DETERMINING THE FATE OF HYBRIDIZED GENOMES IN THE
ALLOPOLYPLOID BRASSICA NAPUS

A Thesis presented to
the Faculty of California Polytechnic State University,
San Luis Obispo

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Biological Sciences

by
Tina Ying Wang
July 2010

© 2010
Tina Ying Wang
ALL RIGHTS RESERVED

COMMITTEE MEMBERSHIP

TITLE: Determining the Fate of Hybridized Genomes in the
Allopolyploid *Brassica napus*

AUTHOR: Tina Ying Wang

DATE SUBMITTED: July 2010

COMMITTEE CHAIR: Edward Himelblau, Ph.D.

COMMITTEE MEMBER: Elena Keeling, Ph.D.

COMMITTEE MEMBER: Peter Jankay, Ph.D.

ABSTRACT

Determining the Fate of Hybridized Genomes in the Allopolyploid *Brassica napus*

Tina Ying Wang

Polyploidy is widely acknowledged as a widespread mechanism in the evolution and speciation of the majority of flowering plants. Allopolyploid forms through interspecific hybridization and whole genome duplication. While allopolyploids may display increased vigor relative to their progenitors, they can also face challenges to fertility following hybridization. Genetic changes in allopolyploids result from recombination between the hybridized subgenomes, which can influence phenotype and ultimately determine fitness of future generations. To study dynamic changes that follow allopolyploid formation, *Brassica napus* lineages were derived by hybridizing *Brassica oleracea* and *Brassica rapa*. Two lineages of *B. napus* were analyzed for genetic and phenotypic changes in the S₂, S₇, and S₁₂ generations. Although these lineages were genetically identical at the time of hybridization, divergence was apparent by the S₂ generation. There was a significant increase in sequence loss across generations within both lineages. Four of six generations from both lineages displayed no significant differences to each other in sequence loss relative to the parental generation. In both lineages, there was a bias towards losing sequences from the *B. oleracea* subgenome. Some individual plants showed novel phenotypes; however, there was no correlation between the examined genetic changes and selected phenotypes.

Keywords: allopolyploids, genetic changes, hybridization, phenotypes

ACKNOWLEDGMENTS

I would like to thank first and foremost my thesis adviser, Dr. Edward Himelblau, who shed light on many angles of this project and for pushing me to succeed beyond what I thought was already over the edge. I could not ask for a better adviser, mentor, and friend. I would like to thank my committee members Dr. Elena Keeling and Dr. Peter Jankay, who have given me great advice in designing my project and data analysis. I would also like to thank Lauren Kearney, who contributed greatly to this project and has stayed by me through some very rough times during data collection. I would also like to thank Maya Benavides, Alan Chavez, and Emily Smith for phenotype data collection during sweltering hours under the merciless sun. In addition, I could not have analyzed my data without the help of statisticians Dr. Rebecca Doerge, Dr. Soma Roy, and Dr. John Walker. Last but certainly not least, I would like to thank Dr. J. Chris Pires and Dr. Robert Gaeta for their input on the experimental design and for providing the initial seeds that jumpstarted my project. For funding, I would like to thank the National Science Foundation Plant Genome Program (grant no. 0077774) and Cal Poly College Based Funds.

TABLE OF CONTENTS

Page

LIST OF TABLES.....	ix
LIST OF FIGURES.....	x

CHAPTER 1: POLYPLOID OVERVIEW

Introduction.....	1
I. Origin and Prominence of Polyploids.....	1
i. <i>Formation of Polyploids</i>	1
ii. <i>Paleopolyploids and Diploidization</i>	3
iii. <i>Plant Polyploids</i>	5
iv. <i>Animal Polyploids</i>	8
II. Implications for Speciation and Genome Evolution.....	8
III. Consequences of Allopolyploidy.....	11
i. <i>Homoeologous Nonreciprocal Transpositions (HNRTs)</i>	13
ii. <i>The Polyploid ‘Ratchet’</i>	14
iii. <i>Epigenetic Changes</i>	15
iv. <i>Phenotypic Variation</i>	18
IV. Studies of Resynthesized Allopolyploids.....	18

CHAPTER 2: SEQUENCE LOSS AND PHENOTYPIC VARIATION WITHIN RESYNTHESED LINEAGES OF *BRASSICA NAPUS* ALLOPOLYPLOIDS

I. Abstract.....	23
II. Introduction.....	24
III. Materials and Methods.....	27
i. <i>Plant Material</i>	27
ii. <i>Genetic Analysis and Microsatellite Fragment Amplification</i>	29
iii. <i>Statistical Analyses of Genetic and Phenotypic Data</i>	30
IV. Results.....	31

i. <i>Detection of Sequence Loss in the Resynthesized Brassica napus Genome</i>	31
ii. <i>Sequence Loss Relative to the Parent</i>	34
iii. <i>Total Sequence Loss in Each Lineage</i>	35
iv. <i>Loss of B. oleracea or B. rapa Subgenome</i>	37
v. <i>Phenotypic Variation Across Selfed Lineages of B. napus</i>	38

V. Discussion

i. <i>Dynamics of Genetic Changes in Resynthesized B. napus</i>	43
ii. <i>Sequence Loss Relative to the Parent</i>	44
iii. <i>Directional Change in Resynthesized B. napus</i>	55
iv. <i>Reappearance of Lost Parental Bands</i>	57
v. <i>Phenotypes and Sequence Loss</i>	58
vi. <i>Directions in Future Research</i>	58

CHAPTER 3: PHENOTYPE DATA

1. Number of leaves at flowering.....	64
2. Plant height of first flower (cm).....	66
3. Raceme height (cm).....	68
4. Total plant height (cm).....	70
5. Number of secondary branches.....	72
6. Seed mass (g).....	74
7. Number of siliques on the primary stem.....	76
8. Number of siliques with more than one seed.....	78
9. Leaf length (cm) of third true leaf.....	80
10. leaf width (cm) of third true leaf.....	82
11. petiole length (cm) of third true leaf.....	84
12. Leaf perimeter (cm) per surface area (cm ²).....	86

BIBLIOGRAPHY	88
--------------------	----

LIST OF TABLES

Table	Page
1. Chi-square goodness-of-fit test.....	34
2. F-tests and <i>P</i> -values.....	41

3. F-tests and P -values for phenotypic variance.....	63
4. Bartlett's test and P -values for phenotypic variance.....	63

LIST OF FIGURES

Figure	Page
1. Several mechanisms that give rise to auto- and allopolyploids.....	3
2. Diploidization of polyploids.....	4
3. Formation of a segmental polyploid.....	7
4. Gametes of duplicated genomes in auto- and allopolyploids.....	12
5. Meiotic products in bivalents and multivalents of allopolyploids.....	13
6. Homoeologous non-reciprocal transpositions (HNRTs) in allopolyploids.....	14
7. Genetic and epigenetic changes in polyploid genomes.....	17
8. Newly formed <i>Brassica napus</i> allopolyploids.....	20
9. Origins of <i>Brassica napus</i> generations.....	28
10. Marker sequence loss or gain.....	32
11. Possible progeny of the S ₂ at the A7 chromosome.....	33
12. Percent sequence lost relative to the parent.....	35
13. Total percent sequence loss.....	36
14. Percent sequence loss relative to diploid progenitors in lineage 5.....	37
15. Percent sequence loss relative to diploid progenitors in lineage 78.....	38
16. Novel phenotypes and phenotypic variation in resynthesized <i>B. napus</i>	40
17. Variation in phenotype at the S ₂ S ₇ and S ₁₂ generation.....	41
18. 3:1 ratio of segregation patterns.	45
19. Possible chromosome patterns found in allotetraploids.....	47
20. <i>Brassica napus</i> chromosome map with 23 microsatellite markers used to map genetic changes and sequence loss.....	48
21. Data analysis of number of leaves at flowering.....	64

22. Data analysis of plant height to first flower (cm).....	66
23. Data analysis of raceme height (cm).....	68
24. Data analysis of total plant height (cm).....	70
25. Data analysis of number of secondary branches.....	72
26. Data analysis of seed mass (g).....	74
27. Data analysis of number of siliques on primary stem.....	76
28. Data analysis of number of siliques with more than one seed.....	78
29. Data analysis of leaf length (L3).....	80
30. Data analysis of leaf width (L3).....	82
31. Data analysis of petiole length (L3).....	84
32. Data analysis of perimeter (cm) per surface area (cm ²).....	86

CHAPTER 1: POLYPLOID OVERVIEW

Introduction

Polyploidy is a genetic state in which an organism possesses more than two complete sets of chromosomes. These sets of chromosome can be derived from duplication of a single genome (autopolyploidy) or by hybridization of distinct genomes (allopolyploidy). Many eukaryotic organisms including fungi, plants, and animals show evidence of polyploidy at some point in their evolutionary history. Polyploidy confers both opportunities and challenges on organisms. The potential success of newly formed polyploids is attributed to their highly plastic genome structure, higher tolerance to mutations, and novel patterns of gene expression (Osborn et al., 2003). Interspecific hybridization in allopolyploids can also increase genetic diversity, allowing new niche exploitation. However, newly formed polyploids face challenges to fertility that originate during meiosis and lead to unbalanced gametes and aneuploid progeny. Despite these challenges, it is clear that polyploidy has played an important role in eukaryotic speciation and biodiversity.

I. Origin and Prominence of Polyploids

Formation of Polyploids

Polyploids can be classified either as allopolyploids or autopolyploids based on the origins of their progenitors (Fig. 1.1). Autopolyploids are formed by genome doubling of a single diploid progenitor prior to gamete formation, resulting in doubling the genome in the progeny (Fig. 1.1A) (Chen, 2007). Formation of allopolyploids

involves interspecific hybridization between different genomes followed by chromosome doubling (Fig. 1.1B) or fertilization of unreduced gametes, or hybridization between genomes that have already duplicated (Fig. 1.1C) (Chen, 2007). The conditions required for the establishment of allopolyploids are (1) the existence of diploid species with distinct genomes, (2) natural ability for the hybridization of two genomes, and (3) a mechanism to increase the fitness of polyploid progenies (Chapman and Burke, 2007). The division between allopolyploids and autopolyploids is dependent on the evolutionary relationship between the diploid parents; the more similar the genomes that are being hybridized, the closer the allopolyploid becomes to an autopolyploid (Comai, 2005). The underlying difference between an autopolyploid and an allopolyploid is that an autopolyploid results from duplication of a single genome, whereas allopolyploids result from interspecific hybridization (i.e. hybridization of two distinct genomes).

Figure 1.1. Several mechanisms that give rise to auto- and allopolyploids. Pairs of homologous chromosomes are shown in each diploid. **(A)** An autopolyploid is formed by whole genome duplication of a single diploid genome. Two paths to allopolyploid formation are shown: **(B)** An allopolyploid is formed by interspecific hybridization of two species followed by chromosome doubling of the sterile hybrid. **(C)** An allopolyploid can also be formed through interspecific hybridization of two autopolyploids.

Paleopolyploids and Diploidization

Many modern organisms that behave genetically as diploids are actually paleopolyploids (ancient polyploids), which were derived from at least one whole genome duplication (WGD) event followed by chromosome rearrangement, gene loss, and genomic reorganization. Most species of flowering plants and vertebrates have descended from ancestors that were once autopolyploids or allopolyploids (Wolfe, 2001).

In fact, several studies have revealed that polyploidization events are common to many eukaryotic lineages. Over evolutionary time, many polyploids have undergone extensive chromosomal reorganization, causing the genome to behave as diploid (i.e. all chromosomes have a single homolog) (Fig. 1.2). Genome reorganization can lead to diploidization. For example, a set of four identical chromosomes derived from WGD diverge into two pairs of homologous chromosomes (Van de Peer, 2009). Although the molecular basis for diploidization is poorly understood, it presumably occurs through DNA sequence changes and/or deletions between homologous regions of chromosomes (Wolfe, 2001).

