Current Biology 22, 248–252, February 7, 2012 ©2012 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2011.12.027

Report

Rapid, Repeated, and Clustered Loss of Duplicate Genes in Allopolyploid Plant Populations of Independent Origin

Richard J.A. Buggs,[1,2,3,*] Srikar Chamala,[1,4] Wei Wu,[5] Jennifer A. Tate,[6] Patrick S. Schnable,[5] Douglas E. Soltis,[1,4] Pamela S. Soltis,[2,4] and W. Brad Barbazuk[1,4]
[1]Department of Biology
[2]Florida Museum of Natural History
University of Florida, Gainesville, FL 32611, USA
[3]School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK
[4]Genetics Institute, University of Florida, Gainesville, FL 32610 USA
[5]Center for Plant Genomics, Iowa State University, Ames, IA 50011, USA
[6]Massey University, Institute of Molecular Biosciences, Palmerston North 4442, New Zealand

Summary

The predictability of evolution is debatable, with recent evidence suggesting that outcomes may be constrained by gene interaction networks [1]. Whole-genome duplication (WGD; polyploidization—ubiquitous in plant evolution [2]) provides the opportunity to evaluate the predictability of genome reduction, a pervasive feature of evolution [3, 4]. Repeated patterns of genome reduction appear to have occurred via duplicated gene (homeolog) loss in divergent species following ancient WGD [5–9], with evidence for preferential retention of duplicates in certain gene classes [8–10]. The speed at which these patterns arise is unknown. We examined presence/absence of 70 homeologous loci in 59 *Tragopogon miscellus* plants from five natural populations of independent origin; this allotetraploid arose ~80 years ago via hybridization between diploid parents and WGD [11]. Genes were repeatedly retained or lost in clusters, and the gene ontology categories of the missing genes correspond to those lost after ancient WGD in the same family (Asteraceae; sunflower family) [6] and with gene dosage sensitivity [8]. These results provide evidence that the outcomes of WGD are predictable, even in 40 generations, perhaps due to the connectivity of gene products [8, 10, 12]. The high frequency of single-allele losses detected and low frequency of changes fixed within populations provide evidence for ongoing evolution.

Results and Discussion

Frequency of Homeolog Loss

Analyses of synthetic polyploids show that rapid gene loss may occur after whole-genome duplication (WGD) [13–22], but we know little about this process in young natural polyploids. *Tragopogon miscellus* (Asteraceae) is a natural allotetraploid (2n = 24; Figures 1A and 1B) that arose approximately 80 years ago (40 generations, because the species is biennial) in eastern Washington and adjacent Idaho, USA, from two diploids introduced from Europe in the early 1900s: *T. dubius* and *T. pratensis*, each with 2n = 12 [23]. *Tragopogon miscellus* has formed repeatedly in different localities from separate populations of the diploid progenitors [11, 23], yielding replicated independent natural allopolyploid lines. Seventy Sequenom MassARRAY assays were used to detect homeolog loss in 59 *T. miscellus* plants from five natural populations (Figure 1; Table 1). An average of 13 (~20%) of the 70 loci investigated in each plant of *T. miscellus* were missing alleles of one homeolog, with high variation (min = 0, max = 33, SD = 6.8). Of these, an average of 8.6 (SD = 5.9) loci per plant were missing both alleles of a homeolog, and 4.4 (SD = 4.3) were missing only one allele. Assuming that immediately after WGD each individual had two alleles from *T. dubius* and two from *T. pratensis* at each pair of homeologous loci (see Figure 1B), on average at least 7.7% (SD = 4.3) of the original 280 allele copies have been lost in an individual plant's lineage since WGD. The five populations studied differ in total frequency of homeolog losses (Table 1), possibly indicating slight differences in the timing of independent origins of *T. miscellus* populations within the last 80 years [24].

