell* **11:** 1769–1784.

Wendel, J.F. 2000. Genome evolution in polyploids. *Plant Mol. Biol.* **42:** 225–249.

Wicker, T., Stein, N., Albar, L., Feuillet, C., Schlagenhauf, E., and Keller, B. 2001. Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. *Plant J.* **26:** 307–316.

WEB SITE REFERENCES

http://www.arabidopsis.org; *Arabidopsis* information resource.