Biological Journal of the Linnean Society, 2004, **82**, 665–674. With 4 figures

Biological relevance of polyploidy: ecology to genomics
Edited by A. R. Leitch, D. E. Soltis, P. S. Soltis, I. J. Leitch and J. C. Pires

Genome redundancy and plasticity within ancient and recent *Brassica* crop species

LEWIS N. LUKENS[1]*, PABLO A. QUIJADA[2], JOSHUA UDALL[2†], J. CHRIS PIRES[2], M. ERIC SCHRANZ[2‡] and THOMAS C. OSBORN[2]

[1]*Department of Plant Agriculture, University of Guelph, Guelph, Ontario, N1G 2W1, Canada*
[2]*Department of Agronomy, 1575 Linden Drive, University of Wisconsin, Madison, WI 53706–1597, USA*

Received 14 July 2003; accepted for publication 5 January 2004

The crop species within the genus *Brassica* have highly replicated genomes. Three base 'diploid' species, *Brassica oleracea*, *B. nigra* and *B. rapa*, are likely ancient polyploids, and three derived allopolyploid species, *B. carinata*, *B. juncea* and *B. napus*, are created from the interspecific hybridization of these base genomes. The base *Brassica* genome is thought to have hexaploid ancestry, and both recent and ancient polyploidization events have been proposed to generate a large number of genome rearrangements and novel genetic variation for important traits. Here, we revisit and refine these hypotheses. We have examined the *B. oleracea* linkage map using the *Arabidopsis thaliana* genome sequence as a template and suggest that there is strong evidence for genome replication and rearrangement within the base Brassicas, but less evidence for genome triplication. We show that novel phenotypic variation within the base Brassicas can be achieved by replication of a single gene, *BrFLC*, that acts additively to influence flowering time. Within the derived allopolyploids, intergenomic heterozygosity is associated with higher seed yields. Some studies have reported that *de novo* genomic variation occurs within derived polyploid genomes, whereas other studies have not detected these changes. We discuss reasons for these different findings. Large translocations and tetrasomic inheritance can explain some but not all genomic changes within the polyploids. Transpositions and other small-scale sequence changes probably also have contributed to genomic novelty. Our results have shown that the *Brassica* genomes are remarkably plastic, and that polyploidy generates novel genetic variation through gene duplication, intergenomic heterozygosity and perhaps epigenetic change.

ADDITIONAL KEYWORDS: *FLOWERING-LOCUS C* – genome evolution – polyploidy.

INTRODUCTION

Polyploidization results in a myriad of phenomena and processes that are both short term and long term in nature. In the short term, two different genomes must rapidly adapt to a common nucleus, both by regulating gene expression for proper development and by regulating chromosome pairing and transmission for cell division (see also Chen *et al*., 2004; Kovarik *et al*., 2004; Pires *et al*., 2004 – all this issue, for genetic changes associated with polyploids). In the longer term, within a polyploid population, redundant genes may retain function, lose function or gain novel functions. Both rapid and long-term polyploid genomic change can cause substantial phenotypic diversity. Polyploids may differ from their parental, diploid species in morphology, physiology, resistance to abiotic

*Corresponding author. E-mail: llukens@uoguelph.ca
Present addresses: †Department of Ecology, Evolution, and Organismal Biology, Bessey, Hall, Iowa State University, Ames, IA, USA; ‡Department of Genetics and Evolution, Max Planck Institute of Chemical Ecology, Jena, Germany.

and biotic stresses as well as other characteristics (Levin, 1983). The regulation and structure of polyploid genomes may also differ in numerous ways from the genomes of their diploid progenitors (Wendel, 2000). The *Brassica* crop species have long been a model system to study the molecular and phenotypic changes associated with both recent and ancient polyploidization events.

The genomes of the base diploid species *B. rapa* ($n = 10$), *B. nigra* ($n = 8$) and *B. oleracea* ($n = 9$) are very similar to each other, and all have a high level of genomic redundancy. Researchers have proposed that different mechanisms generated this redundancy. A recent explanation is that the base *Brassica* genome is comprised of three ancestral genomes that are *Arabidopsis thaliana*-like in structure (Lagercrantz, 1998). The phenotypic effects of the base *Brassica* genomes' duplicated regions have also been widely studied. These regions of shared ancestry have provided the opportunity to examine the fate of duplicate genes that control *Brassica* development. Duplicate genes are expected to be lost at a high rate unless they functionally diverge (e.g. Force *et al*., 1999).

In contrast to the diploid Brassicas, which have an ancient origin, other crop species within the genus [such as *B. napus* ($n = 19$), *B. juncea* ($n = 18$) and *B. carinata* ($n = 17$)] are allotetraploids whose genomes were derived from the recent fusion of two base diploid genomes. U (1935) first proposed this relationship between the polyploid Brassicas and their diploid progenitors. *B. napus* ($n = 19$) can be resynthesized by crossing *B. rapa* ($n = 10$) and *B. oleracea* ($n = 9$), rescuing the embryo and colchicine doubling the resultant amphihaploid (see Pires *et al*., 2004). *Brassica juncea* and *B. carinata* can be resynthesized in a similar way by crossing *B. rapa* and *B. nigra* and *B. oleracea* and *B. nigra*, respectively. Although phylogenies based on RFLP markers from the organelle genomes suggest that the maternal ancestor of *B. napus* is more closely related to *B. montana* than to *B. oleracea* or *B. rapa* (Song & Osborn, 1992), several molecular mapping analyses have confirmed the similarity between the base diploid progenitor genomes and the polyploid genomes. For example, linkage groups N1 to N10 of *B. napus* correspond to the A (*B. rapa*) genome, and N11 to N19 correspond to the C (*B. oleracea*) genome (Parkin *et al*., 1995).