Figure 1.2. Diploidization of polyploids. Diploids undergo a polyploidization event through whole genome duplication. Over evolutionary time, segmental loss of duplicated genes and genome reorganization result in diploidization. Following diploidization, many loci (such as the locus represented by the yellow stripe) are present in multiple copies.

Susumu Ohno's book *Evolution by Gene Duplication* in the 1970s proposed that two or three rounds of polyploidy had occurred in the vertebrate lineage. Comparative genomics confirm that mammalian genomes contain regions of synteny, remnants of ancient duplication events (Wolfe, 2001; Bailey et al., 2004). The frequency of paleopolyploidy in the eukaryotic record suggests that most eukaryotes either benefited from or survived with polyploidy (Comai, 2005).

Plant Polyplids

Evidence from cytogenetic analyses, studies of fossils and extant species, and more recently, comparative genomics show that 60-70% of flowering plants have polyploid ancestry. Interspecific hybridization, required for allopolyploid formation, occurs frequently in plant taxa and have been observed in *Brassica*, *Gossypium*, *Senecio*, *Spartina*, *Tragopogon*, and *Triticum* (Osborn et al., 2003; Meyers and Levin, 2006). Many important crop plants are autopolyploids (e.g. alfalfa, and potato) or allopolyploids (e.g. wheat, oat, cotton, coffee, and canola) (Leitch and Leitch, 2008). Even plants with known small genomes such as *A. thaliana*, have undergone a WGD event (The *Arabidopsis* Genome Initiative, 2000; Adams and Wendel, 2005). Evidence from numerous plant studies show that recurrent genome doubling occurred after the divergence of eudicots from monocots, and a second polyploidization event following the divergence of *Brassica* and *Arabidopsis* from their common ancestor Malvaceae (Adams and Wendel, 2005). Genetic mapping of *Zea maize* (corn) (Ilic et al., 2003), *Glycine max* (soy) (Walling et al., 2006), *Brassica rapa*, and *B. oleracea* (Iniguez-Luy et al., 2009) has shown that these present diploid species are, in fact, paleopolyploids that have undergone

diploidization. The modern maize genome shows evidence that modern maize is a product of a segmental allotetraploid (translocations that occurred between the hybridized subgenome) (Fig. 1.3) formed ~11.4 mya that underwent extensive diploidization (Soltis and Soltis, 1999). Likewise, the *Arabidopsis* and *Lycopersicon* (tomato) genome have been estimated to have undergone a WGD ~112 mya to become a tetraploid (The *Arabidopsis* Genome Initiative, 2000; Blanc and Wolfe, 2004). Although these organisms behave genetically as diploids today, genomic data indicate that many present diploids have descended from a polyploid ancestor.

Figure 1.3 Formation of a segmental polyploid. Genomic changes during homoeologous recombination result in exchanges of genetic material between distinct genomes during meiosis.

Animal Polyploids

Polyploidy is observed presently in fish, amphibians, and other vertebrates, Although polyploidy is much rarer in animals than in plants, there are hundreds of known insects and vertebrate species, mainly fish and amphibians, that are polyploid (Van de Peer, 2009). Species that reproduce by selfing or asexually are more tolerant to polyploidy (Otto and Whitton, 2000). Chromosomal sex determination systems are important to reproduction in animals, whereas most plants (excluding some dioecious plants) do not possess sex chromosomes. One of the biggest barriers for successful establishment of polyploidy in sexual species is the requirement for a genetically compatible partner. However, polyploidy results in either direct disruption of sex determination and sexual development or indirect effects on gene regulation and dosage compensation of sex-linked genes (Dufresne and Hebert, 1994; Otto and Whitton, 2000). Polyspermy (fertilization of an egg by more than one sperm) in humans causes triploidy, and in most cases, forces automatic or early abortion of the fetus due to chromosomal abnormalities (McFadden et al., 1993). The rarity of polyploidy in animals is attributed to the formation of unviable gametes. Despite the problems associated with polyploidy in animals, it has clearly been an important part of animal evolution, as the vertebrate lineage has undergone WGD and present vertebrates are paleopolyploids (Wolfe, 2001).

II. Implications for Speciation and Genome Evolution

There is a correlation between whole genome duplication events and increased rates of speciation or divergence (Soltis et al., 2009). Studies examining newly formed polyploid populations have found higher genetic diversity and higher vigor compared to related diploids (Otto and Whitton, 2000). Since hybrids are sterile due to chromosome

incompatibility during meiosis, WGD allows them to produce viable offspring (Levin, 1982). Polyploidization can also enhance variation with novel allele combinations, creating the conditions for rapid speciation and evolution (Mallet, 2007).

Sympatric speciation is dependent on the appearance of genetic and reproductive barriers between members of a population. Polyploids are often reproductively isolated from their progenitors. For example, a tetraploid and a diploid of the same species will hybridize to form a triploid that is likely to exhibit low fertility. Thus, polyploidy is a mechanism by which reproductive barriers can form (Rieseberg, 1997) and may represent the most common mechanism for sympatric speciation in plants (Otto and Whitton, 2000). In addition to creating barriers to reproduction polyploids may be phenotypically distinct from their progenitors. Novel phenotypes may allow polyploids to exploit niches not available to their progenitors or provide polyploids with adaptive advantages (Schranz and Osborn, 2000). The most compelling evidence that polyploidy confers adaptive advantages is that the timing of WGD in many taxa of flowering plants dates to ~60-70 mya at the approximate time of the massive climatic event of the K/T boundary. The implication is that polyploids may have been better able to adapt to rapidly changing climate conditions while related diploids became extinct, possibly due to higher tolerance of a wide range of environmental conditions (Fawcett et al., 2009). Thus, polyploidy is a driving force in plant evolution with profound effects at the molecular as well as the ecological level.

Indeed, polyploids can survive better in harsh environments, such as extreme altitudes or temperatures, which allow them to prolong their lifespan and increase diversity. For example, plant polyploids may possess traits that increase levels of drought tolerance, apomixis, and pest resistance which may ultimately increase their fitness and

make them better adapted to a wider range of niches (Chen, 2007). In addition, the successful formation of polyploids does not only involve the merging of two genomes, but also requires physiological adjustments. Some metabolic pathways upregulated in recently-formed polyploids overlap with plant stress response. The allopolyploid *A. suecica* plants have altered circadian clock regulation that mediate metabolic and physiological pathways resulting in increased starch and chlorophyll content (Ni et al., 2009). Upregulation of metabolic pathways allows greater vigor and biomass in the *A. suecica* allopolyploid than its diploid progenitors grown under the same environmental conditions. These examples represent preliminary evidence that polyploid organisms have selective advantages over diploid progenitors, particularly under adverse environmental conditions.

Heterosis (hybrid vigor) is a phenomenon in which hybrid phenotypes are more robust than their parents. Merging of genomes in allopolyploid formation increases heterozygosity in the hybrid genome, which appears to contribute to heterosis (Adams, 2007). On the other hand, autopolyploids, organisms that contain the same alleles for all duplicated genes, maintaining homozygosity, tend to be less vigorous than diploid progenitors (Birchler, unpublished). Heterozygosity in the allopolyploid genome contributes to heterosis through the complementation of deleterious alleles and through changes in gene expression relative to diploid progenitors. This is an important mechanism when spatial isolation or a bottleneck is reached, where populations are forced to inbreed and the gene pool cannot be rid of deleterious mutations (Comai, 2005).

III. Consequences of Allopolyploidy

Despite clear evidence that polyploid plants have been successful over evolutionary time, individuals that undergo WGD or hybridization to become polyploid face challenges to genome stability and fertility. Newly formed polyploids experience extensive genomic rearrangements, translocations, changes in gene expression, changes in chromatin structure, and epigenetically induced gene silencing (Adams and Wendel, 2005). Hybridization may allow interspecific reproduction, but does not ensure that the offspring will be fertile.

Homoeologous Chromosome Pairing

Synteny (regions of similarity) found on the chromosomes of closely related species allow recombination between the homoeologous chromosomes (chromosomes that have a common ancestor that coexist in the same allopolyploid genome). Meiotic aberrations caused by spindle and cytokinesis malfunction have been implicated for causes of $2n$ or more chromosomes in gametes (Otto and Whitton, 2000). In an allopolyploid, homoeologous chromosomes form multivalents in meiosis I (instead of bivalents as in a diploid). Multivalent formation can lead to inaccurate segregation and aneuploid gametes (Fig. 1.4), and formation of segmental allopolyploids (Fig. 1.5) (Udall et al., 2005). Crossing over between homeologs may lead to reciprocal exchange (RE) between subgenomes, which are detected in newly resynthesized allopolyploids (Udall et al., 2005; Gaeta et al., 2007). Homoeologous recombination between repetitive regions of chromosomes destabilizes the genome resulting in deletions, chromosomal rearrangements, and gene conversion (Rieseberg, 2001; Gaeta et al., 2007; Gaeta and Pires, 2010), and incorrect segregation of homoeologous pairs can lead to aneuploidy. (Comai et al., 2000) (Fig. 5). As a result, many allopolyploids of closely related species

frequently produce unviable gametes (Comai, 2005). Although allopolyploids often show heterosis, they may have lower fertility, which becomes the limiting factor in reproductive success.

Figure 1.4. Gametes of duplicated genomes in auto- and allopolyploids.

Recombination can lead to multivalent formation, which can leave several chromosomes unpaired resulting in unbalanced gametes. Possible gametes from meiotic aberrations occur in autopolyploid (left) and allopolyploids (right), resulting in aneuploid gametes (Comai, 2005).

Figure 1.5. Meiotic products in bivalents and multivalents of allopolyploids. (A) Bivalent pairing during metaphase I result in equal segregation of chromosomes during meiosis II. **(B)** As multivalents undergo recombination during meiosis I, complex configurations result in homoeologous exchanges, often uneven during cell division. This causes aneuploid gametes after meiosis II, which can account for the loss or gain of whole chromosomes in an individual. Homoeologous non-reciprocal transpositions (HNRTs) as well as reciprocal exchanges (REs) can result from multivalent formation. Shown here are some of the possible gametes.

Homoeologous Nonreciprocal Transpositions (HNRTs)

While genome duplication is a common event leading to polyploid formation, few modern species retain exact duplicate sets of chromosomes. This is due to the accumulation of mutations (e.g. chromosomal deletions and rearrangements) following allopolyploidization. Complications arise during recombination between homoeologous loci during meiosis of cell lineages destined to become gametes. One cause of such mutations is due to homoeologous non-reciprocal transpositions (HNRTs). HNRTs result in an uneven exchange of genetic material between homoeologous chromosomes (Fig. 1.6). This happens in allopolyploids, where homoeologs pair and unequal exchange occurs. Recent molecular analysis of *Brassica* and *Gossypium* show evidence of HNRTs in the progenies of these allopolyploids via non-reciprocal crossovers (Udall et al., 2005; Gaeta et al., 2007; Salmon et al., 2010). In allopolyploids, genetic changes associated with HNRTs have been used to explain the loss of duplicate gene expression and function (Chen, 2007). This loss is due to recombination of homoeologous chromosomes, where there might be unequal exchange of genetic material that causes increased mutations and broken chromosome segments that cannot be paired.

Figure 1.6. Homoeologous non-reciprocal transpositions (HNRTs) in allopolyploids. When homoeologous chromosomes pair during meiosis, recombination can lead to homoeologous non-reciprocal transpositions (HNRTs). As a result, chromosomal material from one homoeolog replaces genetic material of the other.

The Polyploid 'Ratchet'

For an allopolyploid to maintain high fertility, homoeologous pairing during meiosis must be limited. Any genetic change that increases the likelihood that non-homologous chromosome will pair during meiosis will therefore have a negative effect on fertility. REs and HNRTs represent such changes. When RE or HNRT occurs in one generation, meiosis in the next generation is likely to be characterized by greater mispairing, leading to a higher incidence of aneuploid gametes. Gaeta and Pires (2010) coined the term "polyploid ratchet" to describe this phenomena. The term 'polyploid ratchet' is based on Muller's ratchet, which explains the increased accumulation of mutations in selfing or asexual populations (Muller, 1964; Haigh, 1978). A ratchet, a mechanical device that allows movement in only one direction, is an effective analogy for the genetic consequences of allopolyploidy; once a genetic change occurs the likelihood of future deleterious changes increases. REs and HNRTs 'ratchet' the allopolyploid genome towards greater instability.

