DOI: 10.1534/genetics.104.033522

Organ-Specific Silencing of Duplicated Genes in a Newly Synthesized Cotton Allotetraploid

Keith L. Adams,[1] Ryan Percifield and Jonathan F. Wendel

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

Manuscript received July 16, 2004
Accepted for publication September 7, 2004

ABSTRACT

Most eukaryotes have undergone genome doubling at least once during their evolutionary history. Hybridization followed by genome doubling (allopolyploidization) is a prominent mode of speciation in plants, leading to phenotypic novelty and changes in genome structure and gene expression. Molecular events that take place immediately after polyploid formation can be studied using newly synthesized allopolyploids. Here we studied the extent of gene silencing in a newly created and genomically stable allotetraploid cotton, of genotype AAGG, using an AFLP-cDNA display screen. Over 2000 transcripts were screened and ~5% of the duplicated genes in the allotetraploid were inferred to have been silenced or downregulated. Sequencing of 24 AFLP-cDNA fragments revealed genes with a variety of functions. Analysis by RT-PCR showed silencing or a strong expression bias toward one copy for 9 of 13 genes examined. Comparisons of expression patterns among eight organs in the allopolyploid showed that silencing and preferential expression are organ specific. Examination of silencing patterns in two other synthetic polyploids, of genotype AADD, showed that the same gene can be silenced independently in different genotypes. These results provide a detailed portrayal of gene silencing events that can occur following allopolyploidization and suggest epigenetic causal factors.

GENOME doubling has occurred during the evolutionary history of many animals, fungi, and plants, resulting from either duplication of a single genome (autopolyploidy) or combination of two or more differentiated genomes (allopolyploidy) (*e.g.*, Masterson 1994; Nadeau and Sankoff 1997; Wolfe and Shields 1997; Wolfe 2001). Polyploidy is particularly prevalent in plants, where it is an active, ongoing process and a prominent mode of speciation in many groups. Even in groups where modern polyploidization is relatively uncommon, genetic and genomic studies reveal that the evolutionary history of many angiosperms includes multiple rounds of polyploidization (Vision *et al.* 2000; Simillion *et al.* 2002; Blanc *et al.* 2003; Bowers *et al.* 2003; Blanc and Wolfe 2004). Thus, even plants with relatively small genomes, like *Arabidopsis thaliana,* have experienced historical genome doubling. Polyploidization has been shown to cause short-term morphological innovation and novelty (*e.g.*, Stebbins 1950; Levin 1983; Bretagnolle and Lumaret 1995; Comai *et al.* 2000; Schranz and Osborn 2000, 2004) and in addition, the attendant genome doubling provides a massive reservoir of duplicated genes for longer-term evolution of new functions (Stebbins 1950; Stephens 1951; Ohno 1970; Lewis 1980; Force *et al.* 1999; Martin 1999; Wendel 2000; Lynch 2002).

To model the process of allopolyploid formation in nature, newly created synthetic polyploids may be used. Such polyploids are created by hybridizing two species and then doubling the chromosomes. Synthetic polyploids may reveal molecular phenomena that occur immediately upon allopolyploidization. Studies of synthetic Brassica, wheat, and Arabidopsis allopolyploids have shown rapid and nonrandom genomic changes, including sequence elimination and alternations in DNA cytosine methylation (Song *et al.* 1995; Liu *et al.* 1998; Madlung *et al.* 2002; Levy and Feldman 2004). Gene silencing and downregulation of duplicated protein-coding genes (Comai *et al.* 2000; Kashkush *et al.* 2002; He *et al.* 2003) and rRNA genes (Chen *et al.* 1998; Joly *et al.* 2004) have been revealed from studies of synthetic wheat, Arabidopsis, and Glycine allopolyploids. Some silencing events in wheat allotetraploids were inferred to be caused by gene loss (Kashkush *et al.* 2002).

Cotton (Gossypium) has been developed as a particularly useful group for studies of polyploidy (reviewed in Wendel and Cronn 2003; Adams and Wendel 2004). Diploid cottons are classified into eight genome groups based initially on cytology and refined by molecular systematic studies (reviewed in Wendel and Cronn 2003). There are two major branches of Gossypium species, one comprising the New World, D-genome group and the other containing all other genome groups, variously distributed in the Old World (Figure 1). Polyploidiza-

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AY685662–AY685681.

[1]*Corresponding author:* UBC Botanical Garden and Centre for Plant Research, University of British Columbia, MacMillan Bldg., 2357 Main Mall, Vancouver, BC V6T 1Z4, Canada. E-mail: keith.adams@ubc.ca

tion between an A-genome diploid and a D-genome diploid ~1.5 MYA (Senchina *et al.* 2003) created the AD allotetraploid lineage that includes the commercially important *Gossypium hirsutum* (upland cotton) and *G. barbadense* (pima cotton). A number of synthetic cotton allopolyploids have been created (Brubaker *et al.* 1999) that are useful for studying the immediate effects of allopolyploid formation. Previous studies of natural and synthetic cotton polyploids (reviewed in Adams and Wendel 2004), have demonstrated interlocus concerted evolution of rRNA repeats, unequal rates of sequence evolution between the two genomes in the polyploid nucleus, and independent evolution of many protein-coding genes.

In an earlier study of the natural allopolyploid *G. hirsutum* (Adams *et al.* 2003), we showed that for several pairs of homeologs (genes doubled by polyploidy), expression levels of the two copies (hereafter A_t and D_t, referencing the respective diploid genome donors A and D) varied extensively among multiple organs. Most notably, the alcohol dehydrogenase gene *adhA* has been silenced in an organ-specific manner such that only the D_t copy is expressed in petals and stamens, whereas only the A_t copy is expressed in stigmas and styles. Additionally, two genes were shown to display organ-specific silencing and expression biases in a several-generation-old synthetic allotetraploid. These data suggested that extreme biases in duplicate gene expression that are organ specific may accompany allopolyploid formation. However, neither the scale of the phenomenon nor its rapidity of onset during or following allopolyploidization is known.

In this study we used a more global approach (AFLP-cDNA display) to survey the extent of gene silencing in the first generation of a newly created cotton allotetraploid (designated AAGG). For this analysis we selected an allotetraploid line that is known to be genomically additive and stable relative to its diploid parents (Liu *et al.* 2001). Accordingly, biases in gene expression and putative silencing events are unlikely to reflect gene deletions or genomic rearrangements. We show that allopolyploidy induces an immediate and widespread effect on duplicate gene expression that varies by organ type. For a few genes with interesting expression patterns, we extended the analysis to include two additional synthetic allopolyploids of genotype AADD; similarities in silencing patterns among these different genotypes suggests a common underlying mechanistic basis of gene expression alteration.

