Epigenetic phenomena and the evolution of plant allopolyploids

Bao Liu*a and Jonathan F. Wendelb,*

*a Institute of Genetics and Cytology, Northeast Normal University, Changchun 130024, China
b Department of Botany, 353 Bessey Hall, Iowa State University, Ames, IA 50011, USA

Received 24 December 2002; revised 8 January 2003

Abstract

Allopolyploid speciation is widespread in plants, yet the molecular requirements for successful orchestration of coordinated gene expression for two divergent and reunited genomes are poorly understood. Recent studies in several plant systems have revealed that allopolyploid genesis under both synthetic and natural conditions often is accompanied by rapid and sometimes evolutionarily conserved epigenetic changes, including alteration in cytosine methylation patterns, rapid silencing in ribosomal RNA and protein-coding genes, and de-repression of dormant transposable elements. These changes are inter-related and likely arise from chromatin remodeling and its effects on epigenetic codes during and subsequent to allopolyploid formation. Epigenetic modifications could produce adaptive epimutations and novel phenotypes, some of which may be evolutionarily stable for millions of years, thereby representing a vast reservoir of latent variation that may be episodically released and made visible to selection. This epigenetic variation may contribute to several important attributes of allopolyploidy, including functional diversification or subfunctionalization of duplicated genes, genetic and cytological diploidization, and quenching of incompatible inter-genomic interactions that are characteristic of allopolyploids. It is likely that the evolutionary success of allopolyploidy is in part attributable to epigenetic phenomena that we are only just beginning to understand.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Allopolyploidy; Genome evolution; DNA methylation; Gene silencing; Transposable element; Epigenetics

1. Introduction

Polyploidy is an important evolutionary process in plants and some animals (Friedman and Hughes, 2001; Grant, 1981; Gu et al., 2002; Masterson, 1994; McLyshaght et al., 2002; Otto and Whitton, 2000; Soltis and Soltis, 1999). The abundance of allopolyploid plant species in nature suggests a selective advantage conferred by allopolyploids over diploid progenitors, and has implicated polyploidy as an important speciation process. Most proposed explanations for the fitness advantage of polyploids invoke either some form of gene redundancy and the attendant release from functional constraint for one copy and/or divergence leading to novel functions (Harland, 1936; Ohno, 1970), or potentially adaptive genome-wide allelic and/or non-allelic interactions leading to heterozygosity and heterosis (Allard et al., 1993; Pikaard, 2002). Although several aspects of allopolyploidy have been intensively studied over the past century, we still understand relatively little about how two divergent genomes coordinately adjust and evolve to guide growth and development. Yet the mere prevalence of polyploidy constitutes compelling evidence that such adjustments occur and that some fraction of them have positive fitness consequences. Thus, the notion that “polyploidy has contributed little to progressive evolution” (Stebbins, 1971) has been replaced in recent years by a consensus view that polyploidy is a prominent and pervasive force in plant evolution (Leitch and Bennett, 1997; Otto and Whitton, 2000; Soltis and Soltis, 2000; Wendel, 2000).

This altered thinking about polyploidy has collectively been inspired by recent molecular studies conducted in several plant systems, including *Brassica* (Song et al., 1995), wheat (Feldman et al., 1997; Kashkush et al., 2002; Liu et al., 1997, 1998a,b; Ozkan et al., 2001; Shaked et al., 2001), *Arabidopsis* (Comai et al., 2000; Lee...
and Chen, 2001; Madlung et al., 2002), and cotton (Hanson et al., 1999; Hanson et al., 1998; Jiang et al., 1998; Wendel et al., 1995; Zhao et al., 1998). The emerging notion is that allopolyploid genomes are “dynamic” (Soltis and Soltis, 1995) at the molecular level, generating an array of novel genomic instabilities or changes during the initial stages of polyploid formation or over longer time-spans. Some of these alterations are not readily explained by Mendelian principles, but may nonetheless have contributed to the evolutionary success of allopolyploids (Finnegan, 2001a; Liu and Wendel, 2002; Osborn et al., 2003; Pikaard, 2002; Rieseberg, 2001; Soltis and Soltis, 1995; Wendel, 2002; Liu and Wendel, 2000). Consequently, allopolyploidy might well represent a condition whereby evolution is accelerated and fitness is enhanced. Of particular significance are recent findings (Table 1) that epigenetic mechanisms may be involved in the stabilization and evolution of nascent allopolyploids (Finnean, 2001a). Thus, the epigenetic arena may be one in which part of the mystery of the evolutionary success of allopolyploids may be unveiled.

Epigenetics refers to heritable alterations in gene expression that do not entail changes in nucleotide sequence, but which nevertheless may have phenotypic and hence evolutionary consequences. Epigenetic effects can be accomplished by several self-reinforcing and inter-related covalent modifications on DNA and/or chromosomal proteins, such as DNA methylation and histone modifications, and by chromatin remodeling, such as repositioning of nucleosomes. These heritable modifications are collectively termed “epigenetic codes” (reviewed in Richards and Elgin, 2002). Ample evidence has established that programmed global epigenetic control of gene expression is essential to ensure normal growth and development (Finnegan, 2001b; Robertson, 2001; Robertson and Wolfe, 2000; Wolfe and Matzke, 1999), and because of this the epigenetic arena is a vibrant field in current biological research. Given this fundamental importance in diverse organisms, including fungi, plants and animals, it seems appropriate and timely to discuss the possible connections between epigenetics and the still largely mysterious processes of evolution in allopolyploids.

