The effect of formation pathway on allopolyploids between *Brassica carinata, Brassica napus, Brassica juncea* and *Sinapis arvensis*

by
Cleoniki Elizabeth Kesidis

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Science

in

Biology

Carleton University Ottawa, Ontario

© 2018
Cleoniki Elizabeth Kesidis

Abstract

Polyploidization has played a key role in the evolution of plant lineages, but polyploid formation rates and the characteristics of early generation polyploids remain areas where experimental work is needed. The effect of formation pathway on neopolyploid morphology and fertility, especially, is unknown but essential for understanding polyploid formation and establishment. This thesis describes the formation rate of one-step and two-step allopolyploids with *Brassica carinata*, *Brassica juncea* or *Brassica napus* as the maternal parent and *Sinapis arvensis* as the paternal parent using hand and insect-mediated crosses. The fertility and morphology of one-step and two-step allopolyploids were measured for three generations and compared with each other as well as with homoploid hybrids and autopolyploids of the parental species. One-step allopolyploids formed twice as frequently, though once a homoploid hybrid formed two-step allopolyploids were ten times as frequent as one-step allopolyploids. Therefore, hybridization was identified as a limiting factor in allopolyploidization. The two-step allopolyploids were more fertile and more reproductively isolated from parental species. The two types of allopolyploids differed morphologically, but in both types, polyploidy had a large initial effect on morphology that decreased by the third generation while the effect of hybridity endured. Both types of allopolyploids, especially the one-step allopolyploids, showed an unexpectedly high rate of DNA downsizing, decreasing back to the DNA content of *B. carinata* in some cases. In conclusion, allopolyploids formed through different formation pathways differ significantly in fertility, morphology, and

genome stability. Formation pathway should be considered in future work on neoallopolyploids.

Acknowledgements

I thank Connie Sauder and Tracey James for their training and assistance, and Jarred Lau and Tim Garant for their help collecting data. I thank the greenhouse staff at the Ottawa Research and Development Centre, and Drs. Julian Starr, Root Gorelick, and Tyler Smith for discussions on this project. This work was funded by Agriculture and Agri-Food Canada as part of project# J-001029 “Gene flow, diversity and relationships within the Brassicaceae: Focus on Camelina” led by Dr. Sara Martin.

Dedication

For Louis-Philippe, my sunshine.

Table of Contents

Abstract ii

Acknowledgements iv

Dedication v

Table of Contents vi

List of Tables ix

List of Figures x

1. Introduction 1

What is polyploidy and why study it? 1

Polyploidy plays a significant but poorly-understood role in plant evolution and speciation. 1

Definitions of polyploids 5

How do polyploids form? 7

Unreduced gamete formation is affected by genetics and possibly by environment 7

Unreduced gametes are formed by meiotic errors 10

Polyploid formation rates and pathways are largely unknown 12

Formation pathway may affect probability of formation and establishment 15

What are the effects of polyploidy? 16

Hybridization causes more genetic and epigenetic changes than genome doubling 16

The evolutionary advantages of polyploidy 20

Experimental system 23
The triangle of U offers an ideal study system for polyploidy 23
Brassica carinata (Ethiopian mustard; BBCC, 2n=34) 25
Brassica juncea (Indian mustard; AABB, 2n=36) 26
Brassica napus (Canola; AACC, 2n=38) 26
Sinapis arvensis (Wild mustard; SrSr, 2n=18), a persistent weed 26
Research Objectives 27
2. Materials and Methods 29
Plant growth procedure 29
Hand pollinations 30
Insect-mediated pollinations 31
Flow cytometry 31
Autopolyploid induction 32
Characterization of fertility and morphology 33
Allopolyploid formation rate estimation 36
Statistical analysis 38
3. Results 39
One-step allopolyploid formation rates 39
Two-step allopolyploid formation rates 40
Unexpected DNA contents of allopolyploid offspring 40

Reproductive isolation of polyploids 46
Fertility in one-step and two-step allopolyploids over three generations 53
Correlation between fertility and DNA content 56
Morphological characteristics of different types of allopolyploids 58
The contribution of hybridity and polyploidy to allopolyploid morphology 61
Pollen size 73
4. Discussion and Conclusions 76
Formation rates of one-step vs. two-step allopolyploids 76
Genomic distance affects hybridization rate 79
Two-step allopolyploids had higher fertility but fertility declined with generations ... 81
Incomplete reproductive isolation between cytotypes 82
Formation route affects morphology 83
Effects of hybridity on morphology endure but effects of polyploidy fade 84
Dramatic DNA downsizing in allopolyploid offspring 86
Importance of rare individuals 89
Conclusions 90
Future directions 92
5. Supplementary Data 97
References 101