Recent polyploidization events within the Brassicas can cause genomic changes within the derived polyploid genomes, and these genomic changes can have phenotypic consequences. Song *et al*. (1995) hybridized low-copy DNA fragments to cleaved chloroplast and nuclear DNA extracted from resynthesized *B. napus* plants and their progeny. Fragments present in S_1 plants were lost within the S_4, and novel restriction fragments appeared between the S_1 and S_4 generations. Although not all *Brassica* allopolyploids are characterized by genomic change (Axelsson *et al*., 2000), when genomic changes do occur, they can generate phenotypic diversity (e.g. Pires *et al*., 2004). These phenotypic effects may have a profound impact on the ability of new polyploids to become established or selected for use in agriculture.

We have sought to understand the replicated genomic structures of both the recent and the ancient *Brassica* species and the biological mechanisms that produced those structures. In addition, we have sought to understand the molecular basis of polyploid phenotypic variation within the *Brassica* species. Within the base Brassicas, we have found that although several regions of the *Brassica* genomes are triplicated, there is no strong evidence that the diploid Brassicas are ancient hexaploids with ancestral genomes that are similar to that of *A. thaliana*. In addition, a large number of small, perhaps gene-sized rearrangements distinguish *A. thaliana* and the base species *B. oleracea*. Analysis of the replicated transcription factor *FLC* (*FLOWERING-LOCUS C*) in *B. rapa* suggests that within the ancient polyploids, genes that retain their ancestral function and act additively provide novel genetic variation upon which selection can act. Among the recent polyploids, we suggest mechanisms that can account for the genome stability differences among polyploid populations, and mechanisms that can account for the observed genomic changes within recent polyploids. We finally describe analyses indicating that homozygous polyploids generate offspring that are phenotypically variable and that intergenomic heterozygosity contributes to polyploid fitness.

THE SIGNATURE AND PHENOTYPIC EFFECTS OF ANCIENT POLYPLOIDY WITHIN THE DIPLOID *BRASSICA* GENOMES

THE STRUCTURE AND ANCESTRY OF THE BASE *BRASSICA* GENOMES

For several years, researchers have known that different regions of the diploid *Brassica* genomes are highly similar in content. However, interpretations regarding the extent of this genomic replication and the events that caused it vary. Röbbelen (1960) identified six common *Brassica* chromosome types based on characteristics such as the length and the symmetry of chromosome arms. He found that each base species has all six chromosome types, but certain chromosome types are duplicated or triplicated in each species' genome, resulting in different chromosome numbers. By contrast, Truco *et al*. (1996) examined a common set of segregating, low-copy DNA fragments, within mapping populations from the three base *Brassica*

crop species. By interpreting the fragment linkage patterns among the different genomes, they suggested that the base genomes had six ancestral chromosomes but hypothesized that these chromosomes underwent several duplications and rearrangements. As with Truco *et al*. (1996), Lagercrantz & Lydiate (1996) examined the base genomes' linkage groups with shared markers, but they came to a different conclusion. Because linked markers within one species' genome often corresponded to several groups of linked markers from other species' genomes, Lagercrantz & Lydiate (1996) reported that most of the genome of *B. nigra* and the other base Brassicas are in triplicate. In a subsequent paper, Lagercrantz (1998) suggested that the *A. thaliana* genome is similar in complexity to each of the three units of the *B. nigra* genome, but approximately 90 genomic rearrangements have taken place since the divergence of the two species. The idea that the base *Brassica* species are derived from a hexaploid ancestor similar to *A. thaliana* that has undergone several rearrangements has provided a useful null hypothesis for subsequent experiments to compare genome structures.

Some comparisons of chromosomal segments between *A. thaliana* and *Brassica* species find that the same region of the *A. thaliana* genome is present in a multiple of three within the *Brassica* genomes, consistent with the triplication hypothesis (Cavell *et al*., 1998; O'Neill & Bancroft, 2000; Parkin, Lydiate & Trick, 2002). However, whole genome analyses subsequent to the study reported by Lagercrantz (1998) suggest that several regions of the *Brassica* genome deviate from this expectation. Although some regions within the *A. thaliana* genome are similar to three *B. oleracea* fragments, Lukens *et al*. (2003) found that other regions of the *A. thaliana* genome were similar to between zero and seven regions within *B. oleracea.* Lan *et al*. (2000), expanding on the analysis of Kowalski *et al*. (1994), compared the segregation pattern of fragments homologous to *A. thaliana* expressed sequence tags (ESTs) within a single *A. thaliana* and several *B. oleracea* mapping populations. Duplication in the *B. oleracea* genome was strongly suggested because parallel arrangements of duplicated loci between *Brassica* and *Arabidopsis* accounted for 41% of the loci mapped in *Brassica*. By contrast, Lan *et al*. (2000) found that a triplication model explained only 18% of the data, as compared with a random expectation of 14%. Thus, although the diploid *Brassica* species almost certainly evolved from a polyploid ancestor, the current evidence that the base *Brassica* genomes are derived from a hexaploid is not convincing. Nonetheless, whole genome analyses have also been unable to explain confidently the genome structure of *Brassica* relative to *A. thaliana* by other scenarios such as the formation of a tetraploid and subsequent segmental, chromosomal duplications.

There are several reasons why it is difficult to understand the whole genomic relationships between *Brassica* and *Arabidopsis* with confidence, but primarily this is because comparative map data can be misleading and are invariably incomplete. For example, to map the position of three homoeologous alleles within a single mapping population, both parents must differ at all three loci. Thus, to detect linkage among four loci that are present on three homoeologous chromosomes would be very rare. In fact, Lagercrantz (1998) cited the frequent detection of two groups of shared markers as support for the actual presence of three groups of shared markers. Another problem in comparative mapping is the identification of orthologous loci. Lukens *et al*. (2003) pointed out that in comparing genomes one must identify genomic regions with shared most recent common ancestry (MRCA). For example, if one incorrectly infers that a paralogous locus of *Arabidopsis* (A2) is most closely related to the *Brassica* locus (B1), then one incorrectly concludes that a genome change has occurred since the divergence of the two species. In addition, because small sequences may 'shuffle' after the divergence of two genomes (i.e. Fig. 1), one may identify two orthologous sequences, but these sequences may not lie in orthologous regions. Comparative mapping analyses often assume that if a homologous sequence is orthologous between two genomes, then the regions surrounding that sequence are also orthologous. Because single gene deletions and rearrangements are not uncommon, this assumption could lead to over-estimating genome dissimilarity. Finally, individual investigators may interpret data quite differently. When possible, both the criteria for identifying orthologous sequences and the criteria for identifying and evaluating collinear regions should be explicit (Lukens *et al*., 2003).