Epigenetic Changes

Genetic changes can explain the cause of chromosomal deletions over time, but gene silencing can be epigenetically controlled. Epigenetic changes are the immediate effects of allopolyploidization that can reactivate transposons and/or silence duplicated genes, which accelerate further genomic changes following the onset of interspecific hybridization. Changes in gene expression in allopolyploids can result from REs and

HNRTs, but can also be attributed to epigenetics (Fig. 1.7). Chemical modification of DNA results in heritable changes in gene expression. When two different genomes are hybridized, simple doubling of gene expression is rarely observed (i.e. non-additive gene expression is observed). One parental gene will often be expressed while the homoeologous locus will be silenced (Chen, 2007). These changes in gene expression are explained by changes to the chromatin structure in specific regions of the subgenomes; a locus in one genome undergoes epigenetic silencing while the homoeologous locus in the other subgenome is expressed. Studies in *Arabidopsis* have revealed that thousands of genes show non-additive gene expression following polyploid formation (Wang et al., 2006; Ha et al., 2009). It is not clear what mechanism determines whether non-additive expression will occur at a given locus and there is yet a basis for predicting which subgenome will be silenced.

Figure 1.7. Genetic and epigenetic changes in polyploid genomes. Genetic changes include sequence mutations and chromosomal instabilities (deletions/insertions, translocation, transpositions, etc.). Homoeologous chromosomes can exchange equal genetic material, or lose partial regions. Epigenetic changes regulate the chromatin structure of DNA, leading to the activation or silencing of genes. Black bars represent regions of gene inactivation.

Another consequence of the epigenetic changes that follow polyploid formation is mobilization of transposons, which are usually silenced by methylation. WGD events seem to result in the demethylation of some transposable elements, allowing them to move freely within the genome (McClintock, 1984). Rapid changes in retroelement activity destabilize the genome (Parisod et al., 2010). In addition, WGD can also cause

methylation of genes in allopolyploids, silencing genes that are present in multiple copies in the genome. Thus, changes in the polyploid epigenome can influence the success of future generations.

Phenotypic Variation

Genetic and epigenetic changes in allopolyploids will ultimately be reflected in the phenotype. Populations of plant allotetraploids in *Arabidopsis* and *Brassica* show variation in flowering time, fertility, and stature (Comai, 2000; Pires et al., 2004). For any given trait the phenotype can be dominated by one progenitor, be an intermediate between progenitors, or represent new variation relative to the progenitors (Comai, 2000). Novel phenotypes often arise in the generations immediately following hybridization (Pires et al., 2004). These novel phenotypes are heritable, and evolution of *de novo* traits may have contributed to the success and diversification of allopolyploids (Schranz and Osborn, 2000). The dynamic appearance of these novel phenotypes depends on the area in which chromosomal rearrangements are located, and the rate at which deletions in the genome occur.

IV. Studies of Resynthesized Allopolyploids

Investigating naturally existing populations of polyploids at the genetic level is problematic because the exact parental diploid genomes are not known or have diverged since formation of the polyploid. In addition, there could have been multiple polyploidy events that occurred in a species, making it difficult to untangle their evolutionary history (Pires et al., 2004). Thus, polyploids with known progenitors represent effective models for understanding genomic changes and changes in gene expression that follow genome

duplication and hybridization. Polyploid research can also aid efforts to understand the various modes of ecological and evolutionary complexities of speciation (Rieseberg and Willis, 2007). Researchers have turned to recently formed or resynthesized (generated by breeding) polyploids to study the dynamics of polyploid genomes. These polyploids include *Triticum* (wheat), *Gossypium* (cotton), *Brassica*, *Tragopogon*, *Arabidopsis*, *Glycine max* (soybean), and *Nicotiana* (tobacco). These have become experimental model organisms for addressing the fate of duplicated genomes. The known progenitors of these allopolyploids can be hybridized in the lab and the subsequent genetic and epigenetic changes can be observed in early generations (Adams and Wendel, 2005).

The consequences of allopolyploidy on genome evolution, gene expression, and phenotype have been under investigation using the allopolyploid *Brassica napus* as a model system. Lukens et al. (2006) analyzed 50 *B. napus* (n=19) lineages from homozygous genomes of *Brassica oleracea* (egg-donor; n=9) and *Brassica rapa* (pollen donor; n=10). As *B. rapa* (A-subgenome) and *B. oleracea* (C-subgenome) are hybridized, the sterile hybrid (CA) may undergo genome duplication, giving rise to *B. napus* (CCAA). Genome doubling in the hybrids was either induced by colchicine treatment or occurred spontaneously to create the S₀ generation of *B. napus* (Lukens et al., 2006; Yu et al., 2009). Each lineage of *B. napus* represents a unique fertilization event. Because all 50 *B. napus* plants were derived from the same diploid progenitors, they are expected to be genetically identical at the time of hybridization. As lineages of *B. napus* propagate through self-pollination, they begin to diverge both at the genotypic and phenotypic level (Fig. 1.8).

Figure 1.8. Newly formed *Brassica napus* allopolyploids. Doubled haploids *B. oleracea* and *B. rapa* hybridize to produce *B. napus*, which undergoes genome duplication to form the 50 genetically identical lineages at the S₀ generation. Each independently derived and genetically identical lineage was self-pollinated to propagate through the S₅ generations (J. Chris Pires).

The first generation (S₀) of self-pollinating *B. napus* lineages were screened for epigenetic and genetic changes (Lukens et al., 2006). Genetic changes that altered the underlying genetic structure in the S₀ generation were found to be rare as observed using microsatellite markers (i.e. the genome was highly conserved). DNA methylation patterns were extensively altered in the *B. napus* individuals in the S₀ generation relative to the diploid progenitors. The frequency of methylation changes was similar across all ~50

lines in the S_0 generation, indicating epigenetic changes are the immediate effects of allopolyploidization in *B. napus*.

The S_5 generations of the same lineages described above were analyzed for genetic changes, epigenetic changes, gene expression, and phenotypic variation (Gaeta et al., 2007). By the S_5 generation, genetic changes such as chromosomal rearrangements, transpositions, and deletions were common. In contrast to the S_0 generation, individuals from S_5 exhibited a higher frequency of DNA fragment loss at marker loci, most likely due to HNRTs. Although genetic changes were more common in the S_5 relative to the S_0 generation, the epigenetic changes established in the S_0 generation appeared to be fixed throughout subsequent generations. REs and HNRTs in later generations created novel allele combinations, which resulted in further phenotypic variation and divergence among the lineages (Gaeta et al., 2007).

In summary of the observations made on the S_0 and S_5 generations, significant epigenetic changes occur in individual lines immediately after hybridization of the *B. rapa* and *B. oleracea* genomes to create *B. napus*, with rare genetic changes in the S_0 generation. This was followed by increased genetic changes and fixed epigenetic changes observed by the S_5 generation. Genetic and phenotypic analysis included only the S_0 and S_5 generation of *B. napus*. This system provides experiments for further analyses of later generations to observe further changes in the allopolyploid lineages. It is not known at which point during the first five generations the majority of genetic changes occurred. Also, it is not known whether the S_5 genomes have lost the ability to undergo further rearrangements (i.e. the genomes of most lineages have stabilized). Analysis of later generations will add important insights into genetic changes of *B. napus*. It remains to be determined whether lineages maintain the same potential to undergo genetic change. If

genetic changes persist in later generations, it might also have an effect on phenotype. Future research will continue to shed light on the fate of duplicated genomes and elucidate events that influence the evolution of allopolyploids. Genetic changes and phenotypic variation are observed in the S_2 , S_7 , and S_{12} generations within two lineages of the allopolyploid *B. napus*.

CHAPTER 2: SEQUENCE LOSS AND PHENOTYPIC VARIATION WITHIN RESYNTHESIZED LINEAGES OF *BRASSICA NAPUS* ALLOPOLYPLOIDS

Abstract

Polyploidy is widely acknowledged as a widespread mechanism in the evolution and speciation of the majority of flowering plants. Allopolyploid forms through interspecific hybridization and whole genome duplication. Newly formed allopolyploids often exhibit genome instability. Genetic changes in allopolyploids result from recombination between the hybridized subgenomes, which can influence phenotype and ultimately determine fitness of future generations. To study dynamic changes that follow allopolyploid formation, *Brassica napus* lineages were derived by hybridizing *Brassica oleracea* and *Brassica rapa*. Two lineages of *B. napus* were analyzed for genetic and phenotypic changes in the S_2 , S_7 , and S_{12} generations. Although these lineages were genetically identical at the time of hybridization, divergence was apparent by the S_2 generation. There was a significant increase in sequence loss across generations within both lineages. Four of six generations from both lineages displayed no significant differences to each other in sequence loss relative to the parental generation, and one generation showed an approximately three-fold higher percent sequence loss compared to other generations. In both lineages, there was a bias towards losing sequences from the *B. oleracea* subgenome. Some individual plants showed novel phenotypes; however, there was no correlation between the examined genetic changes and selected phenotypes.

Introduction

Polyploidy is an important factor in eukaryotic evolution and appears to have been especially important to the speciation of flowering plants (Masterson, 1994; Ramsey and Schemske, 1998; Otto and Whitton, 2000). Interspecific hybridization of distinct

genomes in allopolyploids has the potential to increase genetic diversity, generate new ecotypes or species, and reinforce or break down reproductive barriers. Novel phenotypes relative to the diploid progenitors often arise in the generations immediately following allopolyploidization (Pires et al., 2004) with considerable hybrid vigor displayed by the allopolyploids (Adams, 2007).

While allopolyploids may display increased vigor relative to their progenitors, they can also face challenges to fertility following hybridization. The diploid subgenomes of an allopolyploid are often closely related. Synteny between the subgenomes can lead to homoeologous pairing of chromosomes during meiosis. Homoeologous pairing can then lead to homoeologous recombination (exchange of syntenic chromosome segments), homoeologous nonreciprocal transpositions (HNRTs) (unequal replacement of a segment of one subgenome with another), and multivalent formation. These events contribute to the formation of aneuploid gametes ultimately reducing the fertility of allopolyploids. Even if fertility is maintained, changes in subgenome chromosome dosage produced by HNRTs and aneuploidy alter gene expression and may potentially alter phenotype (Osborn et al., 2003; Pontes et al., 2004; Gaeta et al., 2007).

Because of the prevalence of plant polyploids, there is interest in understanding the early events following polyploid formation. Diploid progenitors of many polyploid species in natural populations are unknown. Without knowledge of the exact genotype of progenitors, early events shaping the genomes of allopolyploids are difficult to discern. Therefore, the use of resynthesized allopolyploids with well-characterized diploid parents allows direct comparison between known diploid progenitors and their polyploid progeny.

Evidence for rapid genetic changes following hybridization has been observed in resynthesized allopolyploids such as *Tragopogon* (Soltis and Soltis, 1993; Lim et al., 2008), *Arabidopsis* (Madlung et al., 2002; Pontes et al., 2004), *Triticum* (Ozkan et al., 2001; Shaked et al., 2001), *Gossypium* (Wendel et al., 1995), *Brassica* (Song et al., 1993; Song et al., 1995; Schranz and Osborn, 2000; Pires et al., 2004; Lukens et al., 2006; Gaeta et al., 2007), and among natural species (Song et al., 1993). Homoeologous exchange is now thought to be the major mechanism in driving chromosome deletions, novel combinations of alleles, and increased phenotypic variation in allopolyploids (Pontes et al., 2004; Nicholas et al., 2006; Gaeta et al., 2007; Gaeta and Pires, 2010).