List of Tables

Table 1: Source of seeds. 29

Table 2: Data collected for morphological analysis. 34

Table 3: One-step and two-step allopolyploid formation rates. 40

Table 4: ANOVAs and chi-squared tests examining the effect of step and generation on morphological traits. Significant p-values are bolded. 59

Table 5: ANOVAs examining the effect of hybridity and polyploidy on morphological traits that were not affected by type or generation. Significant p-values are bolded. 65

Table 6: ANOVAs examining the effect of hybridity and polyploidy on morphological traits that were only affected by type. Significant p-values are bolded. 66

Table 7: ANOVAs examining the effect of hybridity and polyploidy on morphological traits that were only affected by generation. Significant p-values are bolded. 66

Table 8: ANOVAs examining the effect of hybridity and polyploidy on morphological traits that were affected by both type and generation. Significant p-values are bolded. .. 67

List of Figures

Figure 1: The three pathways of polyploid formation: one-step, triploid bridge, and two-step. Modified from Tayalé and Parisod (2013). 15

Figure 2: The triangle of U. *S. arvensis* is most similar to the B genome. Modified from U (1935). 25

Figure 3: Histograms of parentals, homoploid hybrids and allopolyploids of all three crosses. The expected values are shown in colored lines (maternal parent blue, paternal parent red, homoploid hybrid pink or purple, allopolyploid green). 42

Figure 4: The DNA contents of three generations of one-step allopolyploids (*B. carinata* x *S. arvensis*). Arrows connect maternal parents to their offspring. 43

Figure 5: The DNA contents of three generations of two-step allopolyploids (*B. carinata* x *S. arvensis*). Arrows connect maternal parents to their offspring. 44

Figure 6: The 2C DNA content of *B. carinata*-maternal one-step and two-step allopolyploids over three generations. A) Significance values are based on GLS. B) Vertical lines show expected values for *B. carinata* (blue), *S. arvensis* (red), homoploid hybrids (purple) and allopolyploids (green). 45

Figure 7: Seeds produced by hand crosses between allopolyploids and parentals. Significance letters are based on Kruskal-Wallis tests. BCM: backcross to *B. carinata* (~10 flowers per individual). BCP: backcross to *S. arvensis* (~10 flowers per individual). E: emasculated (~5 flowers per individual). E+S: Emasculated and selfed (~5 flowers per individual). S: Selfed (~5 flowers per individual). SIB: Crossed with polyploid sibling (~5 flowers per individual). SIBD: Crossed with homoploid sibling (~5 flowers per

individual). U: unmanipulated (~5 flowers per individual). (a) and (b) show data from two individuals each, (c) shows data from 13 individuals, and (d) shows data from one individual. 49

Figure 8: Seeds produced by hand crosses between autopolyploids and parentals. Significance letters are based on Kruskal-Wallis tests. BCM: backcross to maternal parent (10-20 flowers per individual). BCP: backcross to *B. carinata* (~20 flowers per individual). E: emasculated (~5 flowers per individual). E+S: Emasculated and selfed (~5 flowers per individual). S: Selfed (~5 flowers per individual). SIB1 or SIB2: Crossed with polyploid sibling (~5 flowers per individual). U: unmanipulated (~5 flowers per individual). (a) shows data from three individuals and (b) shows data from two. 50

Figure 9: Seeds produced by hand crosses between the two second-generation one-step allopolyploids with unexpectedly low 2C DNA contents (~2.3 pg) and their parentals. Significance letters are based on Kruskal-Wallis tests. BCM: backcross to *B. carinata* (~10 flowers per individual). BCP: backcross to *S. arvensis* (~10 flowers per individual). E: emasculated (~5 flowers per individual). E+S: Emasculated and selfed (~5 flowers per individual). SIB: Crossed with the other unexpectedly low second-generation one-step (~5 flowers per individual). U: unmanipulated (~5 flowers per individual). Data from two individuals are included here. 51

Figure 10: Anther orientation in *B. carinata* (a), *S. arvensis* (b), first-generation two-step allopolyploids (c, d) and second-generation two-step allopolyploids (e, f)..... 52

Figure 11: Histogram of all one-step and two-step *B. carinata*-maternal allopolyploids. The red lines indicate the boundaries of the strict definition of allopolyploidy. 53

Figure 12: Pollen viability and seed production for three generations of one-step and two-step allopolyploids, compared to parentals. Significance indications are based on a generalized least square (GLS) model with a factor allowing for each group to have a different variance. 55

Figure 13: Correlations between 2C DNA content and fertility markers. P values are Pearson's correlation coefficient..... 57

Figure 14: Differentiation between types and generations of *B. carinata*-maternal allopolyploids using linear discriminate analysis (LDA). 59

Figure 15: Differentiation of parentals, homoploids, allopolyploids, and autopolyploids using linear discriminate analysis (LDA). 64