The high frequency of single-allele absences found here suggests ongoing evolution in these populations, as does the fact that few homeolog losses are fixed in populations. The rates of homeolog loss are higher than those found previously using a smaller number of cleaved amplified polymorphic sequence (CAPS) markers [25–27], in part because CAPS cannot detect absence of single alleles of one homeolog. Here we also report the first instance in *T. miscellus* of the same homeolog being absent in every individual of a population (the *T. dubius*-derived homeolog of gene 07259_1241 in Garfield). Other studies have shown considerable loss of bands in amplified fragment length polymorphism or restriction fragment length polymorphism profiles of polyploid plants in *Brassica* [16, 19], *Aegilops* [20], *Nicotiana* [28], *Triticum* [13, 21], *Eragrostis* [22], and *Paspalum* [29], but the relationship between loss of bands and true loss of alleles or loci cannot be determined using these methods [see 24].

Patterns of Loss among Loci

The frequency of homeolog loss varied considerably among loci. On average, a locus showed homeolog loss in 11 plants (of 59; excluding cases where neither homeolog was detected), and variation was high (min. = 0, max. = 36, SD = 9.2). Homeolog losses among plants and loci were investigated by average linkage hierachical clustering with a Spearman rank correlation metric. This analysis showed that 12 clusters of loci follow recurrent patterns of presence/absence in populations with unique origins (Figure 1C). Therefore, the subsequent evolution of genes duplicated via polyploidy appears to be repeated in independent lineages. Clustering has been previously reported for genes retained as pairs in *Arabidopsis* [10] and maize [5], but clustering of genes lost in maize was not found despite focused analyses [5]. The discovery reported here of clustered patterns of missing genes in *Tragopogon* is therefore novel.

Patterns Due to Parentage

We investigated whether the maternal versus paternal direction of cross in the origins of populations affected rates of