The objective of the present synthesis is to bring additional awareness to this new and apparently fruitful research direction by summarizing recent findings on various aspects of epigenetic phenomena and mechanisms that likely are relevant to nascent and/or natural allopolyploidy. In the process, we attempt to explicate some potential causes of observed epigenetic changes in plant allopolyploids. In addition, we discuss the possible biological relevance of epigenetic modifications to phenotypic innovation, and thereby to attributes that may be essential to the evolutionary success of newly formed allopolyploid lineages.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Phenomena and references</th>
<th>Phenomenon</th>
<th>Plant material</th>
<th>Plant material</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brassica</td>
<td>DNA methylation pattern changes in anonymous genomic and cDNAs (Song et al., 1995)</td>
<td>DNA methylation pattern changes in anonymous genomic and cDNAs (Song et al., 1995)</td>
<td>Synthetic and natural allotetraploids</td>
<td>Brassica oleracea</td>
<td>(Song et al., 1995)</td>
</tr>
<tr>
<td>Wheat (Triticum-Aegilops)</td>
<td>Nucleolar dominance (Chen and Pikaard, 1997a, Pikaard, 1999; Pikaard, 2000a; Pikaard, 2000b)</td>
<td>Nucleolar dominance (Chen and Pikaard, 1997a, Pikaard, 1999; Pikaard, 2000a; Pikaard, 2000b)</td>
<td>Synthetic and natural allotetraploids</td>
<td>Triticum-Aegilops</td>
<td>(Chen and Pikaard, 1997a, Pikaard, 1999; Pikaard, 2000a; Pikaard, 2000b)</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>DNA methylation pattern changes (Kashkush et al., 2002; Liu et al., 1998a; Liu et al., 1998b; Liu et al., 1998c)</td>
<td>DNA methylation pattern changes (Kashkush et al., 2002; Liu et al., 1998a; Liu et al., 1998b; Liu et al., 1998c)</td>
<td>Synthetic allotetraploids between Arabidopsis thaliana and Cardaminopsis arenosa</td>
<td>Arabidopsis thaliana</td>
<td>(Kashkush et al., 2002; Liu et al., 1998a; Liu et al., 1998b; Liu et al., 1998c)</td>
</tr>
<tr>
<td>Cotton (Gossypium)</td>
<td>Rapid epigenetic gene silencing (Kashkush et al., 2002)</td>
<td>Rapid epigenetic gene silencing (Kashkush et al., 2002)</td>
<td>Natural and synthetic allotetraploids</td>
<td>Gossypium</td>
<td>(Kashkush et al., 2002)</td>
</tr>
<tr>
<td>Triticale</td>
<td>Rapid epigenetic gene silencing (Kashkush et al., 2002)</td>
<td>Rapid epigenetic gene silencing (Kashkush et al., 2002)</td>
<td>Synthetic allopolyploids</td>
<td>Triticale</td>
<td>(Kashkush et al., 2002)</td>
</tr>
<tr>
<td>Rice (Oryza sativa)</td>
<td>Rapid epigenetic gene silencing (Kashkush et al., 2002)</td>
<td>Rapid epigenetic gene silencing (Kashkush et al., 2002)</td>
<td>Synthetic and natural allotetraploids</td>
<td>Oryza sativa</td>
<td>(Kashkush et al., 2002)</td>
</tr>
</tbody>
</table>
2. Methylation repatterning

An integral component of the developmental control of gene expression and the maintenance of genome integrity in a diverse array of organisms is specific, programmed cytosine methylation (Finnegan, 2001b; Finnegan et al., 1998, 2000; Richards, 1997). Hypermethylation is usually a hallmark of heterochromatin and is characteristic of euchromatic gene silencing, whereas hypomethylation is often associated with active gene expression (Martienssen and Colot, 2001; Nö and Bird, 1999). In mammals, although cytosine methylation patterns are dynamic at the whole genome level, as exemplified by erasure and resetting of cytosine methylation in early development (reviewed in Reik et al., 2001), intrinsically defined patterns are essential for normal development. For instance, in mutations in any of the three known DNA methyltransferase (DMTase) genes is lethal during early embryonic stages (reviewed in Geiman and Robertson, 2002). Also, aberrant methylation patterns in promoters of tumor suppressor genes, involving both genome-wide hypomethylation and local hypermethylation, often characterize human tumorgenesis (Jones et al., 2002). Also, aberrant methylation patterns in promoters of tumor suppressor genes, involving both genome-wide hypomethylation and local hypermethylation, often characterize human tumorgenesis (Jones et al., 2002).

In plants, cytosine methylation patterns usually are stably maintained through meiosis and over generations. Experimental disruption of cytosine methylation patterns in plants, though often yielding viable plants, may have conspicuous effects on morphology. This has been demonstrated in Arabidopsis, where reduction in DNA methylation levels through ectopic expression of an antisense DMTase (MET1), or by knock-out mutations in DDM1 (a gene encoding a putative chromatin remodeling factor, whose mutation causes genome-wide demethylation), result in pleiotropic effects on morphology and development and reduced female fertility (Finnegan, 2001b; Finnegan et al., 1998, 2000).

It has been recognized for years that unusual environmental stimuli and passage through tissue culture may cause heritable changes in cytosine methylation patterns in plants (Jablonska and Lamb, 1989; Kaeppler et al., 2000). As a potential primary genome defense system (Yoder et al., 1997), the cytosine methylation machinery may respond to environmental or genomic challenges by causing alterations in methylation that are thought to mediate physiologically meaningful responses to the challenge. Allopolyploidy, by unifying divergent genomes into one nucleus, may constitute such a challenge, or “genomic shock” (sensu McClintock, 1984). This suggestion is supported by experimental evidence, which shows that in several nascent allopolyploid plants, including Brassica (Song et al., 1995), wheat (Kashkush et al., 2002; Liu et al., 1998a; Shaked et al., 2001) and Arabidopsis (Comai et al., 2000; Lee and Chen, 2001; Madlung et al., 2002), allopolyploid formation leads to heritably re-patterned cytosine DNA methylation.

In synthetic Brassica allopolyploids, a low frequency of apparently random DNA methylation changes, including both hypo- and hypermethylation, were shown to occur at anonymous genomic loci and cDNAs, predominantly at CpG sites but also at CpNpG sites (Song et al., 1995). Similar changes in pattern and frequency to those in Brassica were observed in various synthetic wheat allopolyploids (Liu et al., 1998a; Shaked et al., 2001), but the alterations detected were non-random in the sense that the same changes sometimes occurred in multiple, arbitrarily selected individuals, and were even conserved between synthetic and natural plants (e.g., Fig. 1) (Liu et al., 1998a). An alternative interpretation is that some of these apparently directed methylation changes may represent random changes that happen to ameliorate fertility and thus are selected (Rieseberg, 1984).
An additional and potentially important insight into allopolyploidy-stimulated methylation repatterning became apparent from a study of wheat, where using a genome-wide fingerprinting approach, Shaked et al. (2001) showed that cytosine methylation alterations are not only genome-wide in distribution but may significantly differ in frequency between the two constituent genomes. For example, in first generation allotetraploid wheat, of the 11 bands that showed heritable methylation changes (occurred at the F1 hybrid stage and stably maintained after genome doubling), 10 bands are from one parent genome and only one band is from the other parent. Finally, genome-wide and non-random changes in DNA methylation patterns are also observed in synthetic allotetraploid Arabidopsis and Cardaminopsis arenosa, although in this case the total level of CpG methylation remains constant in the allopolyploids and their parents (Madlung et al., 2002).