MATERIALS AND METHODS

Plant materials and RNA extractions: The maternal genome (A) donor *G. arboreum* (accession no. 5265), paternal genome (G) donor *G. bickii* (accession no. 5048), and the derived synthetic AAGG allotetraploid $2(A_2G_1)$ (also designated Hyb-612; Brubaker *et al.* 1999) in the first generation after chromosome doubling were used. Also included were a recently created AADD allotetraploid $2(A_2D_3)$ (gift of J. A. Lee), this generated by colchicine doubling a sterile diploid hybrid formed between the A-genome species *G. arboreum* (A_2) and the D-genome species *G. davidsonii* (D_3), and a synthetic AADD allotetraploid $2(A_2D_1)$ of unknown generation (Beasley 1940) that was also used in Adams *et al.* (2003), with models of its diploid parents, *G. arboreum* (A_2) and *G. thurberi* (D_1). All plants were grown in a greenhouse under common conditions.

For RNA extractions the following floral organs were collected on the morning of flower opening (anthesis) from multiple plants: bracts (epicalyx), sepals, petals, whole stamens, stigmas and styles, ovary walls, and ovules. Young leaves were collected from mature plants. RNA was extracted using a hot borate method as described in Adams *et al.* (2003). RNA quality was inferred by running a small aliquot on a formaldehyde agarose gel. RNAs were treated with the DNA-free kit (Ambion, Austin, TX) to remove residual DNA, and RNA concentrations were estimated using a spectrophotometer. Poly(A)$^+$ RNA was isolated from the total RNA using the PolyATtract kit (Promega, Madison, WI) according to the manufacturer's instructions.

AFLP-cDNA: Poly(A)$^+$ RNA was reverse transcribed using the Retroscript kit (Ambion) and oligo(dT) as a primer to prime off the poly(A) tails of messenger RNAs. Second-strand cDNA synthesis was done using RNaseH (New England Biolabs, Beverly, MA) and DNA polymerase I (Promega) and standard procedures. Multifluorophore AFLP analysis was performed as in Liu *et al.* (2001) using the double-stranded cDNA templates. Enzyme combinations used were as follows: *Eco*RI (labeled) and *Mse*I, *Pst*I (labeled) and *Mse*I, *Eco*RI (labeled) and *Pst*I, and *Pst*I (labeled) and *Eco*RI. Sequences of primers and adapters are given in supplemental Table 1 at http://www.genetics.org/supplemental/. AFLP-cDNA results were visualized using GeneScan software.

AFLP-cDNA and genomic AFLP reactions for manual gels were performed identically to the fluorescent reactions except that primers were labeled with γ-^{32}P instead of fluorophores. Reaction products were run on 6% acrylamide gels and visualized by autoradiography. Bands of interest were cut from dried gels that were carefully marked to ensure that the appropriate band was isolated. DNA was eluted by boiling in 100 μl of water for 5 min. AFLP fragments were amplified by PCR and either sequenced directly or cloned into a PCR-cloning vector followed by sequencing of multiple clones. Sequencing was done at the Iowa State University DNA Sequencing and Synthesis Facility using ABI 3700 DNA sequencers.

RT-PCR and sequence analysis: Reverse transcription was done as described in Adams *et al.* (2003). As controls for DNA contamination, reactions were also performed without reverse transcriptase (RT−), side by side with experimental reactions. One-twentieth of the cDNAs created by first-strand synthesis were used in PCR reactions with 0.5 μM each primer, 2.5 mM $MgCl_2$, and Taq DNA polymerase. Reaction volumes were 30 μl and cycling was done in a MJ PTC-100 thermocycler for 2 min at 94° followed by 30 cycles of 30 sec at 94°, 30 sec at 52°–61°, 1 min at 72°, and then a final 6-min extension at 72°. Primers used for RT-PCR are listed in supplemental Table 2 at http://www.genetics.org/supplemental/. RT-PCR products were sequenced directly using an ABI 3700 DNA sequencer. Chromatogram trace files were analyzed using BioEdit to compare levels of expression from each homeolog using the method of Rauscher *et al.* (2002): The height of each peak was measured using the coordinates system. The relative contribution of the A homeolog (when present) was calculated as the height of the A peak over the sum of the heights of the A and G peaks; likewise the contribution of the G homeolog was calculated as the G peak height over the sum of the A and G heights. Background-corrected averages across all polymorphic sites were reported for each gene. This method has been shown

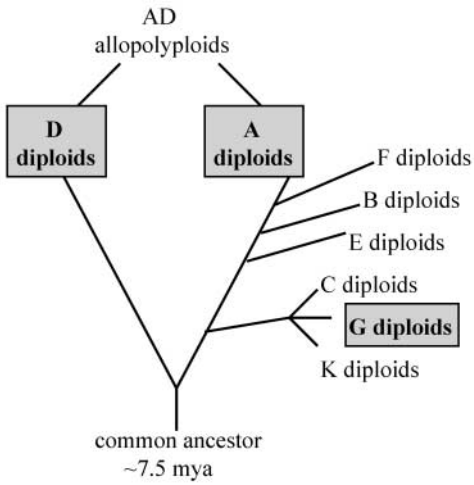

FIGURE 1.—Evolutionary history of the cotton genus (Gossypium), showing the diploids and natural AADD allotetraploid. Genome groups used in this study are shaded. The synthetic allotetraploids were made from AA and GG diploids and from AA and DD diploids.

to accurately determine ratios of homeologs in a mixture of PCR products consisting of both homeologs (RAUSCHER *et al.* 2002).

RESULTS

A survey of duplicate gene silencing upon allopolyploidization: To survey the extent of homeologous gene silencing upon allopolyploid formation in cotton, we performed an AFLP-cDNA display screen using RNA from a newly created AAGG allotetraploid formed between the two diploids *G. arboreum* (A-genome) and *G. bickii* (G-genome; Figure 1). This AAGG allotetraploid shows very low levels and possibly no genomic change following polyploidization (LIU *et al.* 2001; our unpublished data). All AFLP reactions were done in sets of four: the maternal A-genome parent, the paternal G-genome parent, the derived allotetraploid, and a mixture of equal amounts of RNA from the two parents. Two biological replicates (RNA extractions from different plants) were done for each set. Bands were scored only if the AFLP pattern was the same for each set. Four different restriction enzyme combinations and many selective nucleotide combinations were used (see MATERIALS AND METHODS). The initial organ type used for the AFLP-cDNA screen was petals collected on the day of anthesis. AFLP-cDNA reactions were done initially using fluorophore-labeled primers (TET and FAM), run on automated sequencing gels, and analyzed by GeneScan analysis. Only qualitative differences—band presence or absence—were scored.