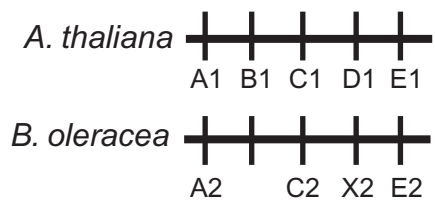

Figure 1. Schematic representation of the relationship between the *B. oleracea* and *A. thaliana* genomes. The horizontal lines represent orthologous, gene-coding regions on two chromosomes, for example *A. thaliana* chromosome 5 and *B. oleracea* linkage group 9 (Lukens *et al*., 2003). The same letters on both chromosomes represent orthologous genes shared between them. Regions with different genes represent areas of local insertions/deletions or translocations.

A large number of rearrangements distinguish the *Brassica* and *Arabidopsis* genomes

Although there are large regions of homology between the compact, sequenced *A. thaliana* genome and the *B. oleracea* genome, recent studies suggest that a very large number of both large and small chromosomal rearrangements distinguish the two genomes. Lukens *et al*. (2003) found evidence for large regions of homology between the *A. thaliana* and *B. oleracea* genomes, but perhaps more remarkably identified a large number of both chromosomal rearrangements and intrachromosomal duplications in *B. oleracea* relative to *A. thaliana*. Lukens *et al*. compared a *B. oleracea* genetic map with the *A. thaliana* genome by sequencing RFLP marker probes linked within *B. oleracea* and by identifying similar sequences within the *A. thaliana* genome sequence. They found that a single *Brassica* DNA sequence is often similar to one region within the *A. thaliana* genome, but other *Brassica* DNA sequences linked to it are similar to different regions in *A. thaliana*. Although this pattern of similarity could be due to the detection of ancient duplications while constructing the *Brassica* genetic map or to incorrect sequence alignment between a *Brassica* sequence and a paralogous sequence in *A. thaliana*, the high frequency of these putative small genomic changes suggests that the genomes differ substantially at many loci. Other recent analyses are consistent with this hypothesis. Ryder *et al*. (2001) identified fragments that behaved in a similar way in comparisons between the positions of mapped markers within the *B. oleracea* genome and their position in the *A. thaliana* genome. Likewise, nucleotide sequence and physical mapping comparisons between *Brassica* and *Arabidopsis* have shown that although the order of genes within a homologous region can be similar, a large number of gene insertions or deletions distinguish the regions. For example, three genes (*ABI1*, *RPS2* and *Ck1*) are collinear within *B. oleracea* and *A. thaliana*. However, an additional gene, *N-myr*, lies between *RPS2* and *Ck1* in *B. oleracea*, but it is absent in *A. thaliana* (Quiros *et al*., 2001). Similarly, the position of genes within a sequenced 222-kb region of *A. thaliana* was compared with the position of genes within *B. oleracea* by constructing a BAC library of the *B. oleracea* genome (O'Neill & Bancroft, 2000). Although all *A. thaliana* genes within this region are present within the *Brassica* genome, there are numerous genes present in *A. thaliana* that are absent from the homologous position in *B. oleracea*, and one gene within *Brassica* was identified in a non-collinear position. We suggest that most regions within the base *Brassica* genome have a structure of gross similarity to one region of *A. thaliana*, but upon a detailed inspection numerous small insertions/deletions and rearrangements distinguish the two genomes.

Effect of polyploidy on ancient *Brassica* genome structure

As described above, the base diploid *Brassica* genomes are very likely derived from ancient polyploids. Several researchers have suggested that polyploid genomes undergo numerous chromosomal changes, such as insertions/deletions and translocations (Otto & Whitton, 2000; Ramsey & Schemske, 2002). Genome rearrangement is also correlated with genome duplication in the base Brassicas. As mentioned, Lukens *et al*. (2003) found a large amount of genome repatterning between *A. thaliana* and *B. oleracea*. However, only a single *A. thaliana* region typically corresponds to a single region in *B. oleracea*, and almost all markers mapped on the *B. oleracea* genetic map have homologous sequences in *A. thaliana*. These results suggest that the *A. thaliana* genome has not undergone genome duplication or sequence loss since its divergence from *B. oleracea*. In contrast to the *A. thaliana* and *B. oleracea* comparison, both mapping and sequence comparisons between *A. thaliana* and its relative *Capsella rubella*, a species that does not have a replicated genome, show very little genome change (Rossberg *et al*., 2001).

Phenotypic effects of gene redundancy within the base *Brassica* genomes

Over time, duplication of whole genomes or genome sequences is expected to be accompanied by widespread gene deletion because duplicate genes may have identical functions (Simillion *et al*., 2002; Bowers *et al*., 2003). However, several similar regions of the base *Brassica* genomes influence flowering time. Quantitative trait loci (QTL) mapping to regions of the *Brassica* genomes that are homologous with the top of *A. thaliana* chromosome 5 can explain much flowering time variation both within annual and between annual and biennial populations of *B. oleracea*, *B. nigra* and *B. rapa* (Lagercrantz *et al.*, 1996; Osborn *et al*., 1997; Bohuon *et al*., 1998; Axelsson, Shavorskaya & Lagercrantz, 2001; Parkin *et al*., 2002; Schranz *et al*., 2002; reviewed in Osborn & Lukens, 2003). These studies have suggested that replicated regions of the *Brassica* genome have retained ancestral functions. Recent analyses suggest that the effects of these 'duplicated' QTL are probably due to single genes that have retained their function and therefore affect flowering time in an additive manner. Two alleles of the *B. rapa* transcription factor *FLC* derived from a biennial parent co-segregate with loci that delay flowering time (Schranz *et al*., 2002). Similarly,

functionally redundant copies of the daylength responsive gene *CONSTANS* (*CO*) have been hypothesized to explain variation among annual *B. rapa* (Axelsson *et al*., 2001). The presence of multiple genes that act in an additive fashion within the base *Brassica* genome suggests that these Brassicas could flourish in a wider range of environments than their diploid progenitors. It would be interesting to see which homoeologous alleles are retained in natural populations that flower at different times and which combinations of alleles have been most widely used to adapt cultivars to various crop production zones.