Given the importance of DNA methylation to gene expression, changes in cytosine methylation resulting from genome merger and/or polyploid formation clearly could have genome-wide epigenetic consequences of relevance to polyploid evolution. This possibility is strengthened by results that have established a close link among various epigenetic codes. Specifically, changes in cytosine methylation usually are consequences rather than causes of other epigenetic modifications (Geiman and Robertson, 2002; Pikaard and Lawrence, 2002). For example, in Neurospora, cytosine methylation in all sequence contexts, CpG, CpNpG and asymmetric, depends on the occurrence of histone H3-lysine9 methylation (Tamaru and Selker, 2001). Similarly, in Arabidopsis, at least some cytosine methylations (CpNpG and asymmetric) are also directed by this histone modification, via the action of plant-specific methyltransferases like chromomethylase (CMT) and/or domains rearranged methylase (DRM) (Jackson et al., 2002). This, together with the correlation between DNA methylation and histone deacetylation (Dobosy and Selker, 2001), suggests that allopolyploidy-induced genomic cytosine methylation changes are likely a manifestation of only part of an interrelated network of global chromatin remodeling and/or histone modifications, probably with variable effects on constituent parental genomes. Thus, it is not difficult to imagine consequences, perhaps including gene silencing and activation as well as transcriptional de-repression and mobilization of transposons, as discussed below.

3. Epigenetic gene silencing

Nascent allopolyploids often are associated with variation and instability in phenotypes that cannot be accounted for by conventional Mendelian transmission genetics or chromosomal aberrations (Comai, 2000; Comai et al., 2000). The affected traits are diverse, including timing of flowering, overall plant habit, leaf morphology, and homeotic transformations in floral morphology (Comai et al., 2000; Schranz and Osborn, 2000). It has been suggested that these allopolyploidy-associated changes in phenotypes are the outcome of altered gene expression due to various causes, including increased variation in dosage-regulated gene expression, altered regulatory interactions, and rapid genetic and epigenetic changes, which are probably conferred by genome-wide interactions (reviewed in Osborn et al., 2003). As noted above, the union of divergent genomes into a single nucleus may disrupt intrinsic regional or global epigenetic controls, leading to genome-wide chromatin remodeling. Consequently, widespread alterations in gene expression might be expected. As demonstrated in several studies, both ribosomal RNA genes and protein-coding genes are subject to epigenetic-mediated modifications in newly formed and natural allopolyploid plants.

3.1. Nucleolar dominance

Nucleolar dominance refers to the phenomenon in hybrids or allopolyploids whereby nucleoli form, in association with ribosomal RNA genes, on chromosomes inherited from only one of the two parents (Pikaard, 1999, 2000a,b). Although this phenomenon had been intensively studied since its discovery in 1934 (Navashin, 1934), its molecular basis is not fully understood. Recently, the phenomenon was studied at the molecular level in natural and synthetic allopolyploids in two plant systems, Brassica (Chen and Pikaard, 1997a,b) and Arabidopsis (Chen et al., 1998). These studies provided the first empirical demonstration of a causal relationship between allopolyploidy, epigenetic modification, and change in gene expression. It was found that nucleolar dominance results from selective silencing of rRNA genes from the non-dominant genome through covalent chromatin modifications (Chen and Pikaard, 1997a,b). Results showed that rRNA transcripts from only one of the parental genomes were detectable in vegetative tissues of both natural and synthetic allopolyploids, thus indicating rapid occurrence of the phenomenon as well as its evolutionary conservation. More significantly, rRNA genes silenced in vegetative tissues are de-repressed in reproductive organs, indicating not only reversibility of the phenomenon, but also differential partitioning of rDNA array expression during allotetraploid plant development. Further experiments revealed that cytosine methylation and histone deacetylation act as partners in the enforcement of rRNA gene silencing in nucleolar dominance (Chen and Pikaard, 1997a). This suggests that the basis of nucleolar dominance in plants is likely the re-establishment of repressive chromatin states following chromatin remodeling induced by hybridization and allopolyploidy.
3.2. Rapid silencing of protein-coding genes

The seminal studies cited above on epigenetic modification and nuclear dominance of highly reiterated ribosomal genes raised the possibility that hybridization and allopolyploidization might similarly induce epigenetic modifications of protein-coding genes. This suspicion has now been confirmed in several model plant systems (Comai, 2000; Comai et al., 2000; Kashkush et al., 2002; Lee and Chen, 2001; Madlung et al., 2002). Two studies by Comai and colleagues (Comai et al., 2000; Madlung et al., 2002) on synthetic allotetraploids between Arabidopsis thaliana \( (2n = 2x = 26) \) and C. arenosa \( (2n = 4x = 32) \) demonstrated that allotetraploid formation caused rapid epigenetic gene silencing that affected a significant fraction (ca. 1%) of the transcriptome, including a variety of genes with different biological functions. At least some of the silencing events were associated with alterations in cytosine methylation status at specific sites, and possibly also were related to altered chromatin architecture. Remarkably, these expression alterations for synthetic allotetraploids are mimicked in the natural counterpart Arabidopsis suecica (Lee and Chen, 2001), formed by hybridization between A. thaliana and C. arenosa. The frequency of silenced genes since allopolyploidy is estimated to be at least 2.5% for this natural species. The silenced genes comprise a variety of RNAs and predicted proteins, including transcription factors and a transposase. An important observation was that silencing can be reversed by treatment with 5-aza-2’-deoxy-cytidine (aza-dC), a chemical specifically blocking the action of DNA methyltransferases and thus demethylation of the genome.

The studies discussed above indicate that speciation via allopolyploidy in Arabidopsis/Cardaminopsis is accompanied by epigenetic gene silencing, which may affect a variety of genes with diverse biological functions. The silencing events may occur rapidly (F2 generation of synthetic allopolyploid) or over a longer evolutionary time span, but their reversibility may be retained in natural allopolyploid species for thousands to perhaps millions of years. Of particular significance is the remarkable similarity or concordance in the silencing patterns between synthetic and natural allopolyploids, which suggests that allopolyploidy not only induces epigenetic changes but that the changes may be visible to natural selection, and judging from their persistence, adaptive.