Five classes of AFLP-cDNA patterns were scored (Figure 2) and tabulated, as follows:

1. A band present in each lane, indicating no polymorphisms between the parents in the restriction enzyme recognition sites or the sites of the selective nucleotides in the AFLP primers. Such loci are uninformative with respect to the expression status of each of the two homeologs in the polyploid, in that band presence may reflect transcripts from either (or both) gene copy.
2. A band present in all lanes except the A-genome parent, caused by a polymorphism between the two parents and indicating that the paternal G_t copy in the polyploid is expressed.
3. A band present in all lanes except the G-genome parent, caused by a polymorphism between the two parents and indicating that the maternal A_t copy in the polyploid is expressed.
4. A band present in only the A and A + G lanes, suggesting that the A_t copy in the polyploid was silenced.
5. A band present in only the G and A + G lanes, suggesting that the G_t copy in the polyploid was silenced.

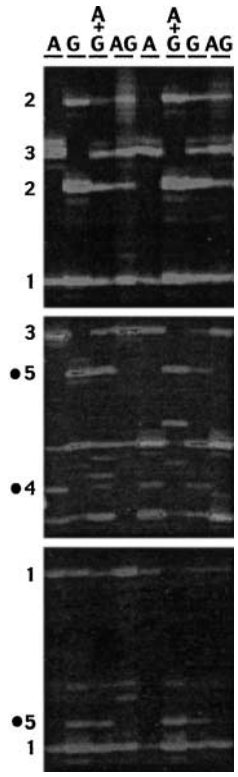

FIGURE 2.—AFLP-cDNA gels. Example fragments of gels run on the ABI 377 with fluorophore-labeled primers are shown. Bands of each class are indicated: (1) a band present in each lane that cannot be scored for the expression status in the polyploid; (2) a band present in all lanes except the A parent, indicating that the maternal G_t copy in the polyploid is expressed; (3) a band present in all lanes except the G parent, indicating that the paternal A_t copy in the polyploid is expressed; (4) a band present in only the A and A + G lanes, suggesting that the A_t copy in the polyploid was silenced; and (5) a band present in only the G and A + G lanes, suggesting that the G_t copy in the polyploid was silenced. Class 4 and 5 bands are highlighted with bullets.

TABLE 1

AFLP-cDNA summary

	Parental origin		
	AA	GG	Total
Petal			
Total bands			2015
Bands polymorphic between parents	318	206	524 (26%)
Bands not present in AG allotetraploid	15	23	38 (7%)
Six-organ comparison			
Total bands			530
Bands polymorphic between parents	85	58	143 (27%)
Bands not present in AG allotetraploid	5	6	11 (8%)

There were no cases of a band being present in all lanes except the AG lanes, which would suggest silencing of both homeologs, nor were there any cases of a band being present in only the AG lanes, which would be suggestive of gene activation following allopolyploidization. Only polyadenylated mRNA transcripts were surveyed; thus transposon activation would likely not be detected.

A total of 2015 cDNA fragments (presumptive loci) were screened using RNAs from petals, with 524 (26%) loci being polymorphic between the parents and therefore informative with respect to the expression status in the polyploid (patterns 2–5 above; Table 1). Thirty-eight bands (7% of the 524 diagnostic loci) were absent in the polyploid lanes (patterns 4 and 5 above), suggesting silencing of either the A_t or the G_t copy. Of the 318 bands that were present in the A lanes but not present in the G lanes and thus diagnostic for expression of the A_t copy in the allotetraploid (patterns 3 and 4 above), 15 (5%) were also not present in the AG allotetraploid, suggesting silencing of the A_t copy. Of the 206 bands that were present in the G lanes but not present in the A lanes and thus diagnostic for expression of the G_t copy in the allotetraploid (patterns 2 and 5 above), 23 (11%) were also not present in the AG allotetraploid, suggesting silencing of the G_t copy. Thus, >50% more of the surveyed genes from the G_t (paternal) genome (11%) may well have been silenced in the allotetraploid compared to those from the A_t (maternal) genome (5%).

To detect genes showing coexpression of both homeologs in petals but silencing of one copy in other organs, the AFLP-cDNA screen was extended to RNAs from seven additional organ types: sepals, stamens, stigmas and styles, ovary walls, ovules, bracts (epicalyx), and leaves. *Eco*RI/*Pst*I-digested cDNAs were used from each organ type along with 10 sets of selective primer combinations that were randomly chosen. A total of 530 loci were screened using these methods, of which 143 were polymorphic between the two parents and thus informative as to the expression status in the allotetraploid. Bands were absent in the polyploid lanes for 11 loci (8% of the polymorphic loci) in at least one organ type (Table 1). Of the 85 bands that were present in the A lanes but not present in the G lanes and thus diagnostic for expression of the A_t copy in the allotetraploid (patterns 3 and 4 above), 5 (6%) were also not present in the AG allotetraploid, suggesting silencing of the A_t copy. Of the 58 bands that were present in the G lanes but not present in the A lanes and thus diagnostic for expression of the G_t copy in the allotetraploid (patterns 2 and 5 above), 6 (10%) were also not present in the AAGG allotetraploid, suggesting silencing of the G_t copy. Thus, as with the petal data, ~50% more genes from the G_t (paternal) genome may well have been silenced in the allotetraploid compared to those from the A_t (maternal) genome.

Identification and verification of putatively silenced genes: Several sets of AFLP-cDNA reactions that revealed one or more putatively silenced loci were performed with γ-^{32}P-labeled primers and run on polyacrylamide gels. For each locus of interest, two bands (one from each replicate) were excised from the gel and the AFLP fragments were sequenced (see MATERIALS AND METHODS). A total of 24 loci were identified representing 23 genes (Table 2). Coincidentally, 1 gene, an ADP-glucose pyrophosphorylase, was identified twice using different restriction enzyme combinations. Seventeen of the 23 genes were inferred to have a homolog in Arabidopsis. The genes have a variety of functions or putative functions (Table 2), including metabolic enzymes, kinases, and DNA-binding proteins. Six loci had no homolog in any other plant; most of these correspond to short AFLP-cDNA fragments (100–150 bp) and contained stop codons in all possible reading frames, suggesting that they represent fragments of the 5′-UTR or 3′-UTR.