GENOMIC AND PHENOTYPIC EFFECTS OF POLYPLOIDY IN *BRASSICA* AMPHIDIPLOID CROP SPECIES

Putative causes for chromosomal rearrangements among recent *Brassica* polyploids

During the early stages of allopolyploidization, two different genomes must be adapted within a common nucleus, a process that has been associated with changes in nuclear genomic DNA (e.g. Song *et al*., 1995; Feldman *et al*., 1997). Parkin *et al*. (1995) suggested that both homoeologous recombination and tetrasomic inheritance could explain the loss of parental alleles among the progeny of a *B. napus* mapping population. Parental alleles within a doubled haploid population of plants derived from F_1 microspores are expected to segregate in a manner consistent with disomic inheritance (Fig. 2A). For several linked loci, however, each plant of the doubled haploid (DH) population had alleles from only one parent. Parkin *et al*. (1995) suggested that a non-reciprocal translocation that occurred during meiosis within a parental plant explains this pattern (Fig. 2B). For other linked loci (on N2/N12), some DH plants had any pair of parental alleles. This pattern of segregation is consistent with tetrasomic inheritance (Fig. 2C) or could arise from non-reciprocal translocation. Of 50 DH lines derived from F_1 microspores, six lines had duplicate copies of linked loci but did not have any copies of homoeologous loci, indicating that non-reciprocal translocations occurred during male gametogenesis within the F_1 plant. An additional line had four copies of N18, indicating that chromosomes failed to disjoin during meiosis. Three lines had duplicated chromosome fragments, and one line appeared to have a deletion. Recently, Udall (2003) and Quijada (2003) reported similar results for linked markers in other DH populations generated from crosses between resynthesized and natural *B. napus* and between natural *B. napus* lines. Again, lines within at least one population were missing both alleles for a given locus but had two doses of the homoeologous locus. Like Parkin *et al*. (1995), Udall (2003) and Quijada (2003) interpreted this segregation pattern as indicative of a non-reciprocal translocation that occurred either within the resynthesized parent or within one of the natural parents.

Homoeologous chromosome exchange and tetrasomic inheritance can also explain the loss of parental fragments after generations of *B. napus* self-fertilization, as described by Song *et al*. (1995). Song *et al*. (1995) resynthesized doubled haploid *B. juncea* (AB) and *B. napus* (AC) plants (S_1 here, but referred to as F2 in the original publication) and selfed the plants for four generations (S_4, referred to as F5). Because the genome of doubled haploid plants is derived from a single haploid genome, these plants would be expected to breed true and be homozygous at all loci in subsequent generations. Within this population, although parental alleles present within an S_2 individual were sometimes absent from an S_4 individual, at least one parental allele was usually present in each individual S_4 plant, as would be predicted if one allele was 'converted' by recombination.

The frequency of DNA fragment changes in Song *et al*.'s study was very high. Although all fragments from the plastid genome were the same between the S_1 and the S_4 generations, within the nuclear genome, between approximately 5% and 13% of the fragments changed. In general, the frequency of homoeologous recombination within the *Brassica* genomes is higher in re-synthesized genomes than in 'natural' polyploid genomes. Homologous chromosomes in synthetic *B. napus* do not pair as accurately as natural *B. napus* chromosomes (Röbbelen, 1960), and linkage analyses confirm this cytological observation. In four natural *B. napus* by natural *B. napus* crosses, approximately 0.4% of recombination events were not between homologous chromosomes (Sharpe *et al*., 1995; Udall, 2003). By contrast, within synthetic *B. napus* genomes, over 1% of total recombination events were not between homologous chromosomes (Parkin *et al*., 1995; Udall, 2003). Similarly, cytogenetic studies show that in contrast to naturally occurring *B. juncea* chromosomes, synthetic *B. juncea* may have univalents and trivalents at a fairly high frequency within the F_1 (Olsson, 1960).

The frequency of homoeologous recombination in *Brassica* may also be correlated with the presence of specific genetic factors. For example, within *B. juncea*, Song *et al*. (1995) observed numerous examples of non-Mendelian inheritance of parental alleles after selfing lines, whereas in another study of *B. juncea*, Axelsson *et al*. (2000) observed only the disomic inheritance of parental alleles. The latter authors derived a population of 60 individuals from a synthetic (*B. rapa* × *B. nigra*) DH plant and backcrossed to a DH