The generality of the Arabidopsis findings with respect to other allopolyploid plant systems is largely unknown, as few comparable studies exist (Liu and Wendel, 2002). Nevertheless, earlier work by Feldman and coworkers (Feldman et al., 1986; Galili and Feldman, 1984) elegantly documented that in natural bread wheat (Triticum aestivum), a young allohexaploid species formed ca. 8500 years ago (Feldman, 2001), genes encoding endosperm storage proteins underwent “a massive and non-random” genetic diploidization, via either gene silencing or dosage compensation. Silencing of this set of genes was also found to occur rapidly in synthetic hexaploid wheat, and sometimes regain of expression of silenced alleles was observed when tetraploid plants were extracted from hexaploid wheat. This observation indicates that silencing is conferred by the addition of the third genome, that it is reversible, and hence, that it most likely has an epigenetic basis.

The scale of the phenomenon and hence its potential level of evolutionary importance is illustrated not just by results from Arabidopsis, but also from recent work involving synthetic and natural allopolyploid wheat and cotton (Gossypium). Using a cDNA-AFLP fingerprinting approach and a synthetic allopolyploid wheat analogous in genomic constitution to natural tetraploid wheat, Kashkush et al. (2002) showed that between 1% and 5% of the total transcriptome experienced silencing within the first generation in the new allotetraploid, and in addition, that novel transcripts were occasionally observed. Similar to the findings in Arabidopsis, epigenetic silencing in wheat is, at least in part, associated with methylation alteration at symmetrical cytosine sites, and a variety of genes with diverse biological functions are affected. Interestingly, all novel transcripts activated by polyploidy that could be assigned a putative function are retrotransposons, suggesting release from suppression (see Section 4).

Adams et al. (2003) used a novel SSCP approach to electrophoretically separate transcripts from the two homoeologues of 40 different genes duplicated by allopolyploidy in cotton. By measuring transcript accumulation in different organs and in both synthetic and natural allopolyploids, they showed a remarkably high level of transcription bias with respect to the duplicated copies, in that 25% of the genes studied exhibited altered expression in one or more organs. The most relevant and surprising result in the present context was the observation of developmentally regulated, organ-specific gene silencing that in some cases was reciprocal, meaning that one duplicate was expressed in one organ (e.g., stamens), while its counterpart was expressed in a different organ (e.g., carpels). Moreover, this organ-specific partitioning of duplicate expression was also evident in synthetic allopolyploids. While not directly shown to be caused by cytosine methylation, the results strongly implicate epigenetic phenomena. In addition, the similarity of silencing patterns observed in natural cotton allopolyploids, estimated to be approximately 1.5 million years old (Senchina et al., 2003; Wendel and Cronn, 2003), and newly formed synthetic cottons implicates either long-term evolutionary maintenance of the epigenetically induced expression state or its fixation through normal mutational processes. The
organ-specific nature of the observations is novel, but it recalls the phenomenon of nucleolar dominance in *Brassica* discussed above.

An additional implication of the *Gossypium* work (Adams et al., 2003) is that epigenetic silencing need not entail modifications in cytosine methylation. This speculation is based on data from a previous survey of over 1100 loci in each of nine sets of synthetic cotton allopolyploids using a methylation-sensitive AFLP approach, where there was no evidence of polyploidy-induced methylation alteration (Liu et al., 2001). These observations are most readily reconciled with the high frequency of biased expression and organ-specific gene silencing observed in allopolyploid cotton (Adams et al., 2003) by invoking other forms of epigenetic control that modulate gene expression, perhaps involving organ-specific alterations in chromatin folding patterns and position effects.

Collectively, studies of the last several years reveal that allopolyploid formation may be accompanied by epigenetic gene silencing that is genomically global and phylogenetically widespread. Moreover, these epigenetic changes may occur with the onset of polyploidy or accrue more slowly on an evolutionary time frame. In at least some cases, rapid epigenetic modifications that arise with the onset of allopolyploidy may be preserved on an evolutionary timescale through multiple speciation events, although it is possible that in the examples from cotton the correspondence in silencing patterns in the synthetic and natural allopolyploids is due to subsequent genetic mutations in the latter that fixed phenotypic states originating during allopolyploid formation.

4. De-repression of dormant transposable elements

Transposable elements (TEs) are mobile DNA sequences that are ubiquitous in all eukaryotes, and particularly abundant in plant genomes (Bennetzen, 2000; Kumar and Bennetzen, 1999). Recent studies provide evidence that TEs are not merely passive genomic inhabitants; instead, they may be a source of genetic diversity because they are capable of transposing in response to environmental and/or genetic cues, thereby modulating gene expression in host genomes (reviewed in Kidwell and Lisch, 2000; Kidwell and Lisch, 2001; Wessler, 2001). These attributes of TEs may be especially consequential in allopolyploids (Liu and Wendel, 2002).

Most TEs, particularly those with high copy numbers, are almost exclusively located in densely packaged and heavily methylated heterochromatic regions, and hence are inactive under normal conditions (Okamoto and Hirochika, 2001; Wessler, 1996). Nevertheless, some TEs can be activated under stress (Beguiristain et al., 2001; Bennetzen, 1996; Grandbastien, 1998; Hirochika et al., 1996; Wessler, 1996), perhaps due to disruption of tightly controlled repressive heterochromatin states.

McClintock predicted long ago that interspecific hybridization could potentially activate dormant TEs, which might cause genome restructuring (McClintock, 1984). Mobilized TEs most likely cause deleterious insertions, particularly under conditions whereby TEs lose their propensity to insert into non-genic heterochromatic regions, such as in the *Arabidopsis ddm1* mutant (Hirochika et al., 2000; Miura et al., 2001) and in tissue culture (Hirochika et al., 1996; Lucas et al., 1995). It thus is conceivable that under diploid conditions, e.g., in a diploid hybrid, enhanced TE activity is likely maladaptive. Polyploidy may be beneficial in this regard, because the harmful effects of TE activity may be buffered by genomic redundancy and hence insertions would be more likely to be tolerated (Matzke et al., 1999; Wendel, 2000).