The AFLP-cDNA data were scored on the basis of visual inspection and thus the data are not quantitative: cases of major downregulation of one homeolog may have been scored as silencing; likewise some of these cases may have been missed due to the presence of a faint band. Furthermore, AFLP-cDNA screens are prone to revealing false positives (*i.e.*, genes inferred to be silenced that actually are expressed). Thus it is important to further characterize expression of putatively silenced genes identified in an AFLP-cDNA display screen. To accomplish this we studied expression of both homeologs for 13 of the identified genes, to determine if one homeolog was silenced, if one homeolog is downregulated, or if there is equal expression and the AFLP-cDNA-based interpretations were erroneous.

To conduct this analysis, 300–600 bp of sequence was first determined from both diploid parents so that polymorphic sites could be identified to facilitate diagnosis of gene expression in the synthetic allotetraploid.

TABLE 2

Sequenced AFLP-cDNA bands

Sequence name	Putative silenced copy	Sequence ID	Similarity amino acid to *A. thaliana* (%)/*E* value	*A. thaliana* GenBank no.	Cotton EST
Em1938	G	Glycoside hydrolase family	60/5*e*-77[a]	NP_567055.1	CO080954
Em2025	A	Flavin-containing monooxygenase	41/2*e*-31	AAF82235.1	CO090813
Em130	G	Choline-phosphate cytidyltransferase	80/*e*-113[a,b]	BAA09642.1	AJ513345
Em145	G	None			None
Em440	A	Putative Ser/Thr protein kinase	58/1*e*-12	AAL24403.1	GH__DEa06H01.f
Em490	A	DNA-binding protein related	57/5*e*-44	CAB87767.1	CO105816
Em396	G	CBL-interacting protein kinase 23	91/1*e*-88	AAK61494.1	AI725753
Em450	G	Splicing factor	80/7*e*-08[a]	AAD52610.1	BF274242
Em460	G	None			None
Em118	G	Expressed protein	70/5*e*-70	NP_568689.1	AW187538
Ep330	G	Calcineurin-like phosphoesterase	44/6*e*-05	NP_175246.1	None
Ep550	G	Zinc finger protein (DHHC type)	76/1*e*-22	NP_187148.2	CO070176
Ep603	A	Expressed protein in Arabidopsis	77/2*e*-70	NP_187365.1	CO108271
Em700	G	ADP-glucose pyrophosphorylase	91/3*e*-48	NP_177629.1	AI726504
Pm122	G	None			None
Pm100	A	None			None
Pm118	G	PHD finger family	50/5*e*-26	NP_177903.4	BF271623
Pm125	A	Calmodulin-binding family	78/2*e*-11	NP_190797.1	BG445035
Pm130	A	None			None
Ep280	A	Ubiquitin-protein ligase 2	63/3*e*-18	NP_177189.1	None
Ep445	G	ADP-glucose pyrophosphorylase	91/2*e*-63	NP_849886.1	AI726504
Pe170	G	None			None
Pe300	G	Kelch repeat containing F-box family	50/0.001[a]	NP_172885.1	BF268638
Ep420	A	Phox domain containing protein	52/3*e*-18	NP_172976.2	None

A and G represent the copies in the allotetraploid derived from the A (maternal) parent or G (paternal) parent.
[a] Hit using corresponding cotton EST sequence.
[b] *Brassica napus*.

By using homology to cotton expressed sequence tags (ESTs), the region sequenced could be lengthened for some short AFLP fragments (Table 2). Sequence divergence between the A-genome and G-genome parental sequences ranged from 0.5 to 3.5%, and the number of polymorphisms for each gene ranged from 3 to 14.

For each of these 13 genes, expression of the A_t and G_t homeologs was quantified as follows: The gene region was amplified by RT-PCR using primers that were perfect matches to both parental copies, and the resulting RT-PCR products were directly sequenced. At each polymorphic site the peaks were evaluated by the method of Rauscher *et al.* (2002) to assess the contribution of each homeologous transcript to the total RT-PCR pool. Nine of 13 examined genes showed silencing or considerable expression bias—a ratio of at least 2:1—of one copy in the organ type where silencing was inferred on the basis of AFLP-cDNA display (petals for all genes except Ep420-ovule and Ep280-stigma and style). Four of nine genes (Em2025, Pm125, Em700, and Ep280) showed silencing of one homeolog, and five genes showed preferential expression toward one copy (Table 3). The silenced genes included a monooxygenase, calmodulin-binding protein, ubiquitin-protein ligase, and ADP-glucose pyrophosphorylase. Both homeologs from four genes (Em396, Em450, Em130, and Ep603) appear to be expressed at comparable levels in petals.

Overall, 9 of 13 genes interpreted from AFLP-cDNA screens as exhibiting silencing were confirmed as having one homeolog either silenced or strongly downregulated; the corollary is that 4 of 13 (or 31%) of the putative silencing events were artifacts of the AFLP screening methodology. This level of artifacts is consistent with or lower than those obtained from AFLP-cDNA screens in previous studies of polyploids. Using this false positive rate and 40 loci scored as silenced (of 524 diagnostic AFLP-cDNA fragments), we infer that ~5% ($9/13 \times 40/524$) of the duplicated loci in allopolyploid cotton experienced immediate silencing (~2.25%) or downregulation of one homeolog (~2.75%) upon polyploidization.

Duplicate gene silencing is organ specific: To determine if there is organ-specific silencing or biased expression of homeologous gene pairs upon allopolyploidization, relative homeolog expression levels of eight gene pairs in eight organ types were assayed by RT-PCR, direct sequencing, and analysis of polymorphic sites as above. Two biological replicates were done for most genes. Four genes showed silencing of one homeolog in some organs but not in others (Table 3; Figure 3). Particularly interesting is gene Em2025 (a flavin-containing mono-

TABLE 3

Expression and silencing patterns of homeologous gene pairs in AAGG allotetraploid

Gene	Function/putative function	Petal	Sepal	Stamen	Stigma and style	Ovary wall	Ovule	Bract	Leaf
Em2025	Flavin-containing monooxygenase	A silen	A silen	A silen	A silen	A silen	A > G	A silen	A < G
Pm125	Calmodulin-binding protein	G silen	G silen	A > G	A > G	G silen	G silen	A = G	A < G
Ep550	DHHC-type zinc finger protein	A > G	G silen	G silen	A > G	G silen	G silen	A > G	A > G
Ep280	Ubiquitin-protein ligase	A = G	A < G	A < G	A silen	A silen	A silen	A silen	ND
Em700	ADP-glucose pyrophosphorylase	G silen	A = G	A = G	A = G	A = G	A = G	A = G	ND
Em490	DNA-binding protein	A < G	A = G	A = G	A = G	A = G	A = G	A = G	A = G
Em440	Ser/Thr protein kinase	A < G	A = G	A = G	A < G	A = G	A = G	A = G	ND
Ep420	Phox domain containing protein	A = G	A < G	A = G	A < G	A < G	A < G	A < G	ND
Em118	Expressed protein	A > G	A = G	A > G	A = G	A = G	A = G	ND	A = G