Brassica napus is an allopolyploid formed by interspecific hybridization of *B. rapa* and *B. oleracea*. Previously, 50 lineages of *B. napus* were derived by hybridizing the rapid-cycling *B. rapa* and *B. oleracea* and then inducing whole genome duplication by colchicine treatment. The *B. rapa* and *B. oleracea* progenitors are doubled haploids, and thus the resulting resynthesized *B. napus* were expected to be genetically identical at the time of hybridization (Lukens et al., 2006). The hybridized *B. napus* lineages were propagated through the S₅ generations, where detailed genetic and epigenetic analysis was performed on S₀ and S₅ (Gaeta et al., 2007). At the S₀ generation, few genetic changes were observed between the 50 lineages. However, the lineages were distinct in terms of DNA methylation (Lukens et al., 2006). When the same lineages were analyzed at the S₅ generation, genetic changes were common (Gaeta et al., 2007). Many genetic changes involved the replacement of markers from one subgenome with that of the other consistent with HNRTs. While genetic changes were common among the 50 lineages, DNA methylation was revealed to have little change from the S₀. These studies showed that epigenetic changes are common immediately following hybridization and whole

genome duplication, while genetic changes occur in subsequent generations. However, it is not known when during the first five generations the genetic changes occurred. Also, it is not known whether genetic changes continue to accumulate after the S_5 .

In this study, we investigated the frequency of genetic changes among the bulked progeny in populations of S_2 , S_7 , and S_{12} between two of the resynthesized *B. napus* lineages described above. These relatively large populations allow for an analysis of the genetic changes that occur in the early (S_2) and later (S_{12}) generations following allopolyploidization. Individuals were analyzed for loss of either subgenome at 23 polymorphic marker loci. Twelve phenotypic traits were screened for each individual. Genetically identical at the time of hybridization, the lineages diverged both at the genotype and phenotype level as they advanced. Genetic changes occurred in all generations tested and loss of the *B. oleracea* subgenome was more common than the loss of the *B. rapa* subgenome. By comparing a single parent with a large population of siblings, it is possible to identify segregation and the appearance of novel chromosomal changes occurring during meiosis in the parent.

Materials and Methods

Plant Material

Hybridization of *Brassica oleracea* (TO1000, egg donor; C-genome) and *Brassica rapa* (IMB218; pollen donor; A-genome) produced 50 resynthesized *B. napus* allopolyploid plants (CCAA) as described previously by Lukens et al. (2006) and Gaeta et al. (2007). *Brassica rapa* and *B. oleracea* are doubled haploids, and thus are expected to be homozygous at every locus. The CA hybrids produced in the original crosses were treated with colchicine to induce genome doubling that produce resulting in the first allopolyploid generation (S_0). Two lineages that maintained high fertility through the S_{10} generation were chosen for this study. These lineages were designated as EL5 and EL78, and here are referred to as “lineage 5” and “lineage 78.” Two to four S_1 , S_6 , and S_{11} plants from lineage 5 and lineage 78 were self-pollinated. The generations in each lineage are not directly descended from one parent plant (Fig. 2.1). Populations of ~35 plants in each the of S_2 , S_7 , and S_{12} generations were derived from a single parent plant with the highest seed set from the S_1 , S_6 , and S_{11} generations.

Figure 2.1. Origins of *Brassica napus* generations. *B. oleracea* (pollen recipient) and *B. rapa* (pollen donor) were hybridized and doubled to produce *B. napus* (S_0). Each lineage was propagated for 12 generations. The generations analyzed here represent pools of ~35 siblings from a single parent. Note that the generations are not generated from a single plant descendent.

Seeds were germinated in germination mix with 1.5 g Osmocote Plant Food fertilizer (The Scotts Company, Marysville, CA, USA) at California Polytechnic State University, San Luis Obispo, CA, USA. Approximately two weeks after germination, plants were transplanted to 2-gallon pots containing commercial planting mix and organized in a completely random design on an open roof of a building from June 2009 to September 2009. Plants were top-watered using a drip irrigation system with 1L/48 hrs. Plants were top-watered with 0.5 mL/L DynaGro™ fertilizer (Dnya-Gro, Richmond, CA, USA), applied to the surface of the soil every seven days once flowering began. All

plants were sprayed with 0.005% Green Light Malathion Insect Spray (Green Light Company, Longview, TX, USA) every 30 days. Pollination bags were used to encourage self-pollination and prevent cross-pollination between individuals.

The third leaf was removed and photographed at the time of flowering for leaf shape analysis. Young and mature leaves were removed for DNA extraction at the time of flowering and frozen in liquid nitrogen. Plants were screened for 12 phenotypes including number of leaves at flowering, plant height to first flower (cm), raceme height (cm), total plant height (cm), number of secondary branches, seed mass (g), number of siliques on the primary stem, and number of siliques with more than one seed. Digital photographs of the third true leaf were used to measure leaf length, leaf width, petiole length, perimeter, and surface area. The ratio of perimeter to surface was used as a proxy for leaf serration. Water was withdrawn after approximately 10 weeks and plants were allowed to dry. Plants were harvested by cutting off the primary stem while still covered with pollination bags, and kept in a drying room for one week. Seeds were collected by manually crushing the dry siliques, and seed mass was determined thereafter.

Genetic Analysis and Microsatellite Fragment Amplification

Genomic DNA was extracted from leaf tissue of plants using DNAeasy Plant Maxi Kit (Qiagen; Valencia, CA, USA). Microsatellite markers from previous *Brassica* mapping projects (Parkin et al., 2005; Iniguez-Luy et al., 2008; Iniguez-Luy et al., 2009) were screened for polymorphism between *B. rapa* and *B. oleracea*, resulting in a collection of 23 polymorphic markers used to assay marker loss within the resynthesized allopolyploids (see Supplemental Figure 2.1). Fragments were amplified using the following conditions: 1.0 U of GoTaq Flexi DNA Polymerase (Promega Corporation;

Madison, WI, USA), 25 mM MgCl₂, 5X Green GoTaq Flexi Buffer, 2.5 μM of dNTPs, 10 μM each of forward and reverse primers, 10 ng of plant DNA, and dH₂O to final volume of 20 μl. PCR conditions were as follows: 95°C for 30 sec; eight cycles of 94°C for 30 sec, 50°C for 30 sec, 72°C for 50 sec; 22 cycles of 89°C for 30 sec, 50°C for 30 sec, 72°C for 50 sec; 72°C for 30 sec. Each amplification included a negative control of 2 μl dH₂O replacing plant DNA, and a synthetic hybrid consisting of a 1:1 mixture of *B. rapa* and *B. oleracea* DNA. Each generation of ~ 35 samples was amplified along with the parent plant. Fragments were separated using a 6.0% polyacrylamide gel for 50 min at 300V. Fragments were visualized using a UV transilluminator and analyzed using Quantity One 4.6.3 Gel Doc EQ (BioRad, Hercules, CA, USA). All *B. napus* samples that produced a fragment pattern different than the synthetic hybrid were repeated for confirmation.

Statistical Analyses of Genetic and Phenotypic Data

For analysis of DNA fragment loss, the null hypothesis of equal proportion of genetic changes was tested between generations within a lineage with analysis of variance (ANOVA) and between lineages with 2-sample t-tests. Phenotypic differences were tested with *F*-tests on equal variances, t-test for differences between the same generations of each lineage, and ANOVA for comparing differences simultaneously across generations in a lineage. A multivariate analysis of variance (MANOVA) was used to test whether the total DNA fragment losses had an effect on the means of all phenotypic traits. All analyses were conducted using Minitab 15.1.20 (Minitab Inc., State College, PA, USA). Microsoft Excel was used to generate graphs and tables corresponding to DNA fragment losses and phenotypes.

Results

Detection of Sequence Loss in the Resynthesized Brassica napus Genome

Previously, 50 independent lineages of *B. napus* had been generated using well-characterized diploid parents *B. oleracea* (TO1000; A-subgenome) and *B. rapa* (IMB218; C-subgenome) (Lukens et al., 2006; Gaeta et al., 2007). Two of these lineages (lineage 5 and lineage 78) were propagated to the twelfth generation by selfing. Individuals at the S_1 , S_6 , and S_{11} generations were self-pollinated. A total of 205 individual *B. napus* plants representing three generations (S_2 , S_7 , S_{12}), were assayed for genetic changes in two lineages. For each generation, ~35 sibling plants were compared to the single allopolyploid parent plant as well as to the diploid progenitors. All plants were screened with 23 polymorphic markers capable of distinguishing the *B. oleracea* and *B. rapa* subgenomes. Loss of either *B. rapa* or *B. oleracea* subgenome sequences were frequent (Fig. 2.2A). In addition, two individuals showed a gain of a marker sequence that was not present in the parent of siblings (Fig. 2.2B).

Figure 2.2. Marker sequence loss or gain. From left to right in each gel: C (*B. oleracea* subgenome), A (*B. rapa* subgenome), SH (synthetic hybrid), S_1 or S_{11} (S_1 or S_{11} parent), S_2 or S_{12} (S_2 or S_{12} progeny). (A) Genetic changes are indicated by the loss of the *B. oleracea* (C-subgenome) in individuals of the S_2 generation. (B) Genetic changes show a gain of *B.*

oleracea (C-subgenome) in an individual of the S₁₂ generation after it was lost in the parental generation.

For each individual and their parent, total sequence loss was determined. Percent loss was calculated by dividing the number of missing marker signals by the number of expected signals. Total percent sequence loss probably represents an over-estimate of the number of genetic changes occurring in individuals because it is likely that the loss of linked markers (scored here as independent events) are in fact a single event. However, in overall genetic changes, there is an underestimation because not all markers are tested to observe deletions in the entire genome. In seven of 205 individuals, adjacent markers were lost simultaneously, probably resulting from a single genetic event. Previous work with these lineages showed that HNRTs were common (Gaeta et al., 2007), thus we speculate that the loss of two or more linked markers is due to HNRTs despite the fact that this cannot be confirmed without dosage sensitive markers. Ten of 23 markers could not be mapped to either the *B. rapa* or *B. oleracea* subgenome and one marker was not mapped to either subgenome. Without map positions for these markers, it is impossible to determine whether loss of these markers correspond to loss of a larger linkage group. Three individuals in the generation of S₂ lineage 5 lost all loci analyzed on A7 of *B. rapa*, indicating replacement of one *B. rapa* chromosome by a homoeologous *B. oleracea* chromosome (i.e. The parents were CCCA for this chromosome) (see discussion and supplemental figure 2.2A). In addition, six individuals in the same generation lost one or more loci on this chromosome. We speculate that the S₁ parent of this generation had a chromosome replacement from *B. oleracea*, resulting in gametes (S₂ progeny) that did not contain A7 of *B. rapa* (Fig. 2.3).

Figure 2.3. Possible progeny of the S_2 at the A7 chromosome. The S_1 might have undergone a chromosome replacement, causing segregation of gametes to produce various combinations of chromosomes in the S_2 progeny.

In five cases, the same marker was absent in five or more siblings (see Supplemental Figure 2A-F). This suggests that the parent of that generation had a fragment deletion affecting one of the two homoeologs of a subgenome. It is expected that 25% of the progeny would be homozygous for this deletion, resulting in a 3:1 ratio ($\chi^2 = < 2.70$; $\alpha = 0.05$, $P = 0.322$). All five cases were consistent with the 3:1 ratio (Table 2.1).

Table 2.1. Chi-square goodness-of-fit test. The expected 3:1 ratio was observed in X cases in which five or more individuals in a single generation lost a particular marker.