*Correspondence: r.buggs@qmul.ac.uk

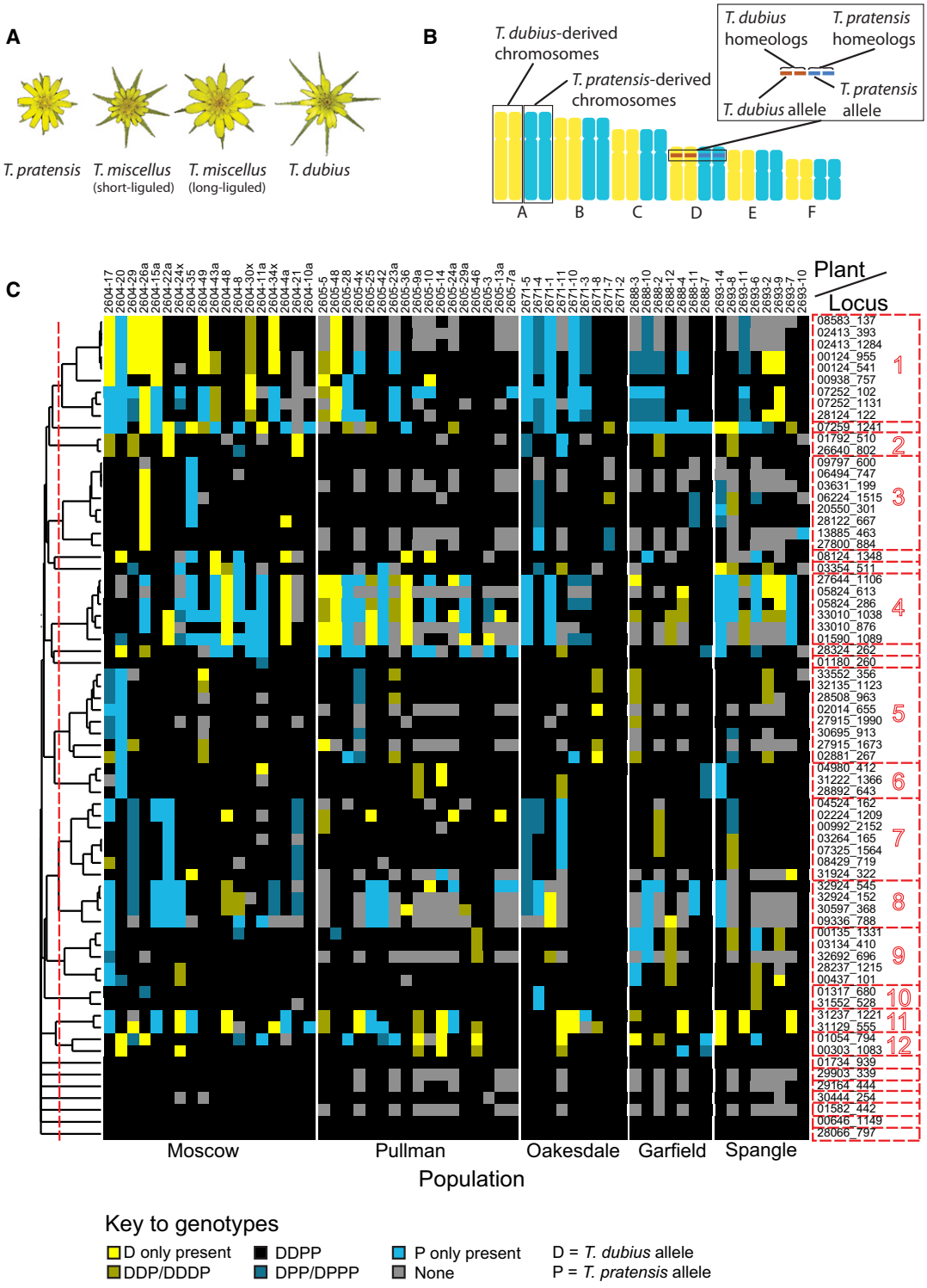

Figure 1. Genome Evolution in Allotetraploid *Tragopogon miscellus*

(A) Sketches of inflorescences L to R: *T. pratensis* (2x), *T. miscellus* (4x, short liguled, with *T. dubius* as paternal parent), *T. miscellus* (4x, long-liguled, with *T. pratensis* as paternal parent), and *T. dubius* (2x).

(B) Chromosome diagram for *T. miscellus*, illustrating terminology.

(C) Genotype calls at 70 loci in 59 plants from five natural populations of *T. miscellus*. Colors show different homeolog presence calls (see Key). Loci are clustered by similarity of pattern of homeolog presence among plants; the tree shows an average linkage hierarchical clustering with a Spearman rank correlation similarity metric; 12 groups arising from this clustering are shown in red. Plants in each population are ordered by number of homeolog losses.

Table 1. Presence/Absence of Homeologs at 70 SNP Loci in Natural Populations of *Tragopogon miscellus*

	Short-liguled					Long-liguled
	Moscow	Oakesdale	Garfield	Spangle	All short-liguled	Pullman
Number of plants	18	9	7	8	*42*	17
Number of data points	1260	630	490	560	*2940*	1190
Genotypes						
D only	7.0%	1.6%	1.2%	4.5%	*4.4%*	4.4%
P only	10.3%	9.5%	6.3%	6.6%	*8.8%*	5.6%
D and P: balanced	72.3%	74.8%	69.4%	62.1%	*70.4%*	69.9%
D and P: one D allele absent	3.3%	6.7%	4.7%	3.2%	*4.3%*	0.8%
D and P: one P allele absent	2.5%	1.9%	4.9%	5.0%	*3.2%*	2.8%
No call	4.6%	5.6%	13.5%	18.6%	*8.9%*	16.6%

loss of homeologs from each parent. Some analyses of gene loss in synthetic polyploids have shown a bias against one parental genome [5, 16, 19, 24]. In the present study, four *T. miscellus* lines had *T. dubius* as the paternal parent, but one (Pullman) had *T. pratensis* as the paternal parent (Figure 1A). In the Pullman population (n = 17), there was a small but nonsignificant bias toward losses of both alleles from the maternal parent (one-tailed Mann-Whitney U test, W = 123, p = 0.2316), and in the other four populations (n = 42), there was a significant bias toward losses of both alleles from the paternal parent (one-tailed Mann-Whitney U test, W = 573.5, p = 0.00272). The Pullman population also has a nonsignificant bias toward the loss of single paternal alleles (one-tailed Mann-Whitney U test, W = 105, p = 0.0575), as did the other four populations (one-tailed Mann-Whitney U test, W = 935, p = 0.34524). Previous CAPS studies [25–27] showed significant bias toward loss of *T. dubius* homeologs in *T. miscellus* populations but were not able to distinguish an effect due to reciprocal parentage.