Figure 16: Stomata length of allopolyploids, autopolyploids, homoploids, and parentals, showing the effect of polyploidy. Significance letters are based on GLS. 69

Figure 17: Trichome density and beak length of allopolyploids, autopolyploids, homoploids, and parentals, showing the effect of hybridity. Significance letters are based on GLS. 70

Figure 18: Petal width and petal width/height of allopolyploids, autopolyploids, homoploids, and parentals, showing the changing effect of polyploidy over generations. Significance letters are based on GLS. 71

Figure 19: Developmental markers of allopolyploids, autopolyploids, homoploids, and parentals, showing little or no effect of polyploidy. Significance letters are based on GLS. 72

Figure 20: Pollen sizes of polyploids, homoploids, and parentals. Significance letters are

based on GLS. 74

Figure 21: Examples of individuals for which histograms of pollen size had two peaks. 75

1. Introduction

What is polyploidy and why study it?

Polyploidy plays a significant but poorly-understood role in plant evolution and speciation.

Polyploidy, defined as a heritable increase in genome copy number (Wood et al., 2009), has been known since Winkler defined it in 1916 (Bennett, 2004). However, recent technological advances have increased interest in polyploidy in plants and have led to many new discoveries (Husband et al., 2013). A major change is that the role of polyploidy in plant evolution has been continuously revised upward in the last two decades. When the *Arabidopsis thaliana* genome was published in 2000, many were surprised to find it had a polyploid origin, and there was again surprise when maize was found to have a polyploid origin as well (Bennett, 2004). Ramsey and Schemske (2002) estimated that 47 to 70% of flowering plants descended from polyploids, but genome sequencing has shown that many plant species considered to be diploids are actually paleopolyploids, leading to the suggestion that all angiosperms have an evolutionary history that includes polyploidization (Scarpino et al., 2014). There may be a diploidy-polyploidy cycle, with polyploidization events followed by diploidization throughout the evolution of eukaryotes (Comai, 2005). This means that polyploidy has a greater evolutionary significance than was previously believed.

Despite this significance, evolutionary effects of polyploidy are not well understood. For instance, it has been widely believed that polyploidy generates diversity, but recent work suggest that diploids speciate at higher rates (Scarpino et al., 2014).

Reproductive isolation between polyploids and their diploid progenitors is also often assumed, which would facilitate speciation, but gene flow between cytotypes is suspected to be larger than previously believed (Soltis et al., 2010). It has also been suggested that polyploidy and its accompanying genome restructuring can be adaptive, but there is little evidence for this (Soltis et al., 2010; Tayalé and Parisod, 2013). Meyers and Levin (2006) proposed that polyploidy is a ratchet, causing lineages to increase in ploidy without being able to decrease, but even this point is contentious (Bennett, 2004). The consequences of polyploidization, hybridization and selection after polyploidization events are difficult to untangle and remain an open question (Levin, 1983; Soltis et al., 2010).

Regardless of the answers to these many debates, polyploidy has significant effects. Particularly in plants, a large number of duplicated genes come from polyploidy. Duplicated genes can lead to individual genes taking on a subset of the role the ancestral gene played, a process known as subfunctionalization, or potentially allow for greater evolutionary divergence of one of the genes leading to emergence of novel functions, a process known as neofunctionalization (Adams, 2007). Polyploidy has been shown to catalyse chromosomal rearrangements, gene loss, interlocus concerted evolution of ribosomal DNA repeats, unequal rates of sequence evolution of duplicated genes, and changes in DNA methylation (reviewed in Adams, 2007). These processes may increase fertility by facilitating bivalent chromosome pairing (Adams, 2007).

Polyploidy is also extremely common. Studying phylogenies, Wood et al. (2009) found that 35% of species are polyploid relative to generic base numbers in angiosperms and ferns and that polyploidy is distributed equitably among genera. They also found that polyploidy was involved in 15% of angiosperm and 31% of fern speciation events, an

estimate that was four times previous estimates and drastically revised upwards the importance of polyploidization in speciation. Their estimate is conservative, however, because some polyploid species may be unknown, multiple polyploidization events can appear as one event on a phylogeny if they involve the same parental species, and some intraspecific polyploids may be missed as cryptic species (Wood et al., 2009). However, polyploidy is an incomplete speciation mechanism. While polyploidization often occurs with speciation, many studies show significant gene flow between ploidy levels – through triploids or other odd ploidies, through unreduced gametes, or through recurrent polyploidization (Parisod et al., 2010). Wood et al. (2009) also found that polyploidy does not result in increased diversification, speciation, or species richness in their angiosperms or fern descendants, and later Scarpino et al. (2014) found diploids to speciate at higher rates. It is possible that correlations between polyploidy and species richness, which led to earlier assumptions that polyploidy increases speciation, are because older genera are more likely to have more species and to have more polyploids, or that polyploids accumulate in genera because they are morphologically not significantly different from their progenitors (Wood et al., 2009).