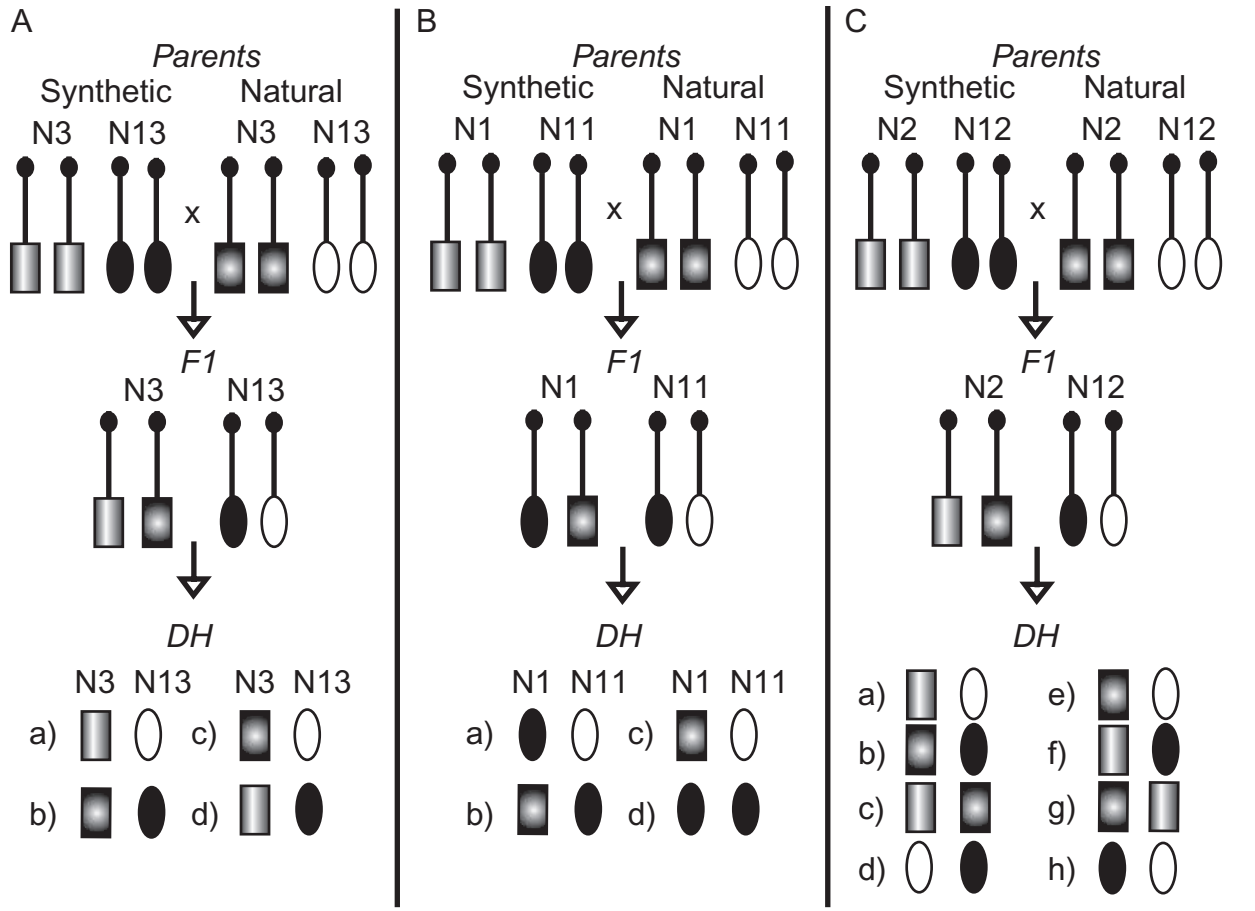

Figure 2. Segregation of homoeologous chromosomes within a synthetic × 'natural' *B. napus* cross followed by doubled haploid production. A, disomic inheritance; homologous and homoeologous chromosomes segregate independently. Segments from all parental chromosomal segments are present. B, non-reciprocal translocations lead to a loss of parental fragments within the progeny. C, tetrasomic inheritance leads to loss of parental segments following meiosis.

B. juncea tester (Fig. 3A). By contrast, Song *et al*. (1995) examined differences among the selfed progeny of a resynthesized polyploid (Fig. 3B). Although homoeologous recombination would be expected to occur more frequently in Song *et al*.'s material because plants were derived from the mixing of two gametes from a synthetic parent, homoeologous recombination alone cannot explain the differences between this material. In Axelsson *et al*.'s population, every polymorphic allele from both parents segregated in a disomic fashion. It is possible that the ability of homoeologous chromosomes to pair faithfully depends on the presence of a locus that may be absent from Song *et al*.'s synthetic parent. Prakash (1974) found that a natural *B. juncea* line homozygous for a translocation also exhibited a high level of multivalent pairing. Thus, it is possible that *B. juncea* lines used in the different studies had different genotypes at a locus such as the *Ph1* locus in wheat (Sears, 1977).

OTHER CAUSES FOR GENOMIC CHANGE WITHIN THE *BRASSICA* SPECIES

Although high rates of homoeologous recombination among resynthesized lines can explain some of the observed instability of the *Brassica* genomes, homoeologous recombination does not account for all changes. Homoeologous recombination cannot explain allelic changes within *Brassica* plants that have not undergone meiosis, nor the presence of novel, non-parental fragments that appear within the progeny of *B. napus* and *B. juncea*. A comparison of the positions of low-copy, largely genic DNA fragments on the *B. napus* genetic map with the position of homologous sequences within the *A. thaliana* genome reveals that low-copy sequences within the *B. napus* genome are more highly dispersed within the *A. thaliana* genome than are low-copy sequences of *B. oleracea* (L. N. Lukens, unpubl.

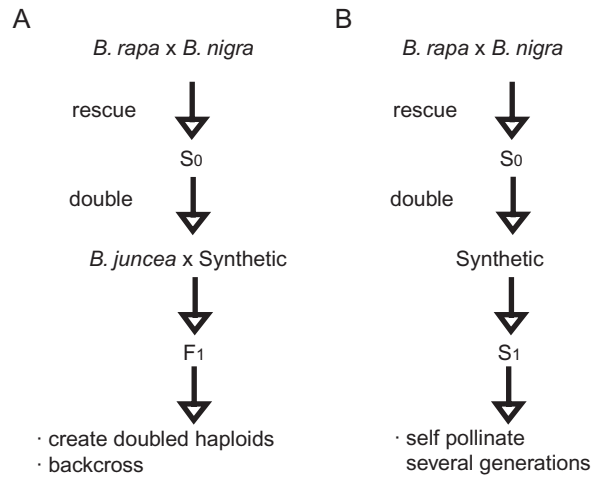

Figure 3. Different schema for determining the inheritance of DNA fragments within resynthesized polyploids. A, base diploid *Brassica* species, in this case *B. rapa* and *B. nigra*, are crossed to create an amphihaploid embryo that is subsequently doubled with colchicine to form an amphidiploid. The doubled haploid (DH) plant is crossed to another *B. juncea* plant. To generate a segregating population, the F_1 is backcrossed or microspores from the F_1 hybrid of the *B. juncea* and synthetic plants are cultured into haploid plantlets whose chromosome complement is subsequently doubled with colchicine. Each plant within the segregating population is screened for fragments at many loci. B, a DH plant is created as above but is selfed for several generations. Each plant within the segregating population is screened for fragments at many loci. Mapping populations of *B. napus* are created in a similar way by rescuing the embryos in a cross between *B. oleracea* and *B. rapa*.