Although a causal relationship between wide hybridization (including allopolyploidy) and TE activity remains to be thoroughly explored in plants, several lines of correlative evidence are consistent with McClintock’s predictions (reviewed in Comai, 2000; Liu and Wendel, 2002). Of particular relevance is the finding that various types of TEs are activated, at least transcriptionally, in the first generation of synthetic allotetraploid wheat plants (Shaked et al., 2001), and both transcriptionally and transpositionally in an intergeneric hybrid and its derivatives of rice and wild rice (*Zizania latifolia*) (unpublished data and Liu and Wendel, 2000).

Because TEs often contain regulatory sequences, their transcriptional activation may alter the activity of nearby genes, due to readthrough or readout transcription, as demonstrated over a decade ago in maize (Barkan and Martienssen, 1991). A compelling example of this phenomenon was recently provided by Kashkush et al. (2003), who showed that in newly synthesized wheat allopolyploids, mobilized TEs altered the expression of adjacent genes. It was found that in first-generation synthetic tetraploid wheat, the Wis 2-1A LTR retrotransposon family was transcriptionally activated. By a novel transcript-assay, i.e., transposon display on cDNAs, chimeric transcripts resulting from both readthrough and readout activities of LTRs were identified. The transcripts included both sense- and antisense-strands of adjacent genes, which led to gene activation and silencing, respectively. Of 360 genomic regions studied that flank members of the Wis 2-1A family, 26 showed altered expression patterns in the allopolyploid. Given the high copy number of this element (tens of thousands) and the possible activation of other TEs, the authors suggested that allopolyploidy-induced transcriptional activation of TEs may have genome-wide epigenetic effects on gene expression.
Additional evidence that wide hybridization causes TE activation comes from a study on wallabies (O’Neill et al., 1998, 1999), where interspecific hybridization between two wallaby species, *Mactopus eugenii* and *Wallabia bicolor*, leads to massive increase in copy number for the KERV-1 retrotransposon. Element activation is accompanied by genome-wide hypomethylation, suggesting epigenetic repression of the element in parental species and de-repression in the hybrid. Although such massive activation of TEs in a diploid hybrid should largely be maladaptive, as discussed above, in this specific case the reinserted elements exclusively targeted the centromeric regions of one parental genome, thus likely minimizing the impact on normal gene expression. On the other hand, as suggested by the authors, the dramatic extension of centromeric regions due to element insertion has led to rapid karyotypic evolution, and thus facilitated reproductive isolation of the hybrid from its parental species.

Although few natural plant hybrids and allopolyploids have been experimentally evaluated for TE activity, it is likely that wide hybridization and allopolyploidy causes activation of dormant TEs, probably as a result of compromised epigenetic, repressive control. The extent and tempo at which these events will occur seems undoubtedly varies among plant species and genome combinations. In general, however, the inherently higher level of tolerance to insertions makes it likely that TEs have played a greater role in genome evolution in allopolyploid than in diploid species.

5. Possible causes of allopolyploidy-induced epigenetic modification

Notwithstanding the ample recent evidence implicating a link between the two seemingly disparate phenomena of allopolyploidy and epigenetics, rather little is known about the underlying causes and mechanisms. In most cases of epigenetic phenomena associated with allopolyploid formation, alteration in cytosine DNA methylation patterns and gene silencing are observed (Comai et al., 2000; Kashkush et al., 2002; Lee and Chen, 2001; Liu and Wendel, 2002; Madlung et al., 2002; Pikaard, 1999, 2000a,b, 2002), but methylation has not always been implicated (Adams et al., 2003; Schranz and Osborn, 2000). Also, as noted in the introduction, methylation alterations may be a secondary effect rather than the proximate cause of epigenetic modification, which instead may have its mechanistic underpinnings in other components of the myriad interactions and processes that affect chromatin states. Given our present ignorance, what is it about genome merger and/or doubling that might provoke epigenetic modification of gene expression? Although a full answer to this question requires a more profound understanding than we are likely to achieve soon of higher order aspects of chromatin structure and its conformational modulations, as well as the myriad proteins and genes involved in mediating epigenetic responses, for the interim the following speculations may be forwarded.

5.1. Intergenomic interaction may directly cause changes in epigenetic gene silencing

A consensus finding from recent studies in plants, fungi and animals is that the presence of nucleic acid repeats (sequence redundancy) and their aberrant interactions (e.g., pairing) constitute a trigger for transcriptional gene silencing (reviewed in Bender, 1998). This process is closely associated with de novo cytosine methylation in promoters of the silenced genes (Vaucheret and Fagard, 2001). It is possible that in a newly formed allopolyploid, certain specific chromosomal regions or sequences are sensitive (perhaps due to unusual composition or structure) to duplication and are prone to interactions (e.g., somatic pairing), which could be perceived by cellular surveillance systems as aberrant. These regions may thereby be targeted by the DNA methylation machinery and related epigenetic silencing systems (Fig. 2).

Intergenomic interactions may also operate pre-meiotically. It was recently shown in *Neurospora crassa* that unpaired chromosomal segments in early meiosis are subjected to a novel type of epigenetic silencing, i.e., meiotic silencing by unpaired DNA (MSUD), whereby the unpaired sequence and all sequences homologous to it are silenced (Shiu and Metzenberg, 2002; Shiu et al., 2001). If a similar mechanism functions in plants, allopolyploid-induced meiotic non-pairing could conceivably cause epigenetic gene silencing. Perhaps pairing between certain homoeologous chromosome regions at pre-meiotic stages would be perceived as un-paired because the strength of normal homologous pairing of the same region would be reduced to the extent that it is below a threshold regarded as normal meiotic pairing (Fig. 2).

5.2. Intergenomic interaction may stimulate formation of expression-altering DNA and/or RNA structures

The foregoing suggests ways that DNA methylation alteration and epigenetic silencing could be induced as a direct consequence of homoeologous chromosome interaction in an allopolyploid. It has been demonstrated that intergenomic interactions in allopolyploid plants also result in rapid structural (genetic) changes (reviewed in Liu and Wendel, 2002; Matzke and Matzke, 1998; Pikaard, 2002; Rieseberg, 2001; Wendel, 2000). One type of change is intergenomic colonization, i.e., spread of DNA from one genome into the other following allopolyploid formation (Hanson et al., 1999; Zhao et al.,...
It can be imagined that some of the displaced DNA segments, when inserted into a non-homologous chromosomal region and particularly in a heterozygous state, may interact with homologous or homoeologous counterparts, and this ecotopic interaction may also be perceived as aberrant- or non-pairing, thereby being targeted for DNA methylation changes and epigenetic gene silencing (Fig. 2). By analogy, in wide hybrids at the diploid level, DNA introgression from related species may have the same consequences (Liu and Wendel, 2000).