The prevalence of polyploidy and the questions about its effects, combined with modern tools such as complete genome sequencing and improved flow cytometry (Kron and Husband, 2015), make plant polyploidy an important question to study. Specific reasons to study polyploidy in plants are reviewed in Bennett (2004) and include:

- Plants make up 90% of the world's biomass. Most if not all of those plants have ancient polyploidizations events in their evolutionary history, many of which are recent events, and many others have tissue-specific somatic polyploidy (Bennett,

2004; Comai, 2005; Soltis et al., 2010; Tayalé and Parisod, 2013; Wood et al., 2009). Therefore, as Bennett (2004) wrote, "life on earth is predominately a polyploid plant phenomenon", and ignoring polyploidy in plants would mean not understanding a huge portion of life on earth. It is also possible that most eukaryotes have gone through many waves of doubling and diploidization, giving polyploidy even more widespread significance (Bennett, 2004; Comai, 2005).

- In the face of Earth's current mass extinction of biodiversity, conservationists need to know the risk to polyploid species compared to diploid species. The loss of an allopolyploid species may mean the loss of multiple genomes, thereby reducing Earth's genetic biodiversity more than the loss of a diploid species (Bennett, 2004).
- The majority of agricultural crops are polyploid (Bennett, 2004; Brownfield and Köhler, 2011). Considering data from 2002, Bennett (2004) found that of the 21 most important crops, 71% are polyploid and they cover 83.7% of cultivated land. Polyploid crops include wheat, rice, maize, soybeans, cotton, sugar cane, potato, alfalfa, oat, chocolate and coffee. Polyploidy majority remained even when Bennett measured only cereals, only pulses, or only fodder, and if importance was judged based on area, production, or monetary value. Therefore, polyploid plants are the base of humanity's food system.
- Due to polyploidy's prevalence in agriculture, most plant breeding involves polyploids. To make polyploidy research useful to plant breeders, it is necessary to know how to predict which genomes can successfully become autopolyploid or allopolyploid (Bennett, 2004). This will require answering basic fundamental

questions on the control and organization of genes and chromosomes (Bennett, 2004), which will also contribute to our general understanding of genetics. Better control and understanding of polyploidy is likely to contribute greatly to plant breeding, in part because genome doubling can rescue otherwise unsuccessful hybrids (Sora et al., 2016). Interspecific hybridization and distant crosses are already known to be a powerful agricultural tool, introducing such traits as male sterility, disease resistance, and stress tolerance (Zhang et al., 2016). Hybridization could improve declining crop diversity and in doing so enrich the human diet (Zhang et al., 2016).

Definitions of polyploids

Polyploidy can be divided into allopolyploidy, resulting from a cross between two species, and autopolyploidy, resulting from genome duplication within a single species or individual (Bennett, 2004; Ramsey and Schemske, 1998). These are not clearly demarked categories or consistently applied definitions, however, and classification of a particular population is often difficult (Comai, 2005). In part this difficulty is due to the difficulty in defining species, with different definitions used in different areas of biology and even within plant biology (Grundt et al., 2006; Soltis et al., 2010). As a result, it may be better to consider a spectrum between autopolyploids and allopolyploids based on the divergence of the parental genomes (Comai, 2005; Ramsey and Schemske, 1998). Autopolyploidy, for example, has been used to describe polyploids resulting from a self cross of a homozygous individual to hybrids between subspecies (Bennett, 2004). Further, some plants identified as autopolyploids may be crosses between cryptic

biological species – which are morphologically identical but reproductively isolated – with doubled genomes to restore fertility (Parisod et al., 2010). Conversely, plants identified as allopolyploids may be crosses between species with very low genetic divergence and display chromosomal behaviour similar to that expected for autopolyploids.

There is a second definition of autopolyploid and allopolyploid based on their chromosome pairing instead of their parental species (Ramsey and Schemske, 1998; Soltis et al., 2010). In this definition, autopolyploids are defined as forming multivalents or having polysomic inheritance while allopolyploids exhibit disomic inheritance. However, using cytological evidence to differentiate between the two groups can be unreliable because other factors affect chromosome pairing and many polyploids have intermediate chromosome pairing patterns (Parisod et al., 2010). It is also very difficult to use chromosome pairing in ancient polyploids, because diploidization may make them appear as allopolyploids when they are not (Soltis et al., 2010). Therefore, the parental species definition will be used here, with the caveat that there are also plants in between allo- and autopolyploidy due to inconsistent or imprecise definitions of species.