Comparisons with Ancient Polyploids

Plants considered to be ancient polyploids show repeated patterns in the genes that are retained as duplicates based on function [5–9], but we do not know whether similar patterns occur in recent polyploids. We compared our data from *T. miscellus* to a study of 18 species from the Asteraceae (to which *Tragopogon* belongs), which have evidence for ancient whole-genome duplications and parallel patterns of subsequent loss or retention of duplicated genes [6]. We asked whether homeolog presence/absence in young *Tragopogon* polyploids paralleled these patterns. The Sequenom assay targets were annotated using Blast2GO with database b2g_may10 (see Table S1B available online). In seven cases, two assays hit different parts of the same gene; within each of these pairs, highly similar patterns of presence/absence were found, and we took an average of the two assays in measuring occurrence of that gene. Eighteen genes in the current study had gene ontology (GO) categories that tended to be lost in Asteraceae [6], and these had mean homeolog loss of 20.5% in *T. miscellus* (SD = 16.1). A comparison of these 18 genes versus all other genes (45 genes, with a mean absence of 14.0% and SD of 14.8) showed significantly higher homeolog loss in the 18 genes (one-tailed Mann-Whitney U test, W = 689.5, p = 0.0427). Thus, gene loss in *T. miscellus*, a young Asteraceae polyploid, appears to repeat patterns found in ancient Asteraceae polyploids. We also compared patterns of gene loss in *T. miscellus* to the model plant *Arabidopsis thaliana* (Brassicaceae), but as with Barker et al. [6] who found Asteraceae species to show different patterns of gene loss versus retention from *Arabidopsis* after ancient WGD, we did not find a similar pattern (see Supplemental Information).

Test of the Gene-Balance Hypothesis

Repeated patterns in the loss versus retention of duplicated genes might be explained and predicted within the realm of evolutionary systems biology, linking the evolution of genes to their function within networks [1, 30]. One such explanation, known as the gene-balance hypothesis [8, 10, 12], holds that genes coding for products that are highly connected—within protein complexes or biochemical pathways—are sensitive to dosage ("dosage-sensitive") in that they must be present in the same number of genomic copies as the genes with whose products they interact. These connected genes are hypothesized to be retained together in duplicate copies to preserve stoichiometry, unlike other ("dosage-insensitive") genes that will revert to singleton status one by one over time [8]. Evidence in support of the gene-balance hypothesis has been found in patterns of whole-genome sequences of putative ancient polyploids [8, 10]. However, the tempo and mode of gene loss and retention are unclear. We used the *Tragopogon* gene descriptions and GO categories to classify (Table S1B) the dosage sensitivity of genes according to Freeling [8]. The mean homeolog absence of 26 putatively dosage-sensitive genes was 4.8% (SD = 5.0), whereas for 24 putatively dosage-insensitive genes, the mean was 10.8% (SD = 8.2); the remaining 13 genes could not be categorized. This difference was statistically significant (one-tailed Mann-Whitney U test, W = 744.5, p = 0.0052), suggesting that dosage sensitivity may play a role in determining the loss versus retention of homeologs.

Clusters of genes (see above and Figure 1C) that were retained together in lineages of *T. miscellus* of separate origin could represent genes whose products interact or classes of genes that are selected for in a similar way. For example, cluster 9 comprises a set of genes that have mostly been retained in duplicate across all populations, and three of the five genes in this cluster are involved in biosynthesis of secondary metabolites (Table S1B). We also found some evidence for functional similarity in genes showing similar patterns of loss: clusters 4 and 11, which showed much gene loss, included two electron carrier proteins and two genes involved in amino acid and nucleotide sugar metabolism, respectively. Clusters of lost genes could also result from the loss of large fragments of chromosomes or perhaps even entire chromosomes [18, 24], on which the "clustered" genes are located. Genomic in situ hybridization karyotyping of *T. miscellus* plants has revealed homeologous recombination

as well as frequent translocations and reciprocal monosomy/trisomy and nullisomy/tetrasomy [31, 32], processes that may provide mechanisms for coordinated homeolog loss. The gene-balance hypothesis and chromosomal linkage hypotheses for patterns of homeolog loss are not mutually exclusive. Several studies have shown that genes with similar patterns of expression and/or function are often physically clustered in the genome [33–36]. Spatial clustering of similar genes could enhance the fitness effects of the loss of the section of chromosome on which they occur, increasing the strength and accelerating the action of natural selection.