A and G represent the copies in the allotetraploid derived from the A (maternal) parent or D (paternal) parent. Homeologous transcript ratios above 2:1 are indicated by > and <. silen, gene silencing; ND, not determined.

oxygenase) that shows silencing of the A_t copy (maternal copy) in most examined organs except the leaves where the A_t copy is also expressed and ovules where there

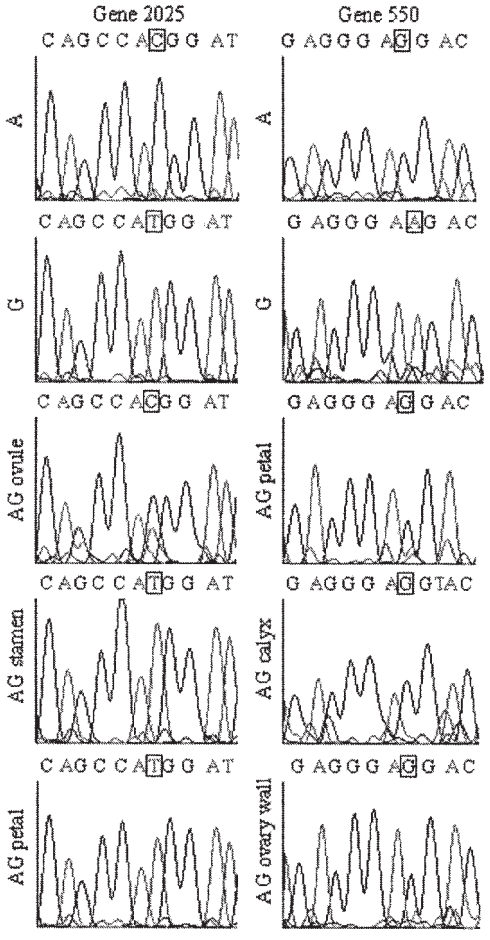

FIGURE 3.—Sequencing chromatograms. Shown are sample regions of the chromatograms from RT-PCR products of gene Em2025 (monooxygenase) in the first column and gene Ep550 (a DHHC-type zinc finger protein) in the second column. Bases polymorphic between the A and G parents are marked with boxes.

actually is preferential expression of the A_t homeolog. Gene Pm125 (a calmodulin-binding protein) shows homeolog silencing in some organs and expression that is biased toward one homeolog in leaves and toward the other homeolog in stamens and in stigmas and styles. The two copies of gene Ep280 (ubiquitin protein ligase) showed a gradient in expression level ratio, ranging from equal expression in petals to A_t silencing in bracts and all three parts of the carpels. Also showing organ-specific silencing of one homeolog (Table 3) are genes Ep550 (a DHHC-type zinc finger protein) and Em700 (an ADP-glucose pyrophosphorylase). Four genes did not show silencing in any organ type, but in some organs there was preferential expression of one copy.

To determine if the cases of gene silencing were caused by allopolyploidization or if they reflect expression of the gene in only one of the two parents in the organ type in question, RT-PCR was performed using RNAs from the diploid parents for each gene and organ type that showed silencing of one copy in the allotetraploid. Controls without reverse transcriptase were done to ensure that amplification was not due to DNA contamination, and some of the genes contain an intron(s). In all cases there was expression of each gene in both parents in the organ type in question (data not shown), indicating that gene silencing occurred as a result of the allopolyploidization process.

Parallel silencing patterns among genotypes: Expression of three genes that showed silencing in at least one organ of the AAGG allotetraploid was assayed in two additional synthetic allotetraploids (between A- and D-genome diploids; genotype AADD) to determine if the same patterns of silencing would be observed in independently generated and genomically divergent allopolyploids. Each of the three allopolyploids has the A-genome species *G. arboreum* as the maternal parent, but they differ with respect to the paternal genome donor. For each gene studied, the sequence of the paternal copy (D_1 and D_3) was determined and compared to the *G. arboreum* sequence to identify polymorphisms, as described above

TABLE 4

Expression and silencing patterns of homeologous gene pairs in synthetic AD allotetraploids

	Gene Em2025			Gene Ep550			Gene Pm125		
Organ	AA2DD3	AA2DD1	AAGG	AA2DD3	AA2DD1	AAGG	AADD3	AADD1	AAGG
Sepal	A silen	A silen	A silen	A < D	A = D	G silen	A = D	A > D	G silen
Petal	A silen	A silen	A silen	A < D	A silen	A > G	A = D	A > D	G silen
Stamen	A silen	A silen	A silen	A < D	A < D	G silen	A = D	A > D	A > G
Stigma/style	A silen	A silen	A silen	A < D	A < D	A > G	A = D	A = D	A > G
Ovary wall	A silen	A silen	A silen	A = D	A = D	G silen	A = D	A = D	G silen
Ovule	A silen	A < D	A > G	A = D	A = D	G silen	A = D	A = D	G silen

A and D represent the copies in the allotetraploid derived from the A (maternal) or D (paternal) parent. Data for the AAGG allotetraploid are listed for comparison. silen, gene silencing.

(only two polymorphisms were observed for genes Ep280 and Em700, and thus these genes were excluded from analysis). RT-PCR was performed using RNAs from seven organs, using two biological replicates for each organ and primers that are prefect matches across the genomes, to avoid PCR amplification bias. RT-PCR products were directly sequenced and analyzed as above.

These experiments revealed some striking similarities across independently synthesized allotetraploids. Most notable is gene Em2025 (a monooxygenase), where the A_t copy was silenced in almost every organ of both AADD allopolyploids, mirroring the patterns observed in the AAGG allotetraploid (Table 4; Figure 4). Thus, the same copy (A_t) of the same gene was silenced in three independently created synthetic allopolyploids of different genotype. Expression of the homeologous pairs of gene Ep550 (a zinc finger protein) ranged from silencing of A_t to coexpression at equal levels, depending on the organ type. There was considerable correspondence in expression patterns between the two AADD allotetraploids, although divergent patterns of homeolog expression were observed in the AAGG allotetraploid (Table 4). Gene Pm125 (a calmodulin-binding protein) showed a contrast between the AAGG polyploid (silencing or biased expression in all organs) and both AADD polyploids (largely or completely equal ratios).