There is also a distinction between neopolyploids, which have recently formed; mesopolyploids, which have a genomic signature of polyploidy, but which have begun to diploidize; and paleopolyploids, which are existing populations resulting from ancient polyploidization events and which have fully diploidized genomes (Tayalé and Parisod, 2013). If all angiosperm are paleopolyploids, then there is a question of a functional definition of polyploidy. Bennett (2004) proposes that polyploid should be measured in copies of a minimal genome, the minimum genetic material needed to survive. Since this

is cumbersome to apply, most researchers define polyploidy as the possession of more than two sets of chromosomes (Bennett, 2004; Parisod et al., 2010; Wood et al., 2009). The latter definition will be used throughout this thesis.

Additional definitions include:

- Aneuploid – an individual which has a chromosome number that is not a multiple of the expected haploid number
- Homoploid hybrid – an individual with the expected chromosome number for a hybridization event without polyploidization

How do polyploids form?

Unreduced gamete formation is affected by genetics and possibly by environment

Most polyploids in nature are formed through unreduced gametes (Kreiner et al., 2017; Ramsey and Schemske, 1998; Soltis et al., 2010; Sora et al., 2016; Szadkowski et al., 2011), which are gametes containing the somatic chromosome number (Brownfield and Köhler, 2011). Unreduced gametes can result in gene flow between species and between cytotypes of the same species (Sora et al., 2016). In addition, polyploids generated from unreduced gametes are usually more fit than those generated from the most frequent method of synthetic polyploid production, somatic doubling (Brownfield and Köhler, 2011). Therefore, understanding how and when unreduced gametes are formed is essential to the study of polyploidy.

It is now known that the production of unreduced gametes is heritable and seems to be governed by only a few genes (Parisod et al., 2010). Zhang et al. (2010) produced interspecific hybrids that give rise to hexaploid wheat when their genomes double and

found that unreduced gamete formation in the hybrids depended on maternal line and, depending on the maternal line, may also depend on paternal line. Sora et al. (2016) also found evidence for a genetic effect on unreduced gamete production rates in a study that resynthesized the natural allopolyploid *Brassica napus*. Zhang et al. (2010) found that most individuals had low rates (0.06-2.17%) of unreduced gamete production but some had much higher rates, with one outlier at 26.7%. This finding was echoed in a different study that compared *Brassica* species and found that on average most individuals produced 1.93% unreduced pollen, some had over 5%, and outliers produced rates of up to 71% and 85% (Kreiner et al., 2017). Zhang et al. (2010) also compared rates over two time points and found the rates were consistent within individuals, which provided evidence for a genetic basis. In addition, backcrossed hybrids between *Brassica napus* and *Sinapis arvensis* had more unreduced gametes than their parents, with the 4^{th} generation backcrossed hybrids producing unreduced gametes at higher rates than the 7^{th} generation. This supported the theory that hybridity can increase unreduced gamete formation rates, likely to rescue noncomplimentary chromosome sets (Sora et al., 2016).

However, the growing evidence for a genetic cause of unreduced gametes and the discovery of many species that produce them frequently gives rise to a paradox: unreduced gametes are assumed to decrease fitness, since they often result in abortion or lower-fertility offspring, so selection should not maintain unreduced gametes in a population (David et al., 2004; Ramsey and Schemske, 1998). Kreiner et al. (2017) found that asexual species produced more unreduced gametes than mixed-mating or outcrossing species. This suggests that unreduced gametes are deleterious but are maintained in species with few opportunities for, and lower efficacy of, selection on sexual processes

(Kreiner et al., (2017). The work of Zhang et al. (2010) and Kreiner et al. (2017) both show that a few individuals with unusually high rates of production of unreduced gametes likely contribute to the majority of polyploid formation. Whether there is any advantage to producing unreduced gametes, such as the facilitation of interspecific hybridization, remains in question. Ramsey and Schemske (1998) suggested that unreduced gametes may be caused by pleiotropic effects of otherwise beneficial genes.

It is widely believed that there is an environmental effect on the production of unreduced gametes, specifically that environmental stress increases their production (Parisod et al., 2010; Ramsey and Schemske, 1998; Sora et al., 2016; Tayalé and Parisod, 2013). Proposed stressors include cycling hot and cold, nutrient deficiency, and herbivory, although there is little evidence for many of these and much of the evidence is anecdotal (Sora et al., 2016). For instance, Sora et al. (2016) subjected plants to leaf wounding and to nutrient limitation to the point of leaf discolouration but found no effect for either. However, it is possible that the duration or severity of these stressors were not high enough. They found that older plants produced more unreduced gametes and suggested that the cause may be water or nutrient stress as the plants became pot bound or that there was a lack of resources as older plants funnelled resources into seed development and away from flower development. Their results call into question the common belief that environmental stress increases the production of unreduced gametes. The effect of herbivory especially seems to be only based on a few studies from the 1930s (Kostoff, 1933; Kostoff and Kendall, 1930, 1929), which involved direct damage to flower buds by mites or viruses (Sora et al., 2016). Mason et al. (2011), however, found evidence that some *Brassica* interspecific hybrids produce more unreduced

gametes when exposed to cold temperatures.