Another trigger for de novo DNA methylation change is the presence of direct or inverted repeats. This has been shown in transcriptional silencing of the endogenous *PAI* gene family (encoding enzymes that catalyze the third step in the tryptophan biosynthetic pathway) in certain ecotypes of *Arabidopsis*, where the presence of an inverted repeat within one member of the gene family (*PAI1*) instigates silencing (Luff et al., 1999; Melquist et al., 1999). In these ecotypes, all four members (two copies of *PAI1*, and *PAI2*, *PAI3*) of the gene family are heavily methylated at all cytosine residues and at least one of the members is silenced. In other ecotypes that do not contain the inverted repeat structure in *PAI1*, all three members (*PAI1*, *PAI2*, and *PAI3*) of the *PAI* family are unmethylated (Luff et al., 1999; Melquist et al., 1999). This clearly shows that in natural *Arabidopsis* accessions, the presence of an inverted repeat, as a result of natural genomic structural changes, is essential and sufficient to cause de novo methylation of both itself and sequences homologous to it, leading to transcriptional silencing.

The relevance of the foregoing to allopolyploid evolution emerges from the possibility that allopolyploid formation could generate inverted repeats during the process of genomic accommodation to doubling. For example, intergenomic interactions may result in reciprocal or unidirectional translocations of chromatin from one genome to the other, as discussed above; under certain circumstances, a transposed segment could be placed adjacent and in an inverted fashion to its homoeologue in the other genome, thus generating an inverted repeat structure. That this is a realistic possibility in allopolyploids is evidenced by recent demonstrations that colonization of the alternative genome may be common following genomic merger, as documented in both cotton (Hanson et al., 1999; Zhao et al., 1998) and wheat (Belyayev et al., 2000).

Another possibility for the generation of inverted repeat structure is allopolyploidy-induced mobilization of TEs, as discussed below. TEs usually intrinsically contain inverted or direct terminal repeats, which make them preferential targets for de novo methylation (Bender, 1998). In fact, it was demonstrated recently by genome-wide profiling of DNA methylation that retrotransposons (a major class of TEs) are the preferential target for CMT3 methyltransferase even when the targets exist as single copy sequences (Tompa et al., 2002). Thus, TE transposition may cause DNA methylation changes at or around the new insertion sites, potentially affecting gene expression nearby. Moreover, TE transposition itself could generate novel inverted sequence structures, e.g., the tail-to-tail insertion of two copies of the same element at or near each other, and/or by restructuring flanking host sequences.

An added consequence of the formation of novel genomic structures generated by allopolyploidy is
changes in DNA methylation and gene silencing via RNA–DNA interactions. For example, if the affected regions are transcribed, due to read-through transcription and/or TE-driven transcription, aberrant RNA species such as double-stranded RNAs could be produced. It has been established recently that such aberrant RNAs may be potent triggers of de novo methylation of DNA sequences (regulatory or coding) homologous to the RNAs, thereby silencing the affected genes via transcriptional or post-transcriptional means (Aufsatz et al., 2002; Martienssen and Colot, 2001; Matzke et al., 2001; Paszkowski and Whitham, 2001; Waterhouse et al., 2001; Wolfe and Matzke, 1999).

What proteins are responsible for establishing, maintaining, and modulating changes of cytosine DNA methylation patterns in plants? Initiation and perpetuation of cytosine DNA methylation patterns are known to require the complementary action of diverse de novo and maintenance DNA methyltransferases. In *Arabidopsis*, two types of *maintenance* methyltransferases have been identified, i.e., methyltransferase 1 (MET1, a homolog of the mammalian maintenance methyltransferase Dnmt1) and Chromomethylase 3 (CMT3, a plant-specific methyltransferase containing a chromodomain), that, respectively, maintain CpG and CpNpG methylation patterns (Bartee et al., 2001; Finnegan et al., 1993, 2000; Lindroth et al., 2001). In contrast, de novo methylation of cytosines in plants, including CpG, CpNpG and asymmetric sites, depends on another distinct type of methyltransferase, i.e., the Domains Re-arranged Methylyases (DRM) (Cao and Jacobsen, 2002; Cao et al., 2000). Demethylation has long been believed to be a passive process, resulting from unavailability or dysfunction of maintenance methyltransferases during cell division on semi-methylated DNAs (Jones and Takai, 2001). The recent finding that demethylation occurs in cold-stressed maize roots where cell division is absent, however, implicates the existence of active demethylases in plants (Steward et al., 2002). Thus, changes in cytosine methylation patterns in allopolyploid plants may be a consequence of activation (higher concentration, increased activity and/or accessibility) or dysfunction (inactivation, titration, and/or inaccessibility) of one or more of these enzymes.

5.3. Could allopolyploidy induce changes in epigenetic codes other than DNA methylation?

Although no study has yet addressed this question, given the mechanistic interdependence of various epigenetic marks, it is likely that changes in DNA methylation and the associated gene silencing are not the only epigenetic consequences of allopolyploidy. Recent studies indicate that proper functioning of DNA methyltransferases requires proper chromatin remodeling and histone modifications (Burgers et al., 2002). This has become particularly evident with the discovery in *Arabidopsis of DDM1*, a member of the yeast SWI2/SNF2-like ATPase family that controls chromatin remodeling (Jeddeloh et al., 1999; Singer et al., 2001). It was found that mutations in DDM1 (ddm1 plants) cause a reduction in genomic DNA methylation of 70% (Jeddeloh et al., 1999). This suggested that the function of DDM1 is to enable access of methyltransferase to particular heterochromatin substrates (Martienssen and Colot, 2001). Identification of a plant-specific methyltransferase that has a chromodomain, CMT3, mentioned above, reinforces the close link between cytosine methylation and chromatin remodeling (Bartee et al., 2001; Lindroth et al., 2001). On the other hand, as discussed earlier, all DNA methyltransferases in *Neurospora* and at least some (those directed to CpNpG and asymmetric sites) in plants act downstream of histone methylation, thus indicating dependence of cytosine methylation on specific histone methylation (Jackson et al., 2002; Tamaru and Selker, 2001). Also, it is known in *Neurospora* that DNA methylation, at least at certain loci, depends on histone hypoacetylation (Selker, 1998). It is clear that DNA methylation is dependent on both chromatin remodeling and histone modifications. But which of the later two epigenetic codes comes first? An elegant recent study shows that histone methylation patterns depend on the function of DDM1 (Gendrel et al., 2002).