Conclusion

This detailed population-level study of homeolog presence/absence in a young natural allopolyploid species—which for the first time documents absences of single allelic copies of homeologs—shows that, in the midst of genome turmoil, clusters of genes tend to be lost or retained together. Furthermore, the patterns of gene loss and retention show repeatability among independently formed lineages within a polyploid species, among polyploids of contrasting age in the same plant family, and in the retention of dosage-sensitive genes. Our data suggest that the evolution of genomes after WGD is to some degree governed by the attributes of gene interaction networks; evolutionary systems biology can therefore make evolution after WGD predictable.

Experimental Procedures

Genomic DNA Sources

Seeds were germinated and grown in a greenhouse at the University of Florida (Gainesville, FL, USA) from the following: five natural populations of *T. miscellus* (Table S2); their diploid parent species, *T. dubius* and *T. pratensis*; and F_1 hybrid plants formed through controlled pollinations of *T. pratensis* with *T. dubius* pollen [37] (crosses 63-1, 79-1, 86-2, 88-5, and 63-4 [37]). Leaf tissue was collected from seedlings four weeks after germination and flash frozen in liquid nitrogen [26]. DNA was extracted from leaf tissue using a modified CTAB protocol.

Single Nucleotide Polymorphism Assays

Sequenom MassARRAY iPLEX genotyping [38] multiplexes locus assays and allows for the detection of losses of single alleles (of which each homeolog normally has two immediately after WGD, see Figure 1B). This method is especially suited for detecting homeologs that differ at only a few nucleotide positions (for details see [38]). Assays were previously designed for 139 putative single nucleotide polymorphisms (SNPs) identified using next-generation sequencing ([38]; Table S1A). These were used to analyze genomic DNA from the five natural populations at Iowa State University, and the traces were analyzed using the Sequenom Typer 4.0 software package (Sequenom, Inc., San Diego, CA, USA). We examined 139 putative SNP loci in 93 plants: 13 *T. pratensis*, 16 *T. dubius*, 59 *T. miscellus*, and five synthetic F_1 hybrids. Of the 139 assays, 88 provided scorable results in genomic DNA.

Primer Polymorphism Analysis

To ensure that our data on gene loss were not affected by polymorphisms within Sequenom primer-binding sites, we analyzed variation in natural *Tragopogon* populations as follows. Using Mosaik Aligner (version 1.1.0020), we aligned Roche 454 reads and Illumina cDNA 36 bp reads from seven *T. dubius* populations (1.26 M 454 and 7.13 M Illumina reads), four *T. pratensis* populations (0.16 M 454 and 6.84 M Illumina reads), and three *T. miscellus* populations (82.5 M Illumina reads) to the *T. dubius* contigs used in the initial design of the Sequenom assays [38]. The sequence reads are available on the NCBI sequence read archive (SRA047022 and SRA009218). Using Gigabayes (version 0.4.1), we counted the number of polymorphisms in all primer-binding sites. The 51 failed assays had on average 1.51 (SD = 1.41) polymorphic sites within their three primer-binding sites. Of the 88 scorable loci, 14 showed polymorphism in *T. dubius* populations and were therefore excluded. Seventy-four loci showed different alleles in *T. dubius* and *T. pratensis* and no polymorphism within each diploid species and appeared suitable for reliably detecting homeolog presence/absence. However, four of these assays showed only one allele to be present in at least one diploid F_1 plant: 33319_126 (2/5 plants), 06494_282 (2/5 plants), 02348_489 (2/4 plants), and 11285_699 (1/3 plants). We designed PCR primers for two of these loci (06494_282 and 02348_489), at different locations from the Sequenom primers, and amplified these in the F_1 plants, finding both alleles to be present by Sanger sequencing. Therefore, we excluded these four Sequenom assays when analyzing the allopolyploid *T. miscellus* populations. These four assays were also found to have an average of 2.00 (SD = 1.83) SNPs within their primer-binding sites in natural populations, whereas the 70 remaining assays had an average of 0.70 (SD = 0.95) such SNPs. For these 70 assays, which we used to detect homeolog loss, there was an insignificant and negative correlation (two-tailed Spearman rank correlation $R_s = -0.0871$, $p = 0.473993$) between the number of SNPs found in the three primer-binding sites and the number of homeolog absences found at each locus, so the varying frequencies of absence found among loci were not due to primer-binding problems.