data). In addition, within the progeny of resynthesized Brassicas, Song *et al*. (1995) detected differences in the size of DNA fragments generated by *Hin*dIII. Because *Hin*dIII is insensitive to methylation in the short symmetrical CpG and CpNpG sequences where most plant DNA methylation occurs, these novel fragments are probably the result of DNA insertion or deletion. Finally, we have also observed the appearance of novel, non-parental, low-copy DNA fragments within 50 synthetic *B. napus* S_1 lines that are derived directly from diploid gametes (L. N. Lukens & J. C. Pires, unpubl. data). Thus, *Brassica* is similar to the *Triticum–Aegilops* group. In wheat, a high frequency of genomic changes occurs prior to meiosis within the F_1 hybrid of two diploid progenitors (Ozkan, Levy & Feldman, 2001; Shaked *et al*., 2001), and parental alleles of coding sequences may disappear or be replaced by novel fragments (Liu *et al*., 1998b).

Although the mechanism responsible for the formation of novel alleles remains unclear, we suggest that small-scale genomic rearrangements may play a role. Within polyploid plants, transposons colonize different genomes (e.g. Hanson *et al*., 2000), and enhanced transposable element activity has been shown to accompany polyploidization in several species (Liu & Wendel, 2002; Levy & Feldman, 2004 – this issue). A transposable element insertion into a locus would create a novel allele. In addition, transposable elements may mobilize genomic DNA. In *Drosophila teissieri* and *D. yakuba*, the sequence of a processed *Adh* messenger RNA was integrated into the nuclear genome and captured several upstream exons and introns of an unrelated gene, creating a novel, chimeric gene, *jingwei* (Long & Langley, 1993). Transcription of sequences flanking the human L1 retrotransposons can also lead to integration of human coding DNA into novel regions of the genome (Moran, DeBerardinis & Kazazian, 1999). These findings support the hypothesis that *Brassica* nuclear coding sequences may rearrange, although other mechanisms of intergenomic exchanges (e.g. Sperisen, Ryals & Meins, 1991) cannot be ruled out. The fact that inbreeding of doubled haploid *Brassica* lines appears to cause greater genomic instability than outcrossing these lines is consistent with increased transposon activity. For example, in *A. thaliana*, plants that are homozygous for *ddm1* mutations that affect DNA methylation do not have a noticeable phenotype until after several generations of selfing (Kakutani *et al*., 1996).

Effects on phenotype: flowering time and seed yield

Several authors have remarked on the phenotypic variation that arises following polyploidization within the Brassicas (e.g. Osborn *et al*., 2003). Schranz & Osborn (2000) quantified this variation for flowering time. They self-pollinated synthetic polyploids that flowered either early or late relative to other plants within a population derived from the same DH parent. The progeny of the selfed plants had significant differences in flowering time, showing that the late- and early-flowering characteristics were heritable (see also Pires *et al*., 2004). The selfed progeny of the extreme early and late lines were also selected for several additional generations. Remarkably, the differences in flowering time between the early- and late-flowering lines became greater after each generation, showing that stable, *de novo* genomic changes occurred in the generations following polyploidization. These lines also differed in life history traits other than flowering time (Schranz & Osborn, 2004).

A significant amount of the heritable phenotypic differences within this and other *B. napus* popula-

tions appears to be due to chromosomal rearrangements. Pires *et al*. (2004) identified a non-reciprocal translocation that accounts for 29% of the flowering time variation among an F_2 population generated from a cross between the early and late *B. napus* lines in the sixth selfed generation. Osborn *et al*. (2003) investigated the segregation of loci within populations derived from crosses between diverse *B. napus* genotypes. In several crosses, multiple RFLP probes to homoeologous regions of linkage groups N7 and N16 indicated the presence of reciprocal translocations (Fig. 4). Translocation heterozygotes have little effect on pollen development, self-incompatibility or seed development, and F_1 hybrids between individuals carrying the translocation often produced high seed yields. However, non-parental configurations of the translocation greatly affected seed yield. Plants that had *B. rapa*-derived (N7) alleles on both N7 and N16 yielded far less seed than plants that had *B. rapa*-derived alleles at (N7) and *B. oleracea*-derived alleles on N16 (Fig. 4). Although we attribute both variation in yield and variation in flowering time in part to translocations within *B. napus*, we suggest that small-scale rearrangements and epigenetic changes in chromatin structure probably also influence the polyploid phenotype. These factors could account for both the phenotypic variation that is unexplained by translocations and the continuous, directional phenotypic change observed in Schranz & Osborn (2000).

CONCLUSION

Numerous mechanisms contribute both to genomic change and to phenotypic novelty within both recent and ancient *Brassica* polyploids. Although we have learned much regarding these mechanisms, our understanding is only in its initial stages. Recently derived *Brassica* polyploid plants that are the offspring of selfed, homozygous parents should be identical at all loci but are nonetheless morphologically diverse. We have only initiated our study of this process. We do not yet know whether DNA gain or loss within the recent polyploid genomes is stochastic or directional. For example, specific sequence changes in wheat have been associated with the promotion of disomic inheritance (Feldman *et al*., 1997; Liu, Vega & Feldman, 1998a). We have also not investigated the genetic factors that influence homoeologous pairing within the *Brassica* polyploids. Finally, although there are conserved regions that are shared between closely related genomes, there are also a remarkable number of small genomic rearrangements, and we suggest that these small genomic rearrangements may be associated with allopolyploidy.