DISCUSSION

Frequency of gene silencing in plant polyploids: The AFLP-cDNA data, combined with follow-up RT-PCR experiments, in this study suggest that silencing or downregulation affects ~5% of duplicated genes in the newly created Gossypium allopolyploid. This value is consistent with studies of synthetic wheat and Arabidopsis allopolyploids (Comai *et al.* 2000; Kashkush *et al.* 2002; He *et al.* 2003); exact comparisons are difficult to make due to varying numbers of AFLP-cDNA fragments surveyed and somewhat differing methodologies used for verification of gene silencing in each study. Our results show that although the AG-genome cotton allotetraploid is genomically stable, gene silencing (inferred to affect ~2.25% of the genes in the allopolyploid) occurs at levels comparable to synthetic allopolyploids characterized by more frequent gene losses and genomic rearrangements. Unexpectedly, about twice as many genes from the paternal G-genome (11%) as from the maternal A-genome (5%) in the allotetraploid cotton were inferred from the AFLP-cDNA gel data to be silenced or downregulated.

The amount of gene silencing or downregulation (5% of surveyed genes) in the newly synthesized cotton allotetraploid used in this study contrasts with a higher value (25%) observed in a previous study of 40 genes in the

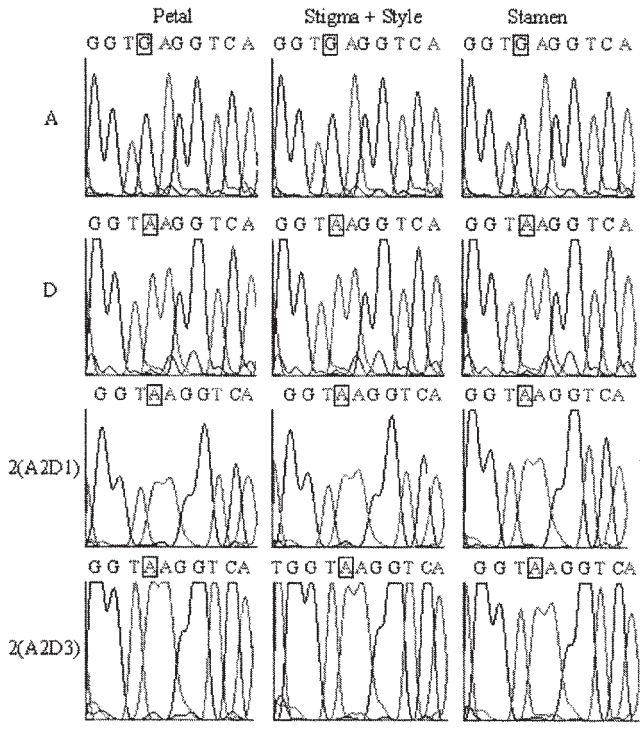

Figure 4.—Sequencing chromatograms. Shown are sample regions of the chromatograms from RT-PCR products of gene Em2025 (monooxygenase) from the A and D parents and the two AD allotetraploids. A base polymorphic between the A and D parents is marked with boxes.

natural allotetraploid *G. hirsutum* (Adams *et al.* 2003). Two factors could account for this difference. The natural cotton allotetraploid was formed 1–2 MYA and thus there has been much more time for changes to occur in gene expression. Second, the AFLP-cDNA screen used in this study was scored on the basis of presence or absence of bands and the method (by itself) is not quantitative and is less sensitive to detection of homeologs that are expressed at unequal levels, in contrast to the SSCP assays used in Adams *et al.* (2003).

Organ-specific silencing of genes duplicated by polyploidy: In addition to quantifying the phenomenon of homeolog silencing after polyploidization in cotton, this study shows that silencing of duplicated genes immediately after allopolyploid formation is organ specific, confirming previous speculations based on data from an allopolyploid that was several generations old (Adams *et al.* 2003). All genes that showed silencing in this study, except for the A_t copy of Em2025 in AA2DD3, were silenced in only some organ(s) and expressed in other organ(s). Considerable expression differences exist among floral whorls, differences that would have been obscured if whole flowers had been used in the experiments. Furthermore, three groups of organs that compose the carpels (ovule, ovary wall, and stigma and style) showed differences in expression *vs.* silencing of homeologous genes. For example, gene Em2025 (a monooxygenase) showed silencing of the A_t copy in the ovary wall and stigma and style, but expression of both copies (albeit not at equal levels) in the ovule. We note that the biased expression patterns of homeologs documented here and previously (Adams *et al.* 2003) were detected despite the fact that the organs sampled are composed of multiple tissues and cell types; that is, even at the relatively crude level of analysis of whole organs, preferential expression between homeologs is observed. The implication is that finer scale analyses of individual tissues and cell types would reveal even a higher level of partitioning of duplicate gene expression than that reported here.

The cases of gene silencing reported in this study were determined to have been postpolyploidization by examining transcript presence in the diploid parents. Preferential accumulation of transcripts from one of two homeologs may also reflect a change in expression following polyploidization, either upregulation of one copy or downregulation of the other copy, or it could reflect retention of parental expression patterns that happened to be different. Likewise, equal expression levels of homeologs in a polyploid could represent no change following polyploidization or it could reflect adjustments in expression level (up- or downregulation) of one or both homeologs to equalize expression. Although transcript presence was assayed in the diploids, accumulation levels were not; hence, distinguishing the foregoing possibilities will require further experimentation.

This study and previous work (Adams *et al.* 2003) have established that organ-specific silencing of homeologous genes can occur upon allopolyploidization and continue over evolutionary time in the natural polyploids. It will be interesting to determine if the phenomenon can be caused solely by the merger of two divergent genomes (hybridization), by examining diploid F_1 hybrids prior to chromosome doubling. We hypothesize that both hybridization and chromosome doubling can serve as causal factors for organ-specific silencing of homeologous genes. It also is of interest to gain a better understanding of the temporal scale of the phenomenon; that is, what are the relative proportions of expression alteration that accompany the earliest stages of polyploid formation (including hybridization) *vs.* those that accrue on an evolutionary timescale following polyploidization, perhaps on the order of millions of years. Toward that end, direct comparisons on a global scale of natural allopolyploid and corresponding synthetic allopolyploids should prove informative.