Genotype Calling

For the 70 assays, we called genotypes using plots generated by Sequenom Typer 4.0 software. For example, Figure S1 shows the Typer 4.0 plot of the height of the T nucleotide (specific to the *T. dubius* genome) peak versus the height of the C nucleotide (specific to the *T. pratensis* genome) peak for all plants in the study for locus 30597_368. The data form six clear clusters. The cluster labeled blue, and called as homozygous "C" by Typer 4.0, contained all the *T. pratensis* diploid plants and some *T. miscellus* plants. The orange cluster, called as homozygous "T" by Typer 4.0, contained all the *T. dubius* diploid plants and some *T. miscellus* plants. The green cluster, called as heterozygous "CT" by Typer 4.0, contained all the F_1 hybrids and many *T. miscellus* plants. The three clusters in red contained *T. miscellus* plants and were not called by Typer 4.0. Two of these red clusters show heterozygotes with dosage biased toward one particular nucleotide. We called such intermediate genotypes as putative cases where a single allelic copy of one homeolog was missing. Thus, for locus 30597_368, we called the cluster of two red points and an orange point (circled in Figure S1) as loss of a single "C" allele (i.e., "CTT" or "CTTT"; the data do not allow us to distinguish between these possibilities, the latter being possible due to gene conversion, nonhomologous recombination or monosomy/trisomy), and the other cluster of five red points as loss of a single "T" allele (i.e., "CCT" or "CCCT").

Clustering Analysis

Patterns of homeolog loss and retention among loci were clustered using Gene Cluster 3.0 (Michiel de Hoon, Human Genome Center, University of Tokyo, Japan) using average linkage hierarchical clustering with a Spearman rank correlation similarity metric. These results were visualized using Java Treeview [39] and arranged in Adobe Illustrator (Adobe Systems, San Jose, CA, USA).

Accession Numbers

The NCBI sequence read archive submission number for the new EST sequences reported in this paper is SRA047022.

Supplemental Information

Supplemental Information includes one figure, two tables, and Supplemental Results and can be found with this article online at doi:10.1016/j.cub.2011.12.027.

Acknowledgments

Funding for this research was provided by the University of Florida and National Science Foundation grants DEB-0614421 to D.E.S. and P.S. Soltis; MCB-0346437 to D.E.S., P.S. Soltis, and J.A.T.; DEB-0919254 to D.E.S., P.S. Soltis, and W.B.B.; and DEB-0919348 to P.S. Schnable. R.J.A.B. has been supported since March 2010 by Natural Environment Research Council Fellowship NE/G01504X/1.

Received: November 9, 2011
Revised: December 9, 2011
Accepted: December 9, 2011
Published online: January 19, 2012