Taken together, the foregoing discussion raises the possibility that allopolyploidy induces global chromatin remodeling, probably by affecting the normal functioning of enzymes like DDM1, which may in turn cause changes in histone modification and downstream DNA methylation patterns. The observation in nascent allopolyploids of both *Arabidopsis* and wheat that changes in DNA methylation patterns are often non-random (Liu et al., 1998a; Madlung et al., 2002; Shaked et al., 2001) seems to further corroborate this suggestion. This is because DNA methyltransferases exhibit no sequence specificity, and thus their action at specific regions must be directed by other cues, such as changes in histone methylation patterns (reviewed in Richards and Elgin, 2002). Thus it seems probable that, apart from cytosine methylation, other epigenetic mechanisms such as chromatin remodeling and histone modifications may also have been induced to change during or following allopolyploid formation.

5.4. Hypomethylation, chromatin remodeling, and TE activity

Transposable elements encompass two main categories, DNA transposons and retrotransposons, that often exist in highly methylated heterochromatic regions (Bender, 1998; Okamoto and Hirochika, 2001; Tompa et al., 2002; Wessler, 1996). Recent characterizations of
DNA hypomethylation mutants (due to mutations in DDM1 or CMT1), indicate that various types of transposons and/or retrotransposons can be transcriptionally and even transpositionally activated, thus reinforcing the close relationship between cytosine methylation and TE repressive control (Hirochika et al., 2000; Lindroth et al., 2001; Miura et al., 2001; Okamoto and Hirochika, 2001). As discussed above, the mutants ddm1 and cmt3 raise the possibility that chromatin remodeling rather than DNA demethylation per se is the primary cause for at least some TE activations. Indeed, Arabidopsis plants mutated in another SWI2/SNF2-like chromatin remodeling factor (like ddm1), MOR-PHEUS’ MOLECULE 1 (MOM 1), also lead to transcriptional activation of retrotransposons that showed no alteration in cytosine methylation status (Amedeo et al., 2000). Thus, transcriptional activation of TEs under allopolyploid conditions, like that observed in synthetic allotetraploid wheat (Kashkush et al., 2002) and some of the latest published work (Madlung et al., 2002). Demethylation is probably the underlying cause for the observed activation of gene expression in wheat (Kashkush et al., 2002), and some of the novel phenotypes in Arabidopsis (Madlung et al., 2002).

6. Implications of epigenetic changes for allopolyploid genome evolution

6.1. Epigenetic change may promote gene diversification

One of the classic explanations for the evolutionary success of allopolyploidy in plants is that allopolyploid species possess a higher capacity to adapt to changing environments than their diploid progenitors (Grant, 1981; Lewis, 1980; Stebbins, 1950). This notion suggests that one important attribute of allopolyploidy may be the capacity to generate de novo genetic/expression variations that could be translated into novel phenotypes. We suggest that the rapid and widespread occurrence of heritable epigenetic changes represents a fundamental addition to the long-standing explanatory concept of diversification or complementation of duplicated genes.

In diploid species, it is well known that genetic variation often is genetically buffered and phenotypically invisible to selection (Rutherford, 2000). Polyploidy might be expected to further reinforce this genetic buffering due to genome-wide redundancy. Yet investigations from various allopolyploid plants have shown that allopolyploid formation may be associated with the generation of a great deal of morphological variation, at least some of which has epigenetic underpinnings. As discussed in previous sections, in newly synthesized Arabidopsis allopolyploids, epigenetically based phenotypic mutants affecting a large number of traits have been recovered (Comai et al., 2000). Similarly, novel variation in flowering time has been reported in progeny of newly synthesized Brassica allopolyploids (Schranz and Osborn, 2000). Although the molecular basis for these variations are yet to be elucidated, it may be more than coincidental that an earlier study also detected DNA methylation pattern changes in coding sequences in these plants (Song et al., 1995).

Epigenetically caused flowering time variation is a particularly noteworthy example of the creative potential of allopolyploidy, because it has direct bearing on a reproductive trait that clearly could affect fitness and speciation. Perhaps the most renowned example of an epigenetic modification that is evolutionarily consequential in plants is the natural mutant in flower symmetry (from wild-type bilateral to radial) in Linaria vulgaris, originally described by Linnaeus more than 250 years ago. It was demonstrated that the molecular basis for this dramatic transformation in flower morphology is hypermethylation and silencing of Lcyc, a gene controlling flower-form in the mutant; the methylated state is heritable and correlated with the mutant phenotype (Cubas et al., 1999). When one extrapolates evolutionarily significant examples of epigenetic regulation of single genes to the entire genome, it becomes clear that allopolyploid lineages may harbor a nearly infinite and latent reservoir of epigenetic/genetic combinations for later release and evaluation by natural selection, perhaps after millions of years (Adams et al., 2003).

Osborn and colleagues recently proposed three epigenetic mechanisms whereby variation in gene expression and novel phenotypes could be generated in allopolyploid plants: increased variation in dosage-regulated gene expression; altered regulatory interactions; and rapid epigenetic changes (Osborn et al., 2003). To these phenomena we would add epigenetic-controlled TE activity, as elegantly exemplified by the recent work of Kashkush et al. (2003) described above. TEs are a significant source of genetic diversity (Kidwell and Lisch, 2000, 2001; Wessler, 2001), perhaps especially in polyploids where de-repression of quiescent elements may be coupled with the elevated tolerance to transposition (Matzke and Matzke, 1998; Wendel, 2000).
The finding that TEs are associated with a diverse array of wild-type plant genes (particularly in regulatory regions) further suggests important roles that TEs may have played in gene and genome evolution (Wessler, 1997, 1998, 2001). For example, many disease-resistance genes contain TEs and their derivatives, implicating a possible role played by TEs in the rapid diversification of R genes (Song et al., 1997). Apart from direct insertions, TE-mediated ectopic recombination may produce swapping of promoter regions and thereby generate novel gene expression patterns. Thus, we are probably just beginning to appreciate the numerous avenues by which TEs might generate genetic novelties, particularly under allopolyploid conditions where insertionial activity may be tolerated to a higher degree than in diploid antecedents.