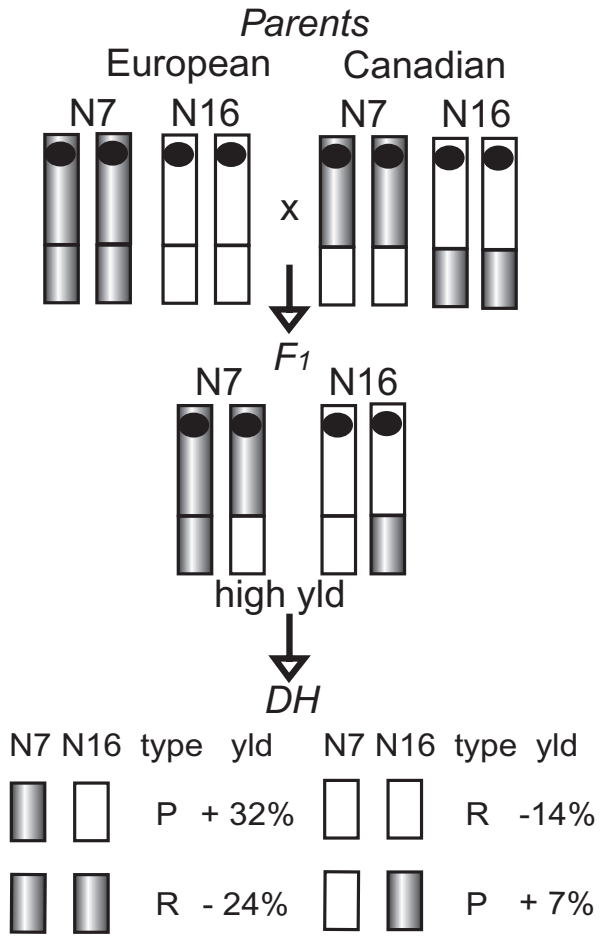

Figure 4. Phenotypic effects of a segregating reciprocal translocation within several *B. napus* lines. Representations of the parental genotypes with and without a translocation are shown at the top of the figure. The depiction of the F_1 genotype assumes disomic inheritance. The squares depicting the doubled haploid represent the chromosome regions that can show translocations. Yields are the percentage differences from the grand mean for the mean yield of all genotypes. R, recombinant genotypes; P, parental genotypes. Recombination within the translocation is not shown.

REFERENCES

Axelsson T, Bowman CM, Sharpe A, Lydiate D, Lagercrantz U. 2000. Amphidiploid *Brassica juncea* contains conserved progenitor genomes. *Genome* **43:** 679–688.

Axelsson T, Shavorskaya O, Lagercrantz U. 2001. Multiple flowering time QTLs within several *Brassica* species could be the result of duplicated copies of one ancestral gene. *Genome* **44:** 856–864.

Bohuon EJ, Keith DJ, Craft JA, Arthur AE, Marshall DF, Lydiate DJ, Kearsey MJ. 1998. The association of flowering time quantitative trait loci with duplicated regions and candidate genes in *Brassica oleracea*. *Genetics* **150:** 393–401.

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

Cavell A, Lydiate D, Parkin I, Dean C, Trick M. 1998. A 30 centimorgan segment of *Arabidopsis thaliana* chromosome 4 has six collinear homologues within the *Brassica napus* genome. *Genome* **41:** 62–69.

Chen ZJ, Wang J, Tian L, Lee H-S, Wang JJ, Chen M, Lee JJ, Josefsson C, Madlung A, Watson B, Lippman Z, Vaughn M, Pires JC, Colot V, Doerge RW, Martienssen RA, Comai L, Osborn TC. 2004. The development of an *Arabidopsis* model system for genome-wide analysis of polyploidy effects. *Biological Journal of the Linnean Society* **82:** 689–700.

Feldman M, Liu B, Segal G, Abbo S, Levy AA, Vega JM. 1997. Rapid elimination of low-copy DNA sequences in polyploid wheat: a possible mechanism for differentiation of homoeologous chromosomes. *Genetics* **147:** 1381–1387.

Force A, Lynch MF, Pickett B, Amores A, Yan Y, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151:** 1531–1545.

Hanson RE, Islam-Faridi N, Crane CF, Zwick MS, Czeschin DG, Wendel JF, McKnight TD, Price HJ, Stelly DM. 2000. Ty1-copia-retrotransposon behavior in a polyploid cotton. *Chromosome Research* **8:** 73–76.

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards E. 1996. Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proceedings of the National Academy of Sciences, USA* **93:** 12406–12411.

Kovarik A, Matyasek R, Lim KY, Skalická K, Koukalová B, Knapp S, Chase M, Leitch AR. 2004. Concerted evolution of 18–5.8–26S rDNA repeats in *Nicotiana* allotetraploids. *Biological Journal of the Linnean Society* **82:** 615–625.

Kowalski S, Lan TH, Feldmann K, Paterson AH. 1994. Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. *Genetics* **138:** 499–510.

Lagercrantz U. 1998. Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics* **150:** 1217–1228.

Lagercrantz U, Lydiate D. 1996. Comparative genome mapping in *Brassica*. *Genetics* **144:** 1903–1910.

Lagercrantz U, Putterill J, Coupland G, Lydiate D. 1996. Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. *Plant Journal* **9:** 13–20.

Lan THI, DelMonte TA, Reischmann KP, Hyman J, Kowalski S, McFerson J, Kresovich S, Paterson AH. 2000. An EST-enriched comparative map of *Brassica oleracea* and *Arabidopsis thaliana*. *Genome Research* **10:** 776–788.

Levin D. 1983. Polyploidy and novelty in flowering plants. *American Naturalist* **122:** 1–25.

Levy AA, Feldman M. 2004. Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization. *Biological Journal of the Linnean Society* **82:** 607–613.

Liu B, Vega JM, Feldman M. 1998a. Rapid genome changes in newly synthesized amphidiploids of *Triticum* and *Aegilops*. II. Changes in low copy coding DNA sequences. *Genome* **41:** 535–542.

Liu B, Vega JM, Segal G, Abbo S, Rodova M, Feldman M. 1998b. Rapid genome changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. I. Changes in low-copy noncoding DNA sequences. *Genome* **41:** 272–277.

Liu B, Wendel J. 2002. Non-Mendelian phenomena in allopolyploid genome evolution. *Current Genomics* **3:** 489–505.

Long M, Langley CH. 1993. Natural selection and the origin of *jingwei*, a chimeric processed functional gene in *Drosophila*. *Science* **260:** 91–94.