References

1. Stern, D.L., and Orgogozo, V. (2009). Is genetic evolution predictable? Science *323*, 746–751.
2. Soltis, D.E., Albert, V.A., Leebens-Mack, J., Bell, C.D., Paterson, A.H., Zheng, C., Sankoff, D., Depamphilis, C.W., Wall, P.K., and Soltis, P.S. (2009). Polyploidy and angiosperm diversification. Am. J. Bot. *96*, 336–348.
3. Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honoré, N., Garnier, T., Churcher, C., Harris, D., et al. (2001). Massive gene decay in the leprosy bacillus. Nature *409*, 1007–1011.
4. Wolfe, K.H. (2001). Yesterday's polyploids and the mystery of diploidization. Nat. Rev. Genet. *2*, 333–341.
5. Woodhouse, M.R., Schnable, J.C., Pedersen, B.S., Lyons, E., Lisch, D., Subramaniam, S., and Freeling, M. (2010). Following tetraploidy in maize, a short deletion mechanism removed genes preferentially from one of the two homologs. PLoS Biol. *8*, e1000409.
6. Barker, M.S., Kane, N.C., Matvienko, M., Kozik, A., Michelmore, R.W., Knapp, S.J., and Rieseberg, L.H. (2008). Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. Mol. Biol. Evol. *25*, 2445–2455.
7. Coate, J.E., Schlueter, J.A., Whaley, A.M., and Doyle, J.J. (2011). Comparative evolution of photosynthetic genes in response to polyploid and nonpolyploid duplication. Plant Physiol. *155*, 2081–2095.
8. Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. Annu. Rev. Plant Biol. *60*, 433–453.
9. Paterson, A.H., Chapman, B.A., Kissinger, J.C., Bowers, J.E., Feltus, F.A., and Estill, J.C. (2006). Many gene and domain families have convergent fates following independent whole-genome duplication events in *Arabidopsis, Oryza, Saccharomyces* and *Tetraodon*. Trends Genet. *22*, 597–602.
10. Thomas, B.C., Pedersen, B., and Freeling, M. (2006). Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. Genome Res. *16*, 934–946.
11. Symonds, V.V., Soltis, P.S., and Soltis, D.E. (2010). Dynamics of polyploid formation in Tragopogon (Asteraceae): recurrent formation, gene flow, and population structure. Evolution *64*, 1984–2003.
12. Birchler, J.A., Riddle, N.C., Auger, D.L., and Veitia, R.A. (2005). Dosage balance in gene regulation: biological implications. Trends Genet. *21*, 219–226.
13. Liu, B., Vega, J.M., and Feldman, M. (1998). Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. II. Changes in low-copy coding DNA sequences. Genome *41*, 535–542.
14. Ozkan, H., Levy, A.A., and Feldman, M. (2002). Rapid differentiation of homeologous chromosomes in newly-formed allopolyploid wheat. Isr. J. Plant Sci. *50*, S65–S76.
15. Kashkush, K., Feldman, M., and Levy, A.A. (2002). Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics *160*, 1651–1659.
16. Gaeta, R.T., Pires, J.C., Iniguez-Luy, F., Leon, E., and Osborn, T.C. (2007). Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell *19*, 3403–3417.
17. Szadkowski, E., Eber, F., Huteau, V., Lodé, M., Huneau, C., Belcram, H., Coriton, O., Manzanares-Dauleux, M.J., Delourme, R., King, G.J., et al. (2010). The first meiosis of resynthesized *Brassica napus*, a genome blender. New Phytol. *186*, 102–112.
18. Mestiri, I., Chagué, V., Tanguy, A.M., Huneau, C., Huteau, V., Belcram, H., Coriton, O., Chalhoub, B., and Jahier, J. (2010). Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity. New Phytol. *186*, 86–101.
19. Song, K., Lu, P., Tang, K., and Osborn, T.C. (1995). Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc. Natl. Acad. Sci. USA *92*, 7719–7723.
20. Shaked, H., Kashkush, K., Ozkan, H., Feldman, M., and Levy, A.A. (2001). Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. Plant Cell *13*, 1749–1759.
21. Nie, L., Han, Z., Lu, L., Yao, Y., Sun, Q., and Ni, Z. (2008). Genomic and genic sequence variation in synthetic hexaploid wheat (AABBDD) as compared to their parental species. Prog. Nat. Sci. *18*, 533–538.