Stochastic *vs.* directed gene silencing: Are some duplicated genes more prone to silencing than others? To address this question we examined expression and silencing patterns in three independently created synthetic allotetraploid lines of different genotypes. Among the more striking observations was the repeated silencing of gene Em2025 (a monooxygenase gene), which exhibited silencing in most or all organs of all three synthetic allopolyploids. Interestingly, the silenced copy was from the maternal (A-genome) parent in all three lines. Also showing similar silencing patterns between the two synthetic AADD allopolyploid lines is the *adhA* gene, where there is minimal or no transcription of one copy in certain floral organs (our unpublished data), and these patterns are almost identical to those observed in the natural polyploid *G. hirsutum* (Adams *et al.* 2003). The above results suggest that silencing of some genes is directed and that chromosomal context can play a role in duplicate gene silencing. Although phenomenologically similar to nucleolar dominance, any of the molecular mechanisms discussed below are potentially responsible for these observations. Evidence for this type of "directed" gene silencing extends beyond the cotton study system. In hexaploid wheat, He *et al.* (2003) examined expression of two genes in three genotypes that are expressed in only one of the parents. They found silencing or downregulation of both genes in two different synthetic AABBDD genotypes as well as the natural Chinese spring wheat.

Patterns of gene silencing across synthetic allopolyploids need not be similar, however. In this study genes Ep550 and Pm125, for example, displayed a more random pattern of silencing and biased expression across all three genotypes (Table 4). Other evidence for this more "stochastic" pattern of gene silencing comes from studies of rRNA and protein-coding genes in synthetic allopolyploid Arabidopsis (Chen *et al.* 1998; Comai *et al.* 2000). Two recent studies are particularly relevant in this regard: Wang *et al.* (2004) showed variation in

silencing patterns of individual genes among newly synthesized allotetraploid Arabidopsis lines, and Joly *et al.* (2004) showed that different accessions of natural polyploid Glycine species exhibit varying expression levels of progenitor rRNA genes. Additional multigenotype comparisons will shed further light on this issue.

Causes and mechanisms of homeologous gene silencing: The gene silencing documented in the newly synthesized allotetraploid cotton is almost certainly epigenetic in nature. This allopolyploid cotton line is genomically stable and additive with respect to its diploid parents, as previously judged by >2000 AFLP markers (Liu *et al.* 2001) and as confirmed here, so genome rearrangements and gene loss are either minimal or nonexistent. Given these observations and what is generally understood about mutation rates, it is unlikely that novel nucleotide substitutions or insertion/deletion events have played a significant role in generating the expression alterations observed in newly created allopolyploids. Hence, epigenetic causes (*sensu lato*) are implicated, including an array of potentially interacting and reinforcing mechanisms, such as hypermethylation of DNA cytosines (Lee and Chen 2001; discussed in Osborn *et al.* 2003); deacetylation, methylation, or other modifications of histones; and positional effects from higher-order changes in chromatin structure (reviewed in Wolffe and Matzke 1999; Richards and Elgin 2002; Liu and Wendel 2003). The latter may include a diverse suite of poorly understood architectural requirements necessitated by the packaging of a suddenly doubled complement of chromosomes in a single nucleus. One can envision, for example, that the genomic spectrum and efficiency of scaffold/matrix attachment regions in an allopolyploid nucleus is rather different from that of a single diploid progenitor; these are known to play an important role in chromatin structure and gene expression (Rudd *et al.* 2004). Specific models for homeologous gene silencing have been proposed involving repeats and LTRs of retroelements (Comai *et al.* 2003). An elegant study in wheat demonstrated that antisense transcripts generated by readout transcription of a retrotransposon caused silencing of adjacent genes (Kashkush *et al.* 2003). It is also possible that small RNAs and RNA interference play a role in gene silencing in polyploids (see also Comai *et al.* 2003). The foregoing factors are not mutually exclusive and mechanisms will probably vary by gene.

From a physiological and evolutionary perspective, genes in polyploids may be silenced or downregulated for various reasons, including dosage (Guo *et al.* 1996; Osborn *et al.* 2003). Interactions of diverged regulatory hierarchies that have been reunited in an allopolyploid may also be responsible for alterations in gene expression (Osborn *et al.* 2003; Riddle and Birchler 2003). Additionally, if homeologous gene sequences have different amino acid sequences, one homeolog might interact better with other proteins in multisubunit complexes (Comai 2000; Adams and Wendel 2004). For example, all of the maternal genes whose sequences have coadapted to each other might be expressed and the paternal copies silenced. These cases are likely to be visible to natural selection and are predicted to be evolutionarily stabilized. In contrast, some expression variation may be functionally and selectively neutral, reflecting instead a side effect of mechanistic processes, such as those discussed in the previous paragraph. Many of these cases are likely to be retained over longer time frames and become exposed to an evolutionary filter only after additional epigenetic and genetic evolution. The scale of the phenomenon of expression alteration, as reported here, suggests that this process will prove to be a significant facet of polyploid evolution.

We thank Curt Brubaker for supplying seed of the synthetic AAGG allopolyploid line and its diploid parents and Josh Lee for his gift of the $2(A_2D_3)$ allopolyploid line and its diploid parents. We thank Josh Moloney and Nevin Sebastian for excellent technical assistance. Keith Adams is postdoctoral fellow of the Jane Coffin Childs Memorial Fund for Medical Research and the Fund aided this investigation. Financial support for this research was provided by a U.S. Department of Agriculture grant to Jonathan Wendel.

LITERATURE CITED

Adams, K. L., and J. F. Wendel, 2004 Exploring the genomic mysteries of polyploidy in cotton. Biol. J. Linn. Soc. **82:** 573–582.

Adams, K. L., R. Cronn, R. Percifield and J. F. Wendel, 2003 Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proc. Natl. Acad. Sci. USA **100:** 4649–4654.

Beasley, J. O., 1940 The origin of American tetraploid *Gossypium* species. Am. Nat. **74:** 285–286.

Blanc, G., and K. H. Wolfe, 2004 Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell **16:** 1667–1678.

Blanc, G., K. Hokamp and K. H. Wolfe, 2003 A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. Genome Res. **13:** 137–144.

Bowers, J. E., B. A. Chapman, J. Rong and A. H. Paterson, 2003 Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature **422:** 433–438.

Bretagnolle, F., and R. Lumaret, 1995 Bilateral polyploidization in *Dactylis glomerata* L. subsp. *lusitanica*: occurrence, morphological and genetic characteristics of first polyploids. Euphytica **84:** 197–207.

Brubaker, C. L., A. H. D. Brown, J. M. Stewart, M. J. Kilby and J. P. Grace, 1999 Production of fertile hybrid germplasm with diploid Australian *Gossypium* species for cotton improvement. Euphytica **108:** 199–213.

Chen, Z. J., L. Comai and C. S. Pikaard, 1998 Gene dosage and stochastic effects determine the severity and direction of uniparental rRNA gene silencing (nucleolar dominance) in *Arabidopsis* allopolyploids. Proc. Natl. Acad. Sci. USA **95:** 14891–14896.