Lineage	Generation	Marker Locus	Marker present: Marker absent	χ^2 value
5	S_2	BRMS 008 (C-subgenome)	29:5	< 2.70 ($P = .165$)
5	S_{12}	fito 095 (C-subgenome)	22:9	< 2.70 ($P = .604$)
78	S_2	fito 095 (C-	29:5	< 2.70 ($P =$

		subgenome)		.165)
78	S ₇	BRMS 034 (C-subgenome)	27:7	< 2.70 (P=.552)
78	S ₁₂	SN 9875 (A-subgenome)	29:5	< 2.70 (P=.165)

Sequence Loss Relative to the Parent

For individuals in all generations, percent sequence loss relative to the parent was averaged (Fig. 2.4). Changes relative to the parent were detected in all generations from both lineages. There were significant differences between the two lineages in S₂ and S₇, but no significant differences between S₁₂ generations (2-sample t-test, $\alpha=0.05$, $P=0.011$, $P=0.001$, $P=0.471$, respectively). The largest percent sequence loss was identified in S₂ of lineage 5 (ANOVA; Tukey 95.0% Simultaneous Confidence Intervals). However, much of the sequence loss in this generation was from A7 of *B. rapa* as described earlier. Four of six generations showed no significant differences of sequence loss relative to the parent in comparison of all generations and lineages. Significant changes across generations appeared in lineage 5 (ANOVA; $\alpha=0.05$, $P=0.000$), but no significant changes were observed throughout lineage 78 (ANOVA; $\alpha=0.05$, $P=0.872$). This shows that two lineages generated from the same diploid progenitors that are presumed to be identical at the time of hybridization, undergo genetic changes at different rates. The two lineages had significant differences between the early generations, but there were no significant differences between S₁₂ generations from both lineages, suggesting the *B. napus* genome becomes more stable as a lineage advances.

Figure 2.4. Percent sequence lost relative to the parent. Error bars represent the standard error of the mean.

Total Sequence Loss in Each Lineage

Based on the observation that each generation displayed novel sequence loss relative to the parent, we predict that the total sequence loss in a lineage should increase incrementally with each generation. To test this prediction, the total sequence loss for each individual was determined and the average sequence loss for each generation was calculated based on deletions from microsatellite markers (Fig. 2.5). Average sequence loss for any given generation reflects the genetic make-up of the parent plus any novel genetic changes appearing in the offspring. For example, the S_{11} parent of lineage 78 displayed loss of three *B. oleracea* subgenome sequences (6.8% of detectable sequences). All S_{12} offspring also lacked these sequences. However, 7 of 34 S_{12} siblings showed loss of additional markers that had been present in the parent. Therefore, the average sequence loss seen among S_{12} individuals ($7.92\% \pm 0.37$) reflects the sequence

loss accumulated during previous generations as well as the novel sequence loss generated during the formation of the S_{12} generation (i.e. during meiosis of the S_{11} parent). Between equivalent generations in the two lineages, S_2 , S_7 , and S_{12} had significant differences in total sequence loss (2-sample t-test; $\alpha = 0.05$ $P = 0.011$, $P = 0.002$, $P = 0.000$, respectively) (Fig. 2.5). The two lineages diverged genetically following hybridization and continue to undergo further changes through the S_{12} generations.

Figure 2.5. Total percent sequence loss. These are from individuals in lineage 78 and lineage 5 generations compared to S_0 . Error bars represent the standard error of the mean.

Loss of B. oleracea or B. rapa Subgenome

The sequences lost in each lineage were not distributed equally between the subgenomes. The genetic sequence lost across generations S_7 and S_{12} from both lineages showed more frequent loss of the *B. oleracea* subgenome (2-sample t-test; $\alpha=0.05$, $P=0.000$ for S_7 and S_{12} in both lineages) (Figs. 2.6 & 2.7). The S_2 generation of both lineages displayed no significant differences in sequence loss between the *B. rapa* subgenome and the *B. oleracea* subgenome (2-sample t-test; $\alpha=0.05$, $P>0.05$).

Figure 2.6. Percent sequence loss relative to diploid progenitors in lineage 5. Error bars represent the standard error of the mean.

Figure 2.7. Percent sequence loss relative to diploid progenitors in lineage 78. Error bars represent the standard error of the mean.

*Phenotypic Variation Across Selfed Lineages of *B. napus**

Previously, phenotypic variation was observed to be greater in S_5 than S_0 (Gaeta et al., 2007). To determine whether genetic changes between individuals in a generation contributed to phenotype and whether variation within a lineage increases as sequence loss increases, 12 phenotypes were measured. The phenotypes included the number of leaves at flowering, plant height to first flower, raceme height, total plant height, number of secondary branches, seed mass, number of siliques on the primary stem, and number of siliques with more than one seed. The third true leaf was used to measure leaf length, leaf width, petiole length, perimeter, and surface. The ratio of perimeter to surface area served as a proxy for leaf serration. The number of DNA fragment losses did not have an effect on the overall phenotypic means (Wilks' Lambda; $F=0.948$; $P=0.63$, $\alpha=0.05$) nor did the losses correlate with any individual phenotype ($P > 0.05$). The averages of seven

of twelve phenotypes were significantly different between the two lineages (Wilks' Lambda; $F=40.374$; $P=0.000$, $\alpha=0.05$). However, no trend was observed in which one lineage or generation was consistently greater (i.e. for a given phenotype, lineage 5 had a greater mean in S_7 while lineage 78 had greater a mean in S_{12} (See Supplemental Figure 2.1A-C). The variance in five of twelve phenotypes was significantly different between equivalent generations in the two lineages (see Supplemental Table 2.1). Novel phenotypes were detected in individuals of several generations, which were not seen many of their siblings. For example, individuals in the S_7 of lineage 78 displayed novel variation in flower color and shape, and flowering time (Fig. 2.8B & C). For yellow flower color, there was no significant difference from a 3:1 ratio ($\chi^2 = < 2.70$; $\alpha=0.05$, $P=0.322$). There was also significant variation in leaf morphology among all generations and lineages (Fig. 2.8D). When possible, the genetically inherited novel phenotypes were confirmed in selfed progeny (data not shown). Since not all plants were fertile, some phenotypes could not be confirmed genetically.

Figure 2.8. Novel phenotypes and phenotypic variation in resynthesized *B. napus*. (A) Typical *B. napus* napus flowers are white (left). A bright yellow-flowered plant (right) was identified in generation S_7 of lineage 78. (B) Flowers varied in size (left and right) and color (center, pale yellow). (C) Variation in flowering time represented by an early flowering plant (left) and a late flowering plant (right). Both plants shown are the same age and both are siblings from the S_7 generation of lineage 78. (D) Rosette-dwarf phenotype observed in a single plant from the S_7 generation of lineage 78. (E) Variation in leaf size and serration of the third true leaf. A representative leaf from the *B. rapa* progenitor (far left) and *B. oleracea* progenitor (far right) are shown.

Supplemental Table 2.1. F-tests and P-values. Phenotypic variance between two lineages of resynthesized *Brassica napus* allopolyploids at equivalent generations of S₂, S₇, and S₁₂.

Trait Measured	5-2 σ^2	78-2 σ^2	F	p	5-7 σ^2	78-7 σ^2	F	p	5-12 σ^2	78-12 σ^2	F	p
No. Leaves at Flowering	1.38	1.07	1.68	0.146	1.21	1.89	0.39	0.007	1.15	1.01	1.28	0.474
Plant Height to First Flower (cm)	5.97	3.47	2.97	0.003	3.96	5.88	0.46	0.027	4.96	4.77	1.08	0.820
Raceme Height (cm)	10.60	6.70	2.50	0.010	8.37	13.00	0.39	0.008	5.97	5.19	1.33	0.418
Total Plant Height (cm)	13.60	8.03	2.88	0.003	9.15	12.40	0.52	0.060	6.98	6.42	1.18	0.632
No. Secondary Branches (cm)	0.65	0.83	0.62	0.174	0.88	1.02	0.75	0.409	0.88	0.85	1.09	0.798
Seed Mass (g)	0.24	0.55	0.19	0.000	0.55	1.92	0.08	0.000	0.64	0.23	7.78	0.000
No. Siliques on Primary Stem	8.85	9.74	0.83	0.591	9.58	10.40	0.83	0.599	8.81	8.87	0.99	0.973
No. Siliques With More Than One Seed	15.60	22.10	0.50	0.053	17.30	19.80	0.74	0.398	22.20	25.40	0.76	0.448
Leaf 3 Length (cm)	1.74	1.77	0.97	0.925	2.21	1.56	1.87	0.072	2.16	0.85	6.50	0.000
Leaf 3 Width (cm)	1.10	0.98	1.26	0.519	1.05	0.75	1.99	0.051	0.84	0.85	0.99	0.974
Leaf 3 Petiole Length (cm)	1.10	1.36	0.65	0.222	0.38	1.04	1.70	0.126	1.27	0.67	3.54	0.000
Leaf 3 Perimeter/SA (cm/cm ²)	0.72	0.64	1.25	0.526	0.48	0.26	3.39	0.001	0.43	0.22	3.98	0.000

Supplemental Figure 2.1A. Variation in phenotype at the S₂ generation. Error bars present the standard error of the mean.

Supplemental Figure 2.1B. Variation in phenotype at the S_7 generation. Error bars present the standard error of the mean.

Supplemental Figure 2.1C. Variation in phenotype at the at the S_{12} generation. Error bars present the standard error of the mean.

Discussion

*Dynamics of Genetic Changes in Resynthesized *B. napus**

The allopolyploid genome is dynamic in that it undergoes a duplication event that lead to genomic changes, causing divergence between lineages. Since the *B. napus* (CCAA-genome) lineages analyzed here were derived from the same doubled haploid varieties of *B. oleracea* (C-subgenome) and *B. rapa* (A-subgenome), our null hypothesis was that both lineages would contain identical genomes. This null hypothesis was rejected; there were already significant genetic differences between the lineages by the S_2 generation in terms of percent sequence loss identified by 23 polymorphic markers. In fact, sequence changes were detected in all generations tested suggesting the genome

rearrangements that initiate following polyploidization continue for at least 12 generations. Results are consistent with evidence from previous studies with resynthesized *B. napus* allopolyploids. Previously, genetic changes were shown to be more abundant in the S_5 generation than in the S_0 generation (Gaeta et al., 2007) but evidence presented here along with work by others (Szadkowski et al., 2010) show that genetic changes are present by the S_1 generation. Though they are genetically identical at the time of formation, lineages of resynthesized *B. napus* begin to diverge immediately.

Sequence Loss Relative to the Parent

In all generations tested, individuals were identified with sequence loss not present in the parent. Genetic changes observed in a generation could be explained by novel mutation or by segregation of altered parental chromosomes. We detected 20 cases in which a single marker was lost in an individual that was present in the parent and all siblings. These could represent mutations that occurred in the parent; however, this is unlikely. For any individual to show loss of a signal from one subgenome requires that both gametes involved in forming that individual lack the same chromosomal segment. It is unlikely that two identical independent mutations would occur at the high frequency with which we observed these events (~10% of individuals). A more likely explanation is that the parent contained reciprocal exchanges (REs) between the subgenomes at the analyzed locus. In this scenario, individuals missing one subgenome would be expected to represent 1/16 of the progeny (Gaeta and Pires, 2010).

Six cases were identified in which a loss of the same marker sequence was common to multiple siblings in a generation. These changes are likely to be the result of segregation of parental chromosomes with HNRTs (Fig. 2.9; see Supplemental Figure 2.2

A-F). Such changes are expected to segregate 3:1, with $\frac{1}{4}$ of individuals lacking one subgenome sequence. All six cases show no significant difference from a classical Mendelian 3:1 ratio.

Figure 2.9. 3:1 ratio of segregation patterns. Segmental replacement due to HNRTs in the parental chromosome lead to segregation patterns of 3:1 in the progeny, with $\frac{1}{4}$ of individuals partially lacking one subgenome sequence.