Lukens L, Zou F, Lydiate D, Parkin I, Osborn T. 2003. Comparison of a *Brassica oleracea* genetic map with the genome of *Arabidopsis thaliana*. *Genetics* **146:** 359–372.

Moran JV, DeBerardinis RJ, Kazazian HH. 1999. Exon shuffling by L1 retrotransposition. *Science* **283:** 1530–1534.

O'Neill C, Bancroft I. 2000. Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homoeologous to sequenced regions of chromosomes 4 and 5 of *A. thaliana*. *Plant Journal* **23:** 233–243.

Olsson G. 1960. Species crosses within the genus *Brassica*. II. Artificial *Brassica juncea*. Coss. *Hereditas* **46:** 171–222.

Osborn TC, Butrulle DV, Sharpe AG, Pickering KJ, Parkin IAP, Parker JS, Lydiate DJ. 2003. Detection and effects of a homoeologous reciprocal transposition in *Brassica napus*. *Genetics* **165:** 1569–1577.

Osborn TC, Kole C, Parkin IAP, Sharpe AG, Kuiper M, Lydiate DJ, Trick M. 1997. Comparison of flowering time genes in *Brassica rapa*, *B. napus*, and *Arabidopsis thaliana*. *Genetics* **146:** 1123–1129.

Osborn TC, Lukens L. 2003. The molecular genetic basis of flowering time variation in *Brassica* species. In: Nagata T, Tabata S, eds. *Brassicas and legumes: from gene structure to breeding*. Berlin: Springer-Verlag, in press.

Otto SP, Whitton J. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* **34:** 401–437.

Ozkan H, Levy AA, Feldman M. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops–Triticum*) group. *Plant Cell* **13:** 1735–1747.

Parkin IAP, Lydiate DJ, Trick M. 2002. Assessing the level of collinearity between *Arabidopsis thaliana* and *Brassica napus* for *A. thaliana* chromosome 5. *Genome* **45:** 1–11.

Parkin IAP, Sharpe AG, Keith DJ, Lydiate DJ. 1995. Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). *Genome* **38:** 1122–1131.

Pires JC, Zhao J, Schranz ME, Leon EJ, Quijada PA, Lukens LN, Osborn TC. 2004. Flowering time divergence and genomic rearrangements in resynthesized polyploids (*Brassica*). *Biological Journal of the Linnean Society* **82:** 675–688.

Prakash S. 1974. Probable basis of diploidization of *Brassica juncea* Coss. *Canadian Journal of Genetics and Cytology* **16:** 232–234.

Quijada PA. 2003. Introgression of germplasm from winter into spring *Brassica napus*: detection and confirmation of

quantitative trait loci. PhD dissertation, University of Wisconsin.

Quiros CF, Grellet F, Sadowski J, Suzuki T, Li G, Wroblewski T. 2001. *Arabidopsis* and *Brassica* comparative genomics: sequence, structure, and gene content in the *ABI1–RPS2–CK1* chromosomal segment and related regions. *Genetics* **157:** 1321–1330.

Ramsey J, Schemske DW. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* **33:** 589–639.

Röbbelen G. 1960. Beiträge zur Analyse de *Brassica*-Genomes. *Chromosoma* **11:** 205–228.

Rossberg M, Theres K, Acarkan A, Herrero R, Schmitt T, Schumacher K, Schmitz G, Schmidt R. 2001. Comparative sequence analysis reveals extensive microcolinearity in the *Lateral suppressor* regions of the tomato, *Arabidopsis*, and *Capsella* genomes. *Plant Cell* **13:** 979–988.

Ryder CD, Smith LB, Teakle GR, King GJ. 2001. Contrasting genome organization: two regions of the *Brassica oleracea* genome compared with collinear regions of the *Arabidopsis thaliana* genome. *Genome* **44:** 908–817.

Schranz ME, Osborn TC. 2000. Novel flowering time variation in the resynthesized polyploid *Brassica napus*. *Journal of Heredity* **91:** 242–246.

Schranz ME, Osborn TC. 2004. De novo variation in life-history traits and responses to growth conditions of resynthesized polyploid *Brassica napus* (Brassicaceae). *American Journal of Botany* **91:** 174–183.

Schranz ME, Quijada P, Sung S-B, Lukens L, Amasino R, Osborn TC. 2002. Characterization and effects of the replicated flowering time gene *FLC* in *Brassica rapa*. *Genetics* **162:** 1457–1468.

Sears ER. 1977. An induced mutant with homoeologous pairing in common wheat. *Canadian Journal of Genetics and Cytology* **19:** 585–593.

Shaked H, Kashkush K, Ozkan H, Feldman M, Levy AA. 2001. Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* **13:** 1749–1759.

Sharpe AG, Parkin IAP, Keith DJ, Lydiate DJ. 1995. Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). *Genome* **38:** 1112–1121.

Simillion C, Vandepoele K, Van Montagu MCE, Zabeau M, Van de Peer Y. 2002. The hidden duplication past of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* **99:** 13627–13632.

Song K, Lu P, Tang K, Osborn TC. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proceedings of the National Academy of Sciences, USA* **92:** 7719–7723.

Song K, Osborn T. 1992. Polyphyletic origins of *Brassica napus*: new evidence based on organelle and nuclear RFLP analyses. *Genome* **35:** 992–1001.

Sperisen C, Ryals J, Meins F. 1991. Comparison of cloned genes provides evidence for intergenomic exchange of DNA in the evolution of a tobacco *glucan endo-1,3-B-glucosidase* gene family. *Proceedings of the National Academy of Sciences, USA* **88:** 1820–1824.

Truco M, Hu J, Sadowski J, Quiros CF. 1996. Inter- and intra-genomic homology of the *Brassica* genomes: implications for their origin and evolution. *Theoretical and Applied Genetics* **93:** 1225–1233.

U N. 1935. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Japanese Journal of Botany* **7:** 389–452.

Udall J. 2003. A genetic study of oilseed *Brassica napus* L. Mapping chromosome rearrangements and quantitative trait loci. PhD dissertation, University of Wisconsin.

Wendel JF. 2000. Genome evolution in polyploids. *Plant Molecular Biology* **42:** 225–249.