Comai, L., 2000 Genetic and epigenetic interactions in allopolyploid plants. Plant Mol. Biol. **43:** 387–399.

Comai, L., A. P. Tyagi, K. Winter, R. Holmes-Davis, S. H. Reynolds *et al.*, 2000 Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. Plant Cell **12:** 1551–1567.

Comai, L., A. Madlung, C. Joseffson and A. Tyagi, 2003 Do the different parental "heteromes" cause genomic shock in newly formed allopolyploids? Philos. Trans. R. Soc. Lond. B **358:** 1149–1155.

Force, A., M. Lynch, F. B. Pickett, A. Amores, Y.-L. Yan *et al.*, 1999 Preservation of duplicate genes by complementary, degenerative mutations. Genetics **151:** 1531–1545.

Guo, M., D. Davis and J. A. Birchler, 1996 Dosage effects on gene expression in a maize ploidy series. Genetics **142:** 1349–1355.

He, P., B. Friebe, B. Gill and J.-M. Zhou, 2003 Allopolyploidy alters gene expression in the highly stable hexaploid wheat. Plant Mol. Biol. **52:** 401–414.

Joly, S., J. Rauscher, S. Sherman-Broyles, A. Brown and J. Doyle, 2004 Evolutionary dynamics and preferential expression of homeologous 18S–5.8S–26S nuclear ribosomal genes in natural and artificial *Glycine* allopolyploids. Mol. Biol. Evol. **21:** 1409–1421.

Kashkush, K., M. Feldman and A. A. Levy, 2002 Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics **160:** 1651–1659.

Kashkush, K., M. Feldman and A. A. Levy, 2003 Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nat. Gen. **33:** 102–106.

Lee, H.-S., and Z. J. Chen, 2001 Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids. Proc. Natl. Acad. Sci. USA **98:** 6753–6758.

Levin, D. A., 1983 Polyploidy and novelty in flowering plants. Am. Nat. **122:** 1–25.

Levy, A. A., and M. Feldman, 2004 Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization. Biol. J. Linn. Soc. **82:** 607–614.

Lewis, W. H. (Editor), 1980 *Polyploidy: Biological Relevance*. Plenum, New York.

Liu, B., and J. F. Wendel, 2003 Epigenetic phenomena and the evolution of plant allopolyploids. Mol. Phylogenet. Evol. **29:** 365–379.

Liu, B., J. M. Vega and M. Feldman, 1998 Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. II. Changes in low-copy coding DNA sequences. Genome **41:** 535–542.

Liu, B., C. L. Brubaker, G. Mergeai, R. C. Cronn and J. F. Wendel, 2001 Polyploid formation in cotton is not accompanied by rapid genomic changes. Genome **44:** 321–330.

Lynch, M., 2002 Genomics—gene duplication and evolution. Science **297:** 945–947.

Madlung, A., R. W. Masuelli, B. Watson, S. H. Reynolds, J. Davison *et al.*, 2002 Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. Plant Physiol. **129:** 733–746.

Martin, A. P., 1999 Increasing genomic complexity by gene duplication and the origin of vertebrates. Am. Nat. **154:** 111–128.

Masterson, J., 1994 Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science **264:** 421–424.

Nadeau, J. H., and D. Sankoff, 1997 Comparable rates of gene loss and functional divergence after genome duplications early in vertebrate evolution. Genetics **147:** 1259–1266.

Ohno, S., 1970 *Evolution by Gene Duplication*. Springer-Verlag, New York.

Osborn, T. C., J. C. Pires, J. A. Birchler, D. L. Auger, Z. J. Chen *et al.*, 2003 Understanding mechanisms of novel gene expression in polyploids. Trends Genet. **19:** 141–147.

Rauscher, J. T., J. J. Doyle and A. H. D. Brown, 2002 Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploid complex. Mol. Ecol. **11:** 2691–2702.

Richards, E., and S. Elgin, 2002 Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell **108:** 489–500.

Riddle, N. C., and J. A. Birchler, 2003 Effects of reunited diverged regulatory hierarchies in allopolyploids and species hybrids. Trends Genet. **19:** 597–600.

Rudd, S., M. Frisch, K. Grote, B. C. Meyers, K. Mayer *et al.*, 2004 Genome-wide in silico mapping of scaffold/matrix attachment regions in *Arabidopsis* suggests correlation of intragenic scaffold/matrix attachment regions with gene expression. Plant Physiol. **135:** 715–722.

Schranz, M. E., and T. C. Osborn, 2000 Novel flowering time variation in the resynthesized polyploid *Brassica napus*. J. Hered. **91:** 242–246.

Schranz, M. E., and T. C. Osborn, 2004 De novo variation in life-history traits and responses to growth conditions of resynthesized polyploid *Brassica napus* (Brassicaceae). Am. J. Bot. **91:** 174–183.

Senchina, D. S., I. Alvarez, R. C. Cronn, B. Liu, J. Rong *et al.*, 2003 Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. Mol. Biol. Evol. **20:** 633–643.

Simillion, C., K. Vandepoele, M. C. E. Van Montagu, M. Zabeau and Y. Van de Peer, 2002 The hidden duplication past of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **99:** 13627–13632.

Song, K., P. Lu, K. Tang and T. C. Osborn, 1995 Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc. Natl. Acad. Sci. USA **92:** 7719–7723.

Stebbins, G. L., 1950 *Variation and Evolution in Plants*. Columbia, New York.

Stephens, S. G., 1951 Possible significance of duplication in evolution. Adv. Genet. **4:** 247–265.

Vision, T. J., D. G. Brown and S. D. Tanksley, 2000 The origins of genomic duplications in *Arabidopsis*. Science **290:** 2114–2117.

Wang, J., L. Tian, A. Madlung, H.-S. Lee, M. Chen *et al.*, 2004 Stochastic and epigenetic changes of gene expression in Arabidopsis polyploids. Genetics **167:** 1961–1973.

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

Wendel, J. F., and R. C. Cronn, 2003 Polyploidy and the evolutionary history of cotton. Adv. Agron. **78:** 139–186.

Wolfe, K. H., 2001 Yesterday's polyploids and the mystery of polyploidization. Nat. Rev. Genet. **2:** 333–341.

Wolfe, K. H., and D. C. Shields, 1997 Molecular evidence for an ancient duplication of the entire yeast genome. Nature **387:** 708–713.

Wolffe, A. P., and M. A. Matzke, 1999 Epigenetics: regulation through repression. Science **286:** 481–486.

Communicating editor: J. A. Birchler