22. Mecchia, M.A., Ochogavía, A., Pablo Selva, J., Laspina, N., Felitti, S., Martelotto, L.G., Spangenberg, G., Echenique, V., and Pessino, S.C. (2007). Genome polymorphisms and gene differential expression in a 'back-and-forth' ploidy-altered series of weeping lovegrass (*Eragrostis curvula*). J. Plant Physiol. *164*, 1051–1061.
23. Soltis, D.E., Soltis, P.S., Pires, J.C., Kovarik, A., Tate, J.A., and Mavrodiev, E. (2004). Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): cytogenetic, genomic and genetic comparisons. Biol. J. Linn. Soc. *82*, 485–501.
24. Gaeta, R.T., and Chris Pires, J. (2010). Homoeologous recombination in allopolyploids: the polyploid ratchet. New Phytol. *186*, 18–28.
25. Tate, J.A., Joshi, P., Soltis, K.A., Soltis, P.S., and Soltis, D.E. (2009). On the road to diploidization? Homeolog loss in independently formed populations of the allopolyploid *Tragopogon miscellus* (Asteraceae). BMC Plant Biol. *9*, 80.
26. Tate, J.A., Ni, Z.F., Scheen, A.C., Koh, J., Gilbert, C.A., Lefkowitz, D., Chen, Z.J., Soltis, P.S., and Soltis, D.E. (2006). Evolution and expression of homeologous loci in *Tragopogon miscellus* (Asteraceae), a recent and reciprocally formed allopolyploid. Genetics *173*, 1599–1611.
27. Buggs, R.J.A., Doust, A.N., Tate, J.A., Koh, J., Soltis, K., Feltus, F.A., Paterson, A.H., Soltis, P.S., and Soltis, D.E. (2009). Gene loss and silencing in *Tragopogon miscellus* (Asteraceae): comparison of natural and synthetic allotetraploids. Heredity *103*, 73–81.
28. Anssour, S., Krügel, T., Sharbel, T.F., Saluz, H.P., Bonaventure, G., and Baldwin, I.T. (2009). Phenotypic, genetic and genomic consequences of natural and synthetic polyploidization of *Nicotiana attenuata* and *Nicotiana obtusifolia*. Ann. Bot. *103*, 1207–1217.
29. Martelotto, L.G., Ortiz, J.P.A., Stein, J., Espinoza, F., Quarin, C.L., and Pessino, S.C. (2007). Genome rearrangements derived from autopolyploidization in *Paspalum* sp. Plant Sci. *172*, 970–977.
30. Koonin, E.V., and Wolf, Y.I. (2006). Evolutionary systems biology: links between gene evolution and function. Curr. Opin. Biotechnol. *17*, 481–487.
31. Lim, K.Y., Soltis, D.E., Soltis, P.S., Tate, J.A., Matyasek, R., Srubarova, H., Kovarik, A., Pires, J.C., Xiong, Z., and Leitch, A.R. (2008). Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (Asteraceae). PLoS ONE *3*, e3353.
32. Chester, M., Gallagher, J.P., Symonds, V.V., Silva, A.V.C.d., Mavrodiev, E.V., Leitch, A.R., Soltis, P.S., and Soltis, D.E. (2012). Extensive chromosomal variation in a recently formed natural allopolyploid species, *Tragopogon miscellus* (Asteraceae). Proc. Natl. Acad. Sci. USA, in press.
33. Al-Shahrour, F., Minguez, P., Marqués-Bonet, T., Gazave, E., Navarro, A., and Dopazo, J. (2010). Selection upon genome architecture: conservation of functional neighborhoods with changing genes. PLoS Comput. Biol. *6*, e1000953.
34. Osbourn, A. (2010). Gene clusters for secondary metabolic pathways: an emerging theme in plant biology. Plant Physiol. *154*, 531–535.
35. Hurst, L.D., Pál, C., and Lercher, M.J. (2004). The evolutionary dynamics of eukaryotic gene order. Nat. Rev. Genet. *5*, 299–310.
36. Chen, W.H., de Meaux, J., and Lercher, M.J. (2010). Co-expression of neighbouring genes in Arabidopsis: separating chromatin effects from direct interactions. BMC Genomics *11*, 178.
37. Tate, J.A., Symonds, V.V., Doust, A.N., Buggs, R.J.A., Mavrodiev, E., Majure, L.C., Soltis, P.S., and Soltis, D.E. (2009). Synthetic polyploids of *Tragopogon miscellus* and *T. mirus* (Asteraceae): 60 Years after Ownbey's discovery. Am. J. Bot. *96*, 979–988.
38. Buggs, R.J.A., Chamala, S., Wu, W., Gao, L., May, G.D., Schnable, P.S., Soltis, D.E., Soltis, P.S., and Barbazuk, W.B. (2010). Characterization of duplicate gene evolution in the recent natural allopolyploid *Tragopogon miscellus* by next-generation sequencing and Sequenom iPLEX MassARRAY genotyping. Mol. Ecol. *19* (*Suppl 1*), 132–146.
39. Saldanha, A.J. (2004). Java Treeview—extensible visualization of microarray data. Bioinformatics *20*, 3246–3248.