Three individuals in the S_2 generation of lineage 5 lost all markers tested on A7 (*B. rapa* subgenome) (see Supplemental Figure 2A). A possible explanation is that the S_1 parent had a chromosome replacement (CCCA) or deletion (CCA). Many gametes from such an S_1 parent would be expected to lack an A7 chromosome resulting in some progeny lacking A7 (Fig. 2.9). During meiosis, bivalents form when gametes contain one chromosome from each subgenome (AC). However, chromosomes that contain HNRTs

and aneuploid gametes cause uneven proportions of chromosomes from each subgenome (Fig. 2.10). During meiosis in a CCCA or CCA individual, the A chromosome is more likely to pair with a C homoeolog than when homologous A chromosomes are present. This A-C pairing increases the likelihood of homoeologous exchange and HNRTs. In the S_2 generation of lineage 5, we observe six individuals missing one or more A7 markers (the individuals missing all A7 markers are not counted in this group). Taken together, this suggests that a chromosome dosage change in the parent led to a high frequency of progeny with complete and partial aneuploid of the A7 chromosome. It is important to note that the high percent sequence loss relative to the parent observed in the S_2 generation of lineage 5 (the largest observed in any generation tested; approximately a three-fold increase compared to other generations) is a reflection of mostly A7 segregation and aneuploidy.

Figure 2.10. Possible chromosome patterns found in allotetraploids. Under normal meiosis with bivalents, the meiotic products contain one chromosome from each genome. However, complex configurations such as multivalents formed in meiosis result in aneuploid cells or segmental chromosomes due to HNRTs. These possible abnormal gametes contribute to a positive feedback loop, resulting in an irreversible accumulation of mutations.

Supplemental Figure 2.1. *Brassica napus* chromosome map with 23 microsatellite markers used to map genetic changes and sequence loss. A1-A10 represents the *B. rapa* genome, while C1-C9 represents the *B. oleracea* genome of lineage 5 S_2 . Note: not all markers were able to located on both sets of the parental diploid chromosomes. The map distances of the markers are not to an exact scale, rather they are approximations.

Supplemental Figure 2.1A. *Brassica napus* chromosome map with 23 microsatellite markers used to map genetic changes and sequence loss. A1-A10 represents the *B. rapa* genome, while C1-C9 represents the *B. oleracea* genome of lineage 5 S₂. Colored circles represent the number of individuals in which the indicated sequence was lost, boxed markers represent those that were lost in the parent. Note: not all markers were able to located on both sets of the parental diploid chromosomes. The map distances of the markers are not to an exact scale, rather they are approximations.

Supplemental Figure 2.1B. *Brassica napus* chromosome map with 23 microsatellite markers used to map genetic changes and sequence loss. A1-A10 represents the *B. rapa* genome, while C1-C9 represents the *B. oleracea* genome of lineage 5 S₇. Colored circles represent the number of individuals in which the indicated sequence was lost, boxed markers represent those that were lost in the parent. Note: not all markers were able to located on both sets of the parental diploid chromosomes. The map distances of the markers are not to an exact scale, rather they are approximations.

Supplemental Figure 2.1C. *Brassica napus* chromosome map with 23 microsatellite markers used to map genetic changes and sequence loss. A1-A10 represents the *B. rapa* genome, while C1-C9 represents the *B. oleracea* genome of lineage 5 S_{12} . Colored circles represent the number of individuals in which the indicated sequence was lost, boxed markers represent those that were lost in the parent. Note: not all markers were able to located on both sets of the parental diploid chromosomes. The map distances of the markers are not to an exact scale, rather they are approximations.

Supplemental Figure 2.1D. *Brassica napus* chromosome map with 23 microsatellite markers used to map genetic changes and sequence loss. A1-A10 represents the *B. rapa* genome, while C1-C9 represents the *B. oleracea* genome of lineage 78 S₂. Colored circles represent the number of individuals in which the indicated sequence was lost, boxed markers represent those that were lost in the parent. Note: not all markers were able to located on both sets of the parental diploid chromosomes. The map distances of the markers are not to an exact scale, rather they are approximations.

Supplemental Figure 2.1E. *Brassica napus* chromosome map with 23 microsatellite markers used to map genetic changes and sequence loss. A1-A10 represents the *B. rapa* genome, while C1-C9 represents the *B. oleracea* genome of lineage 78 S₇. Colored circles represent the number of individuals in which the indicated sequence was lost, boxed markers represent those that were lost in the parent. Note: not all markers were able to located on both sets of the parental diploid chromosomes. The map distances of the markers are not to an exact scale, rather they are approximations.

Supplemental Figure 2.1F. *Brassica napus* chromosome map with 23 microsatellite markers used to map genetic changes and sequence loss. A1-A10 represents the *B. rapa* genome, while C1-C9 represents the *B. oleracea* genome of lineage 78 S₁₂. Colored circles represent the number of individuals in which the indicated sequence was lost, boxed markers represent those that were lost in the parent. Note: not all markers were able to located on both sets of the parental diploid chromosomes. The map distances of the markers are not to an exact scale, rather they are approximations.

Chromosomal changes resulting from RE and HNRTs are common in resynthesized *B. napus* (Osborn et al., 2003; Pires et al., 2004; Udall et al., 2005; Gaeta et al., 2007). Here, by examining progeny of a single allopolyploid, parent we observe the effects of segregation of chromosomes altered by RE and HNRTs. However, our methods can only detect loss of a subgenome at a particular locus, and not changes in dosage at marker loci. For example, chromosome substitutions have been observed such that the *B. napus* genome becomes CAAA, CCCA, AAAA, CCCC, CAA, or CCA

(aneuploid genotypes). Instead, dosage insensitive markers used in this study only detected deletions of the *B. oleracea* and *B. rapa* subgenomes as complete losses of AA or CC. Therefore, these observations may underestimate the total genetic change present in the parents and progeny.

The percent sequence loss increased at different rates in the two lineages. There were significant differences between the lineages at the S_2 and S_7 generations. No significant differences between the S_{12} generations were detected. This may indicate that the rate of changes from one generation to the next may start to stabilize sometime between the S_7 and S_{12} generation. The lineages did not undergo changes at the same rate, though they were genetically identical at the time of hybridization.

Directional Change in Resynthesized B. napus

As mentioned above, genetic changes were detected in every generation examined. Therefore it is not surprising that a significant increase in total genetic change relative to the S_0 was observed for both lineages, where S_{12} had the largest percent of loss. This suggests irreversible accumulation of genomic reorganizations and modifications that some have described as rapid genome evolution in newly formed allopolyploids (Ozkan et al., 2001; Pontes et al., 2004; Lim et al., 2008).

There was a bias towards losing the *B. oleracea* subgenome rather than the *B. rapa* subgenome across both lineages and all generations. Gaeta et al. (2007) also found that genetic changes in the S_5 occurred marginally more frequently in the *B. oleracea*. Specifically, at chromosomes A1/C1, there was a significant bias toward C1 over A1 fragment losses, but neither A2/C2 nor A3/C3 had any differences in loss between the subgenomes. Bias towards the loss of one subgenome has also been identified in studies

of the allopolyploid *Brassica juncea*. This allopolyploid showed significant directional loss of the *Brassica nigra* subgenome relative to the *B. rapa* subgenome. Such bias would make *B. juncea* more similar to *B. rapa* than to *B. nigra*. Genomic variation by allopolyploidization of the wheat-rye hybrid triticale was found to have greater sequence loss of the genome in rye (*Secale cereale* L.) than wheat (*Triticum* spp.) (Ma et al., 2004; Ma and Gustafson, 2006). The directional change towards one parental subgenome suggests the intriguing possibility that some condition predetermines which parental genome is lost at a greater rate. These studies suggest that changes following allopolyploidization are nonrandom and directional (Ozkan et al., 2001).

Not all studies of resynthesized allopolyploid have shown directional loss of one subgenome. Song et al. (1995) did not detect any directional changes in AACC and CCAA *B. napus* allopolyploids, and found that there are higher levels of cytoplasmic-nuclear genome compatibility in these allopolyploids. However, these studies only analyzed nine siblings in each of two generations. This small sample size might have skewed a normal distribution of a given generation (i.e. observations of non-directional changes could have been due to random chance in data collection). We analyzed only CCAA allopolyploids rather than reciprocal allopolyploids (AACC and CCAA) as in Song et al. (1995), but we included over 30 individuals in each generation. This gives more even distribution to each generation, allowing more significant comparisons to be made. More importantly, we observed significant directional change in one subgenome, highlighting the importance of analyzing many allopolyploids to accurately assess the dynamics of genetic changes.

Since the lineages showed directional loss of the *B. oleracea* subgenome, it might be expected that plants resemble the *B. rapa* subgenome progenitor phenotypically. In

one case, flower color showed directional change in several individuals. *Brassica rapa* have yellow flowers, whereas *B. oleracea* flowers are white. *Brassica napus* flowers are white indicating that flower color alleles derived from *B. oleracea* subgenome are dominant. Loss of the *B. oleracea* sequence at the flower-color locus would cause a shift toward the *B. rapa* phenotype (yellow). We found 11 individuals in S₇ of lineage 78 that displayed yellow flowers. The white and yellow flower ratio in this generation displayed no significant difference between a 3:1 ratio. This ratio suggests that the S₆ parent was heterozygous for the *B. oleracea* allele for flower color leading to segregation of yellow and white flowers in the progeny.

Reappearance of Lost Parental Bands

Two individuals (one in each of the S₁₂ generations) appear to have regained a subgenome signal that was absent in the S₁₁ parent and in all siblings. There is no clear mechanism to explain the reappearance in the progeny of sequences absent in the parents. Pollen contamination can be ruled out as an explanation for the reappearance of bands because the individuals that showed reappearance at one marker did not show reappearance at other markers absent in the parent. However, others have observed this same phenomenon in *B. napus* and other species. It has been suggested that a small frequency of aberrant meioses resulting in broken chromosome fragments could result in a gain of previous fragments due to recombination and/or gene conversion (Song et al., 1995). AFLP and RFLP analysis on triticale showed novel bands that were not present in either of its parents, wheat (*Triticum*) and rye (*Secale*) (Ma et al., 2004).

Phenotypes and Sequence Loss

Both lineages tested showed increase of genetic changes as the lineage advanced. Thus, it was expected that phenotypic variation would increase correspondingly in later generations. However, none of the phenotypes showed a directional trend in terms of variation when lineages or generations were compared. We did not find that total sequence loss had an overall effect on the means of overall phenotypes. In addition, genetic changes had no effect on any individual phenotypes across all generations. The microsatellite markers we used might not have corresponded to any loci that influenced the phenotypes that were analyzed.

An allopolyploid such as wheat can lose up to 15% of its genome (Shaked et al., 2001). However, sequence elimination of allopolyploids in the wheat group was found mostly in non-coding regions of the genome. Thus, most genetic changes in wheat are not expected to influence the phenotype. Since phenotypic variation did not correlate with genetics changes in *B. napus* in this study, it is possible that many of the genetic changes reported here map to non-coding regions.

Directions in Future Research

Many of the phenotypes investigated here are influenced by known genes. Future attempts to identify correlations between genetic changes and phenotype may focus on specific candidate genes. One example is the *FLC3* gene known to have an effect on flowering time (Pires et al., 2004). Currently, the S₂, S₇ and S₁₂ generations are being screened for genetics changes at the *FLC3* loci of the *B. rapa* and *B. oleracea* subgenomes. It will be interesting to determine whether variation in flowering time in any of the generations correlates with *FLC3* genotype. When the full sequenced genome

of *B. napus* is published, it will be possible to locate genes and genetic distances between markers so that the effect of genes on phenotypes may become more strongly correlated.

In our microsatellite analysis, we observed rare genetic changes from the S_6 to the S_7 progeny. The selfed seeds of S_7 in lineage 5 could be grown to observe whether subsequent generations show changes or if this particular lineage has a stable genome. If the S_6 generation of lineage 5 was truly stable, we would expect the S_7 plants of this lineage to inherit this stability. Few or no genetic changes would be expected in the S_8 generation if this were the case. The same set of microsatellite markers could be used to for genetic analysis and phenotype screening of S_8 to confirm whether it is significantly different from S_7 .

Using dosage sensitive markers such as RFLPs, loss or gain of one single chromosome generated by HNRTs can be identified. In addition, fluorescence *in situ* hybridization (FISH) can enable visualization of loss or addition of one copy of the chromosome by detecting the exact number of chromosomes missing within the allopolyploid genome. These methods will allow greater accuracy of detecting genetic changes and linked deletion events.