This conflict between new results and established beliefs is not uncommon in polyploidy research. Polyploidy research was conducted through the 1900s, but faded out of popularity later in that century and then became popular again in the 2000s when new technology expanded research possibilities (Soltis et al., 2010). This means that much of the new research is being compared to studies many decades older, and sometimes speculations about polyploidy based on those older studies became general beliefs but were supported by little or no experimental data (Husband et al., 2013). Sora et al.'s (2016) observation that little evidence supports the common belief that herbivory increases unreduced gamete production is a perfect example of this issue in the second wave of polyploidy research. A second example is new data finding no evidence for the common belief that polyploids are more prevalent in extreme environments, a belief seemingly based on a few older papers that reference very few examples (Ehrendorfer, 1980; Martin and Husband, 2009; te Beest et al., 2012). Heterogeneity between plant groups may also contribute to varying results (Husband et al., 2013), and conclusions drawn from one study system may not indicate general truths about polyploidy.

Unreduced gametes are formed by meiotic errors

The mechanisms of unreduced gamete production are becoming well understood (Sora et al., 2016). Defects in meiosis, which usually cause abortions, can also often produce viable unreduced gametes (Brownfield and Köhler, 2011; Ramsey and Schemske, 1998). Some of these errors, listed in a review by Brownfield and Köhler (2011) that looked at genetic mutations leading to unreduced gametes in *Arabidopsis*

thaliana and reviewed more briefly in Ramsey and Schemske (1998) include:

- Errors in meiosis I, which can cause univalents and unbalanced segregation in meiosis I or separation of sister chromatids in meiosis I.
- Cell cycle errors that interfere with the control required to complete meiosis. An example of this kind of error is meiosis II not occurring.
- Spindle orientation errors which result in chromosomes that were separated in meiosis I regrouping in meiosis II, in a plant with simultaneous cytokinesis. Interestingly, *Arabidopsis* mutations that cause these errors result in unreduced gametes in male meiosis, where the spindle should be perpendicular to produce a tetrahedral arrangement, but not in female meiosis, which aims for a linear arrangement.
- Cytokinesis errors, for instance when cytokinesis does not occur. Some nuclei in the cell may fuse before mitosis begins, and go on to form unreduced gametes.

These mechanisms of unreduced gamete production can be divided into two types: first division restitutions (FDR), where the unreduced gamete has non-sister chromosomes, and second division restitution (SDR), where the unreduced gamete has sister chromosomes (Brownfield and Köhler, 2011). FDR maintains the heterozygosity of the parent, but SDR greatly reduces it (Brownfield and Köhler, 2011). Understanding these mechanisms is important for plant breeding, because unreduced gametes are useful for crosses between ploidy levels – which have already been used to introduce lower cyanide content and disease and pest resistance to crops – and to generate new polyploids, which can increase genetic diversity and heterosis (Brownfield and Köhler, 2011). One advancement that would be especially useful for plant breeding would be a

mutation that resulted in an unreduced gamete with sister chromatids (SDR) that had no recombination, and therefore were completely homozygous, but such a mutation is not yet known (Brownfield and Köhler, 2011).

While the mechanisms of unreduced gamete formation are becoming well understood, much remains unknown. Questions remain unanswered about the incidence, magnitude, determinants, and favorable conditions of unreduced gamete formation, as well as the relative importance of genetics and environment (Sora et al., 2016). The opportunity to answer these questions is opening up, however, especially as flow cytometry techniques improve. Kron and Husband (2015), for instance, recently developed a technique to quickly and reliably detect unreduced pollen, which will hopefully result in more research on the frequency and causes of unreduced gamete formation across the plant kingdom.

Polyploid formation rates and pathways are largely unknown

Ramsey and Schemske (1998) estimated the formation rate of autopolyploids at 10^{-5} per generation, and suggested that the rate of formation of allopolyploids would likely be less than the autopolyploid formation rate unless there was a high rate of interspecific hybridization. However, this is based on very little data and the formation rates of allopolyploids are largely unknown (Tayalé and Parisod, 2013). Formation rates of autopolyploids are also unknown (Parisod et al., 2010), and made more difficult to calculate by the fact that autopolyploids are often so morphologically similar to their diploid progenitors that they can be indistinguishable (Soltis et al., 2007).