An added consequence of epigenetic change is that of permanent genetic change. It has long been recognized that methylated cytosines spontaneously deaminate to thymines at a high frequency (Bird, 1980; Gonzalgo and Jones, 1997), and thus methylation changes accompanying polyploidization create numerous opportunities for novel genetic mutations via this avenue. Also, insertional mutagenesis by mobilized TEs create permanent genetic change, as might increased recombination resulting from compromised epigenetic controls. Thus, allopolyploidy is likely to be associated with increased genetic mutation rates and structural changes.

6.2. Epigenetic change may contribute to genetic and cytological diploidization

Genetic diploidization refers to the phenomenon whereby the expression level of genes in a polyploid is often reduced, by gene silencing and/or dosage compensation, to a level comparable to its diploid progenitors (Soltis and Soltis, 1993). It has been assumed that genetic diploidization is an evolutionary process accomplished by slow mutational decay. As discussed above, epigenetic gene silencing can occur virtually instantaneously following allopolyploidy. It is therefore possible that genetic diploidization, at least for some genes, could occur rapidly by epigenetic means. A clear demonstration of this phenomenon is the epigenetic, reciprocal silencing and organ-specific partitioning of duplicate gene expression following allopolyploid formation in cotton (Adams et al., 2003). It is even possible that rapid genetic diploidization for certain genes may be essential to initial stabilization of nascent allopolyploids.

A pivotal requirement for the success of many allopolyploids is diploid-like meiotic behavior, essential to disomic inheritance and full fertility (Feldman, 1993; Moore, 2002). Although it is plausible that genetic systems conferring this trait, such as the Phl gene in polyploid wheat, may emerge in the course of evolution, alternative mechanisms must be responsible for enforcing exclusive bivalent formation during the nascent stages of allopolyploid evolution. It was proposed by Feldman and colleagues that the rapid and widespread elimination of specific sequences at the F1 hybrid stage and first generations after allopolyploid formation could generate immediate and non-random divergence between homoeologous chromosomes, and thus provide a physical basis for homologous chromosome recognition (Feldman et al., 1997; Liu et al., 1998a,b; Ozkan et al., 2001; Shaked et al., 2001). Little is known about mechanisms whereby targeted sequences are recognized, nor how they become eliminated. A recent study demonstrated that programmed sequence elimination in Tetrahymena is directed by an epigenetic mark, i.e., methylation of histone H3 at Lysine 9 (Taverna et al., 2002). It is thus possible that an epigenetic mechanism is ultimately responsible for rapid and non-random sequence elimination observed in new wheat allopolyploids, and hence contributes to their rapid cytological diploidization. Moreover, as also was demonstrated in wheat, DNA methylation pattern changes accompanying polyploid formation are differential between the constituent genomes (Shaked et al., 2001). Therefore, it is reasonable to deduce that this epigenetic differentiation between genomes might directly or indirectly contribute to homologous chromosome recognition and cytological diploidization.

6.3. Epigenetic change may facilitate intergenomic coordination

Another potential but significant role that rapid epigenetic change may have played is the quenching of incompatible genetic interactions. The sudden union of divergent genomes of distinct parental species will disrupt intrinsic regulatory and developmental harmonies, possibly causing myriad incompatibilities at many levels (Rieseberg, 2001). Although chromosomal incompatibility may be overcome by genome doubling, unfavorable genetic interactions cannot be purged through Mendelian segregation under allopolyploid conditions (Rieseberg, 2001). Epigenetic modifications may provide a means to overcome deleterious genetic interactions in nascent allopolyploids. Negative allelic interactions may be rapidly ameliorated by hypermethylation and silencing or by epigenetic modification or remodeling of the regional chromatin states. Given the potential for innumerable epigenetic/genetic recombinations during allopolyploid speciation and the strong selection that might be associated with such epigenetic modification, one can readily envision the rapidity with which novel epigenetic states might become evolutionarily stabilized.
7. Conclusions and perspectives

Epigenetic phenomena pervade all aspects of plant development and probably are more important in polyploid plants than presently recognized. The often non-random changes observed in natural allopolyploid species and their synthetic counterparts, together with the similarity of observed phenomena among diverse allopolyploid plant systems, suggest that allopolyploidy-associated epigenetic phenomena bear direct relevance to genome evolution and adaptation. Although epigenetic mechanisms likely have played essential roles in both the initial stabilization and long-term evolutionary success of allopolyploidy, relatively little is understood about mechanisms and controls, and even less about how the spectrum of epigenetic changes translates into phenotypic variation for modulation by natural selection. Thus, at present, we can only speculate about the full significance of epigenetics to evolution.

Among the various and possible epigenetic modifications associated with allopolyploidy, at present we know most about changes in cytosine methylation and gene silencing. It is not known whether other epigenetic codes, such as histone modification, have undergone heritable changes following allopolyploidy. Given the interrelatedness and interdependence of the various epigenetic codes, this seems likely. With the availability of the various mutants in Arabidopsis that are defective for each of the epigenetic codes, and advanced techniques like chromatin immunoprecipitation, it is a realistic goal to empirically evaluate these possibilities in the near future. In addition, of the methylation changes described in allopolyploids, to date only alterations at symmetric sites have been studied. Because of the potential defensive role of non-symmetric cytosine methylation (which is a characteristic of RNA-directed DNA methylation) in plant genomes (Martienssen and Colot, 2000), it will be interesting to survey whether this type of cytosine methylation is also subject to allopolyploid-induced modification. Further elucidation of the mechanisms and prevalence of epigenetic phenomena associated with allopolyploidy will undoubtedly further our understanding of the evolutionary process at levels ranging from the gene to the environment.

Acknowledgments

We thank J. Chris Pires for comments on the manuscript. Research in the authors’ laboratories has been supported by the United States National Science Foundation; the United States-Israel Binational Science Foundation; the United States Department of Agriculture; the Iowa State University Plant Sciences Institute; and the China National Science Fund for Distinguished Young Scholars.

References