In addition to detecting sequence loss, epigenetic change is another phenomenon to observe in the 12 generations of *B. napus*. The dynamics of epigenetic changes in the *B. napus* genome has not been followed throughout several generations past the S_5 generation. Previous analysis of these *B. napus* lineages at the S_5 generation revealed little change in CpG methylation (Gaeta et al., 2007). Using the material developed for this project, it would be possible to analyze DNA methylation in the S_2 to determine if epigenetic changes are exclusive to the S_0 or whether these changes continue for several generations. It would also be interesting to determine whether this reduction in

epigenetic change persists into the S₁₂ generation or beyond. A PCR-based approach to distinguishing methylated and unmethylated sequences has been developed and this could be used to screen the generations developed for this study.

CHAPTER 3: PHENOTYPE DATA

Phenotype Analysis

Previously, 50 independent lineages of *B. napus* had been generated using well-characterized diploid parents *B. oleracea* (TO1000; A-genome) and *B. rapa* (IMB218; C-genome) (Lukens et al., 2006; Gaeta et al., 2007). Two of these lineages (lineage 5 and

lineage 78) were propagated to the twelfth generation and twelve phenotypes were measured. These phenotypes include number of leaves at flowering, height of first flower (cm), raceme height (cm), total plant height (cm), number of secondary branches, seed mass (g), number of siliques on the primary stem, and number of siliques with more than one seed. The third true leaf was used to measure leaf length (cm), leaf width (cm), and petiole length (cm). The ratio of leaf perimeter (cm) per surface area (cm^2) was determined as a proxy for leaf serration. Here we present an index of statistical analysis for each phenotype. An analysis of variance (ANOVA) was performed to test for significant differences across generations in a lineage (i.e. comparison of lineage 5, generation S2 to lineage 5, generation S7, etc.). A 2-sample t-test was performed to test for significant differences between lineages at equivalent generations (i.e. comparison of S2 of lineages 5 and 78). F-tests were used as a measurement for variance among each phenotype in comparison of equivalent generations between lineages (Table 3.1). Bartlett's test was used to compare variances among generations in each lineage (Table 3.2). The F-test compares variances between two variables, while Bartlett's test compares multiple variables simultaneously.

No cases were observed in which sequence loss correlated with selected phenotypes. However, genetic analyses used polymorphic markers rather than coding sequences. Further genetic analyses could be done to determine the effect of specific loci on phenotypic variation. As genomic resources are developed for *B. napus*, it will be possible to identify candidate genes encoding functions that are likely to influence some of the phenotypes analyzed for this project. For example, the location of the *FLC3* gene has been genetically mapped in the *B. napus* genome, and can be amplified for polymorphism identification to observe correlations with the number of leaves at

flowering (Pires et al., 2004) Studies in *Arabidopsis thaliana* have revealed many genes involved in flowering time, leaf serration, fertility, and stature, which can all be candidates corresponding to genetic changes (The *Arabidopsis* Initiative, 2000).

Table 3.1. F-tests and *P*-values for phenotypic variance. Variance between two lineages of resynthesized *Brassica napus* allopolyploids at generations S_2 , S_7 , and S_{12} .

Trait Measured	5-2 σ^2	78-2 σ^2	F	p	5-7 σ^2	78-7 σ^2	F	p	5-12 σ^2	78-12 σ^2	F	p
No. Leaves at Flowering	1.38	1.07	1.68	0.146	1.21	1.89	0.39	0.007	1.15	1.01	1.28	0.474
Plant Height to First Flower (cm)	5.97	3.47	2.97	0.003	3.96	5.88	0.46	0.027	4.96	4.77	1.08	0.820
Raceme Height (cm)	10.60	6.70	2.50	0.010	8.37	13.00	0.39	0.008	5.97	5.19	1.33	0.418
Total Plant Height (cm)	13.60	8.03	2.88	0.003	9.15	12.40	0.52	0.060	6.98	6.42	1.18	0.632
No. Secondary Branches (cm)	0.65	0.83	0.62	0.174	0.88	1.02	0.75	0.409	0.88	0.85	1.09	0.798
Seed Mass (g)	0.24	0.55	0.19	0.000	0.55	1.92	0.08	0.000	0.64	0.23	7.78	0.000
No. Siliques on Primary Stem	8.85	9.74	0.83	0.591	9.58	10.40	0.83	0.599	8.81	8.87	0.99	0.973
No. Siliques With More Than One Seed	15.60	22.10	0.50	0.053	17.30	19.80	0.74	0.398	22.20	25.40	0.76	0.448
Leaf 3 Length (cm)	1.74	1.77	0.97	0.925	2.21	1.56	1.87	0.072	2.16	0.85	6.50	0.000
Leaf 3 Width (cm)	1.10	0.98	1.26	0.519	1.05	0.75	1.99	0.051	0.84	0.85	0.99	0.974
Leaf 3 Petiole Length (cm)	1.10	1.36	0.65	0.222	0.38	1.04	1.70	0.126	1.27	0.67	3.54	0.000
Leaf 3 Perimeter/SA (cm/cm ²)	0.72	0.64	1.25	0.526	0.48	0.26	3.39	0.001	0.43	0.22	3.98	0.000

Table 3.2. Bartlett's test and *P*-values for phenotypic variance. Variation across two lineages of resynthesized *Brassica napus* allopolyploids at generations S₁, S₂, and S₃.

Trait Measured	5-2 σ^2	5-7 σ^2	5-12 σ^2	Bartlett's test	p	78-2 σ^2	78-7 σ^2	78-12 σ^2	Bartlett's test	p
No. Leaves at Flowering	1.38	1.21	1.15	1.17	0.557	1.07	1.89	1.01	17.32	0.000
Plant Height to First Flower (cm)	5.97	3.96	4.96	5.39	0.068	3.47	5.88	4.77	8.78	0.012
Raceme Height (cm)	10.60	8.37	5.97	2.50	17.620	6.70	13.00	5.19	30.91	0.000
Total Plant Height (cm)	13.60	9.15	6.98	14.37	0.001	8.03	12.40	6.42	15.25	0.000
No. Secondary Branches (cm)	0.65	0.88	0.88	3.70	0.157	0.83	1.02	0.85	3.70	0.157
Seed Mass (g)	0.24	0.55	0.64	24.46	0.000	0.55	1.92	0.23	26.46	0.000
No. Siliques on Primary Stem	8.85	9.58	8.81	0.29	0.866	9.74	10.40	8.87	0.29	0.866
No. Siliques With More Than One Seed	15.60	17.30	22.20	4.14	0.126	22.10	19.80	25.40	4.14	0.126
Leaf 3 Length (cm)	1.74	2.21	2.16	2.11	0.349	1.77	1.56	0.85	2.11	0.349
Leaf 3 Width (cm)	0.66	1.05	0.84	2.54	0.281	0.98	0.75	0.85	2.54	0.281
Leaf 3 Petiole Length (cm)	1.10	0.38	1.27	1.73	0.422	1.36	1.04	0.67	1.70	0.422
Leaf 3 Perimeter/SA (cm/cm ²)	0.72	0.48	0.43	9.95	0.007	0.64	0.26	0.22	40.72	0.000

1. Number of leaves at flowering

In lineage 5, S_7 was larger than S_2 , S_{12} was larger than S_2 and there were no significant differences between S_{12} and S_7 (Fig. 3.1A&B). In lineage 78, S_7 was larger than S_2 , and S_{12} , and there were no significant differences between S_2 and S_{12} . Across equivalent generations between two lineages, lineage 78 was larger in S_2 , S_7 , and there were no significant differences between the generations of the two lineages in S_{12} .

Figure 3.1A. Data analysis of number of leaves at flowering. Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.1B. Data analysis of number of leaves at flowering. Interval plot of the number of leaves at flowering. Error bars represent the standard deviation of the mean.

2. *Plant height to first flower (cm)*

In lineage 5, S_7 was larger than S_2 and S_{12} , but S_2 was larger than S_{12} (Fig.3.2A&B). In lineage 78, S_7 was larger than S_2 , S_{12} was larger than S_7 , and S_2 was larger than S_{12} . Across equivalent generations between two lineages, lineage 5 was larger in S_2 , lineage 5 was larger in S_7 , and lineage 78 was larger in S_{12} .

Figure 3.2A. Data analysis of plant height to first flower (cm). Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.2B. Data analysis of plant height to first flower (cm). Interval plot of plant height to first flower. Error bars represent the standard deviation of the mean.

3. Raceme height (cm)

In lineage 5, S_7 was larger than S_2 , S_{12} was larger than S_2 , and there were no significant difference between S_7 and S_{12} (Fig. 3.3A&B). In lineage 78, S_7 was larger than S_{12} , and there were no significant differences between other generations. Across equivalent generations between two lineages, lineage 78 was larger in S_2 and S_7 , and lineage 5 was larger in S_{12} .

Figure 3.3A. Data analysis of raceme height (cm). Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.3B. Data analysis of raceme height (cm). Interval plot of raceme height. Error bars represent the standard deviation of the mean.

4. Total plant height (cm)

In lineage 5, S_7 was larger than S_2 , and there were no significant differences between S_7 and S_{12} , or S_2 and S_{12} (Fig. 3.4A&B). In lineage 78, there were no significant differences throughout generations. Across equivalent generations between two lineages, there were no significant differences between equivalent generations.

Figure 3.4A. Data analysis of total plant height (cm). Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.4B. Data analysis of total plant height (cm). Interval plot of total plant height. Error bars represent the standard deviation of the mean.

5. *Number of secondary branches*

In lineage 5, there were no significant difference between S_2 , S_7 and S_{12} (Fig.3.5A&B). In lineage 78, S_7 was larger than S_2 , S_{12} was larger than S_7 , and there were no significant differences between S_{12} and S_2 . Across equivalent generations between two lineages S_2 : lineage 78 was larger in S_2 , lineage 78 was larger in S_7 and no significant differences between S_{12} .

Figure 3.5A. Data analysis of number of secondary branches. Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.5B. Data analysis of number of secondary branches. Interval plot of secondary branches. Error bars represent the standard deviation of the mean.

6. *Seed mass (g)*

In lineage 5, S_7 was larger than S_2 , S_{12} was larger than S_2 and there were no significant differences between S_7 and S_{12} (Fig. 3.6A&B). In lineage 78, S_7 was larger than S_2 , S_{12} was larger than S_7 , and there were no significant differences between S_{12} and S_2 . Across equivalent generations between two lineages, lineage 78 was larger in S_2 , lineage 5 was larger in S_7 , and lineage 78 was larger in S_{12} .

Figure 3.6A. Data analysis of seed mass (g). Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.6B. Data analysis of seed mass (g). Interval plot of seed mass (g). Error bars represent the standard deviation of the mean.

7. *Number of siliques on the primary stem*

Lineage 5, S_7 was larger than S_2 , S_{12} was larger than S_2 , and there were no significant differences between S_{12} and S_7 (Fig. 3.7A&B). In lineage 78, S_7 was larger than S_2 , S_{12} was larger than S_7 , and there were no significant differences between S_{12} and S_2 . Across equivalent generations between two lineages, lineage 78 was larger in S_2 , lineage 5 was larger in S_7 , and no significant differences between equivalent generations in S_{12} .

Figure 3.7A. Data analysis of number of siliques on primary stem. Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.7B. Data analysis of number of siliques on primary stem. Interval plot of number of siliques on primary stem. Error bars represent the standard deviation of the mean.

8. *Number of siliques with more than one seed*

In lineage 5, S_{12} was larger than S_2 , and there were no other significant differences across other generations (Fig. 3.8A&B). In lineage 78, S_{12} was larger than S_2 , but no significant differences across any other generations. Across equivalent generations between two lineages, lineage 78 was larger in S_2 , lineage 5 was larger in S_7 , and no significant differences between equivalent generations in S_{12} .