There are three routes to spontaneous allopolyploid formation through unreduced

gametes (Tayalé and Parisod, 2013) (Figure 1). The first pathway, called one-step polyploidization or bilateral polyploidization, occurs when two unreduced gametes merge, one from each parent (Ramsey and Schemske, 1998; Tayalé and Parisod, 2013). The second pathway, called unilateral polyploidization or the triploid bridge, occurs when an unreduced gamete from one parent merges with a reduced gamete from the other parent, resulting in a triploid. The tetraploid is then formed from the triploid intermediary either backcrossing to a parent or crossing with another triploid (Husband, 2004; Ramsey and Schemske, 1998; Tayalé and Parisod, 2013). Using simulation and synthesis of research on *Chamerion angustifolium*, Husband (2004) found that triploids were involved in the formation of 62% of tetraploids and that even partially fit triploids contributed to the establishment or even fixation of tetraploid populations. The third pathway, called two-step polyploidization or the homoploid bridge, occurs when two reduced gametes merge to form a homoploid hybrid, which then produces a tetraploid offspring. This third pathway may be quite common, since homoploid hybrids produce more unreduced gametes than non-hybrids. In an analysis of several studies, Ramsey and Schemske (1998) found hybrids produced fifty times more unreduced gametes than non-hybrids. The effective frequency may be even higher due to the fact that reduced gametes in hybrids often have errors such as aneuploidy that may make them inviable; indeed, in their review they found a higher rate of two-step polyploid production than would be expected just from the unreduced gamete production rate of hybrids that they estimated. However, they warn that the unlikelihood of producing viable, fertile hybrids may make allopolyploids from this pathway less common than expected. In addition to these three pathways, polyploids can be formed from other polyploids, for example allopolyploids

from a cross between autopolyploids of different species (Ramsey and Schemske, 1998). A review of the literature found no studies that have explicitly calculated and compared formation rates through different pathways.

Allopolyploidy formation may be more likely between some genomes than between others. The probability of genomes being able to contribute to a successful allopolyploid appears to depend on the relatedness of the genomes (Levin, 2013). This observation has led to the hypothesis that plant genomes merge, diverge, and merge again (Tayalé and Parisod, 2013). Levin (2013) suggested that a good system for understanding allopolyploid formation is one in which the genomes are relatively closely related. This observation is supported by the fact that recently evolved allopolyploids are often found in human-disturbed habitats where the introduction of non-native species related to native species has increased hybridization and allopolyploidy (Ramsey and Schemske, 1998).

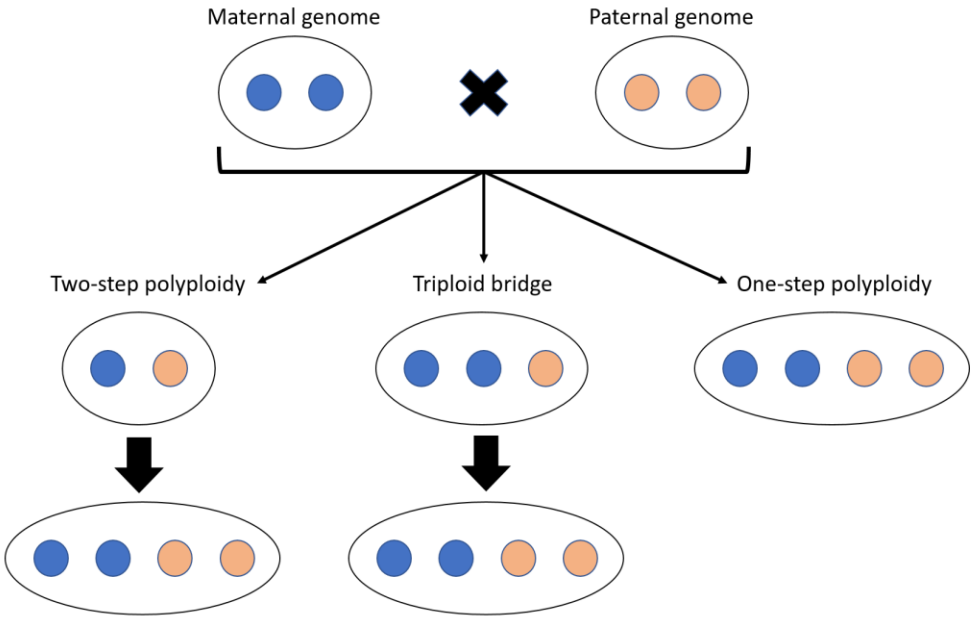


Figure 1: The three pathways of polyploid formation: one-step, triploid bridge, and two-step. Modified from Tayalé and Parisod (2013).

Formation pathway may affect probability of formation and establishment

When the genomes of two species come together in an allopolyploid, the results appear to be repeatable, with separate formation events resulting in similar morphology and genome restructuring. For example, Buggs *et al.* (2014) found that the gene expression of parental diploids is often maintained in allopolyploids. The allopolyploids in the genus *Tragopogon* L. form repeatedly in nature, exhibiting the same homeolog loss and repeated patterns of tissue-specific silencing (Soltis et al., 2009). In studies of wheat, genetic and epigenetic changes were also found to be repeatable and to mimic changes found in nature (Levy and Feldman, 2004). Additionally, naturally-occurring polyploids like *Brassica napus* can be resynthesized (Song et al., 1993; Zhang et al., 2004).