Figure 3.8A. Data analysis of number of siliques with more than one seed. Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.8B. Data analysis of number of siliques with more than one seed.
Interval plot of the number of siliques with more than one seed. Error bars represent the standard deviation of the mean.

9. *Leaf length (cm)*

In lineage 5, S_7 was larger than S_2 , S_{12} was larger than S_2 , and no significant differences between S_7 and S_{12} (Fig. 3.9A&B). In lineage 78, there were no significant differences across any other generations. Across equivalent generations between two lineages, lineage 78 was larger in S_2 , lineage 5 was larger in S_7 , and Lineage 5 was larger in S_{12} .

Figure 3.9A. Data analysis of leaf length (L3). Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.9B. Data analysis of leaf length (L3). Interval plot of leaf length (L3). Error bars represent the standard deviation of the mean.

10. *Leaf width (cm)*

In lineage 5, S_2 was larger than S_7 , while there were no significant differences between S_7 and S_{12} (Fig. 3.10A&B). In lineage 78, there were no significant differences throughout the lineage. Across equivalent generations between two lineages, S_2 was larger in lineage 78, S_7 was larger in lineage 5, and S_{12} was larger in lineage 5.

Figure 3.10A. Data analysis of leaf width (L3). Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.10B. Data analysis of leaf width (L3). Interval plot of leaf width (L3). Error bars represent the standard deviation of the mean.

11. *Petiole length (cm)*

In lineage 5, S_7 was larger than S_2 , S_{12} was larger than S_2 , and there were no significant differences between S_7 and S_{12} (Fig. 3.11A&B). In lineage 78, there were no significant differences between any other generations. Across equivalent generations between two lineages, there were no significant differences between equivalent generations in S_2 , lineage 5 was larger in S_7 , and lineage 5 was larger in S_{12} .

Figure 3.11A. Data analysis of petiole length (L3). Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.11B. Data analysis of petiole length (L3). Interval plot of petiole length (L3). Error bars represent the standard deviation of the mean.

12. *Perimeter (cm) per/surface area (cm²)*

In lineage 5, there were no significant differences between any generations (Fig. 3.12A&B). In lineage 78, S_{12} was larger than S_2 , but there were no significant differences between any other generations. Across equivalent generations between two lineages, lineage 5 was larger in S_2 , lineage 5 was larger in S_7 , and no significant differences between equivalent generations in S_{12} .

Figure 3.12A. Data analysis of perimeter (cm) per surface area (cm^2). Arrows point to the generation in which the mean was significantly larger. Lines with no arrowed ends represent no significant differences between generations.

Figure 3.12B. Data analysis of perimeter (cm) per surface area (cm²). Interval plot of perimeter (cm) per surface area (cm²). Error bars represent the standard deviation of the mean.

BIBLIOGRAPHY

- Adams KL (2007) Evolution of duplicate gene expression in polyploid and hybrid plants. J Hered 98 (2): 136-141**
- Adams KL , Wendel JF (2005) Polyploidy and genome evolution in plants. Curr Opin Plant Biol 8: 135-141**
- Bailey JA, Baertsh R, Kent WJ, Haussler D, Eichler EE (2004) Hotspots of mammalian chromosomal evolution. Genome Biol 5: R23**
- Blanc G, Wolfe KH (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. Plant Cell 16: 1679-1691**
- Chapman MA, Burke JM (2007) Genetic divergence and hybrid speciation. Evolution 61: 1773-1780**
- Chen ZJ (2007) Genetic and Epigenetic Mechanisms for Gene Expression and Phenotypic Variation in Plant Polyploids. Annu Rev Plant Biol 58: 377-406**
- Comai L (2000) Genetic and epigenetic interactions in allopolyploid plants. Plant Molec Biol 43: 387-399**
- Comai L (2005) The advantages and disadvantages of being polyploid. Nature 6: 836-846**
- Comai L, Tyagi. AP, Winter. K, Holmes-Davis. R, Reynolds. SH, Stevens. Y, Byers B (2000) Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. Plant Cell 12: 1551-1567**
- Dufresne F, Hebert PDN (1994) Hybridization and origins of polyploidy. Proc Biol Sci 258: 141-146**
- Fawcett JA, Maere S, Van de Peer Y (2009) Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. Proc Natl Acad Sci USA 106: 5737-5742**
- Gaeta RT, Pires JC (2010) Homoeologous recombination in allopolyploids: the polyploid ratchet. New Phyt 186: 18-28**
- Gaeta RT, Pires JC, Iniguez-Luy FL, Leon E, Osborn TC (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell 19: 3403-3417**
- Ha M, Kim ED, Chen JZ (2009) Duplicate genes increase expression diversity in related species and allopolyploids. Proc Natl Acad Sci USA 106: 2295-2300**
- Haigh J (1978) The accumulation of deleterious genes in a population- Muller's ratchet. Theor Popul Biol 14: 251-267**

- Ilic K, SanMiguel PJ, Bennetzen JL (2003) A complex history of rearrangement in an orthologous region of the maize, sorghum and rice genomes. *Proc Natl Acad Sci USA* 100: 12265-12270
- Iniguez-Luy FL, Lukens LN, Farnham MW, Amasino RM, Osborn TC (2009) Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling *Brassica rapa* and *B. oleracea*. *Theor Appl Genet* 120: 31-43
- Iniguez-Luy FL, Voort AV, Osborn TC (2008) Development of a set of public SSR markers derived from genomic sequence of a rapid cycling *Brassica oleracea* L. genotype. *Theor Appl Genet* 117: 997-985
- Leitch AR, Leitch IJ (2008) Genome plasticity and the diversity of polyploid plants. *Science* 320: 481-483
- Levin DA (1982) Polyploidy and novelty in flowering plants. *Amer Nat* 122: 1-25
- Lim KY, Soltis DE, Soltis PS, Tate JA, Matyasek R, Srubarova H, Kovarik A, Pires JC, Xiong Z, Leitch AR (2008) Rapid chromosome evolution in recently formed polyploid in *Tragopogon* (Asteraceae). *PLoS ONE* 3: 1-13
- Lukens LN, Pires JC, Leon E, Vogelzang R, Oslach L, Osborn TC (2006) Patterns of sequence loss and cytosine methylation within a population of newly resynthesized *Brassica napus* allopolyploids. *Plant Physiol* 140: 336-348
- Ma X, Fang P, Gustafson JP (2004) Polyploidization-induced genome variation in triticales. *Genome* 47: 839-848
- Ma X, Gustafson JP (2006) Timing and rate of genome variation in triticales following allopolyploidization. *Genome* 49: 950-958
- Madlung A, Masuelli RW, Watson B, Reynolds SH, Davison J, Comai L (2002) Remodeling of DNA methylation and phenotypic transcriptional changes in synthetic *Arabidopsis* allotetraploids. *Plant Physiol* 129: 733-746
- Mallet J (2007) Hybrid speciation. *Nature* 446: 279-283
- Masterson J (1994) Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* 264: 421-424
- McClintock B (1984) The significance of responses of the genome to challenge. *Science* 226: 792-801
- McFadden DE, Kwong LC, Yam IYL, Langlois S (1993) Parental origin of triploidy in human fetuses: evidence for genomic imprinting. *Hum Genet* 92: 465-469
- Meyers LA, Levin DA (2006) On the abundance of polyploids in flowering plants. *Evolution* 60: 1198-1206

- Muller HJ (1964) The relation of recombination to mutational advance. *Mutation Research* 1: 2-9
- Ni Z, Kim E, Ha M, Lackey E, Liu J, Zhang Y, Sun Q, Chen ZJ (2009) Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* 457: 327-333
- Nicholas SD, Mignon GL, Eber F, Coriton O, Monod H, Clouet V, Huteau V, Lostanlen A, Delourme R, Chalhoub B, Ryder CD, Chevre AM, Jenczewski E (2006) Homoeologous recombination plays a major role in chromosome rearrangements that occur during meiosis of *Brassica napus* haploid. *Genet* 175: 487-503
- Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, Lee H, Comai L, Madlung A, Doerge RW, Colot V, Martienssen. RA (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet* 19: 141-147
- Otto SP, Whitton J (2000) Polyploid incidence and evolution. *Annu Rev Genet* 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
- Parisod C, Alix K, Just Jeremy, Petit M, Sarilar V, Mhiri C, Ainouche ML, Chalhoub B, Grandbastien M (2010) Impact of transposable elements on the organization and function of allopolyploid genomes. *New Phyt* 186: 37-45
- Parkin IAP, Gulden SM, Sharpe AG, Lukens LN, Trick M, Osborn TC, Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* 171: 765-781
- Pires JC, Zhao J, Schranz ME, Leon EJ, Quijada A, Lukens LN, Osborn TC (2004) Flowering time divergence and genomic rearrangements in resynthesized *Brassica* polyploid (Brassicaceae). *Biol J Linnean Soc* 82: 675-688
- Pontes O, Neves N, Silva M, Lewis MS, Madlung A, Comai L, Viegas W, Pikaard CS, Wessler SR (2004) Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. *Proc Natl Acad Sci USA* 52: 18240-18245
- Ramsey J, Schemske DW (1998) Pathways, mechanisms and rates of polyploid formation in flowering plants. *Annu Rev Ecol Syst* 29: 467-501
- Rieseberg LH (1997) Hybrid origins of plant species. *Annu Rev Ecol Syst* 28: 359-389
- Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends Ecol Evolut* 7: 351-358
- Rieseberg LH, Willis JH (2007) Plant Speciation. *Science* 17: 910-914

- Salmon A, Flagel LE, Ying B, Udall JA, Wendel JF (2010) Homoeologous nonreciprocal recombination in polyploid cotton. *New Phyt* 186: 123-134
- Schranz ME, Osborn TC (2000) Novel flowering time variation in resynthesized polyploid *Brassica napus*. *J Hered* 91: 242-246
- 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
- Soltis DE, Soltis PS (1993) Molecular data and the dynamic nature of polyploidy. *Crit Rev Plant Sci* 12: 243-273
- Soltis DE, Soltis PS (1999) Polyploidy: recurrent formation and genome evolution. *Tree* 14: 348-352
- Song K, Lu P, Tang K, Osborn TC (1995) Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proc Natl Acad Sci USA* 92: 7719-7723
- Song K, Tang K, Osborn TC (1993) Development of synthetic *Brassica* amphidiploids by reciprocal hybridization and comparison to natural amphidiploids. *Theor Appl Genet* 86: 811-821
- Szadkowski E, Eber F, Huteau V, Lode M, Huneau C, Belcram H, Coriton O, Manzanares-Dauleux MJ, Delourme R, King GJ, Chalhoub B, Jenczewski E, Chevre AM (2010) The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phyt* 186: 102-112
- The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796-815
- Udall JA, Quijada A, Osborn TC (2005) Detection of chromosomal rearrangements derived from homeologous recombination in four mapping populations of *Brassica napus* L. . *Genetics* 169: 967-979
- Van de Peer Y (2009) The evolutionary significance of ancient genome duplications. *Nat Rev Genet* 10: 725-732
- Walling JG, Shoemaker R, Young N, Mudge J, Jackson S (2006) Chromosome-level homeology in paleopolyploid soybean (*Glycine max*) revealed through integration of genetic and chromosome maps. *Genet* 172: 1893-1900
- Wang J, Tian L, Lee HS, Chen ZJ (2006) Nonadditive regulation of *FRI* and *FLC* mediates flowering-time variation in *Arabidopsis* allopolyploids. *Genet* 173: 965-974

- Wendel JF, Schnabel A, Seelanan T (1995) Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc Natl Acad Sci USA 92: 280-284**
- Wolfe KH (2001) Yesterday's polyploids and the mystery of diploidization. Nature Rev Genet 2: 333-341**
- Yu ZY, Haage K, Streit VE, Gierl A, Torres Ruiz AT (2009) A large number of tetraploid *Arabidopsis thaliana* lines generated by a rapid strategy, reveal high stability of neo-tetraploid during consecutive generations. Theor Appl Genet 118: 1107-1119**