However, there is evidence that different methods or formation may affect the resulting genome structure and the rate at which fertility recovers. For example, Szadkowski *et al.* (2011) found differences in the meiotic behavior of *Brassica oleracea* L. and *Brassica rapa* L. allopolyploids formed by unreduced gametes or by somatic doubling via colchicine. Those produced by unreduced gametes had more translocations between homoeologous chromosomes, which would lead to more genome instability, but the somatic doubled plants had larger translocations, which decrease stability more than small ones (Szadkowski et al., 2011).

This raises the possibility that polyploids formed by unreduced gametes through different pathways may have different results for genome structure, organization and meiotic behavior. Allopolyploids that form via a pathway that results in a faster recovery

of fertility will be more likely to successfully establish populations. It is currently unknown if there are differences between allopolyploids produced by unreduced gametes through the different pathways shown in Figure 1 (Tayalé and Parisod, 2013).

What are the effects of polyploidy?

Hybridization causes more genetic and epigenetic changes than genome doubling

Polyploidy is a large-scale mutation that has significant effects. Epigenetic changes are common in allopolyploids (Parisod et al., 2010), including gene silencing and upregulation caused by DNA methylation, retrotransposon activation, and histone modifications (Adams, 2007). Gaeta and Pires (2010) have proposed that polyploidy may involve a ratchet, where translocation mutations accumulate and eventually result in genome instability and sterility. In addition, it is not unusual for there to be a bias against the expression of genes from one parent in allopolyploids, and this bias is immediate (Adams, 2007; Szadkowski et al., 2011). The reason that genes are silenced, downregulated, or upregulated in polyploids likely varies by organism or gene; potential reasons include accommodations for increased gene dosage, altered regulatory networks, epigenetic remodeling, interactions between homeologs from different parents, or a side effect of other mechanisms (Adams, 2007). Bennett (2004) has observed that new genes may not be necessary to manage polyploidization if all angiosperm are indeed paleopolyploids; instead, genes from previous rounds of polyploidization could be reused to manage new polyploidy events.

Adams (2007) identifies two main types of changes in neopolyploids. The first is silencing of one homeolog or a large bias against one. This can occur immediately and

sometimes synthetic neopolyploids have the same patterns of expression of natural polyploids with the same progenitor species, suggesting that the changes are repeatable (Adams, 2007). The changes can also vary by generation in neopolyploids (Wang et al., 2004). The second change is non-additive gene expression, which is studied by measuring the products of both homeologs simultaneously. This change can also be immediate, and the nonadditive difference may be quite large (Adams, 2007). Some of these changes have been shown to affect phenotype (Adams, 2007). While the majority of these studies look at RNA transcripts, Albertin et al. (2006) demonstrated that protein abundance is also affected (Adams, 2007). Levy and Feldman (2004) describe three changes in neopolyploids: non-random elimination of both coding and non-coding DNA, epigenetic changes including methylation, and activation of retroelements which alters gene expression – all of which occur in the formation of wheat allopolyploids. They emphasize that elimination of non-coding sequences will increase the differences between homoeologous chromosomes and contribute to diploidization.

These changes in gene expression in polyploids may be organ-dependent, which could lead to subfunctionalization, which occurs when gene function is partitioned so that both copies of a duplicated gene are necessary for survival (Adams, 2007). Adams et al. (2003) found different homeologs silenced in different organs of *Glossypium hirsutum*. Subfunctionalization may be a frequent fate of allopolyploid homeologs, and would contribute to speciation by requiring individuals to have two copies of a gene and thus making inter-population hybrids less likely to be viable (Adams, 2007). Subfunctionalization can also drive diploidization (Le Comber et al., 2010).

Despite the large-scale changes inherent in polyploidy, most of the genome

restructuring following allopolyploid formation is in fact a result of hybridization, not of genome doubling itself (Adams, 2007; Hegarty et al., 2006; Rieseberg, 2001). Genome doubling may even mitigate the extensive gene expression changes of homoploid hybrids (Hegarty et al., 2006), an idea contrary to the common belief that polyploidy is detrimental (Stebbins, 1971). In one study, 89% of gene expression changes in allopolyploids were found to be due to hybridization (Albertin et al., 2006). Hybridization and backcrossing without polyploidization can generate large amounts of genetic variation, including large morphological diversity and novel phenotypic mutations (Zhang et al., 2016). Homoploid hybrids also produce offspring with a wide range of DNA contents (Zhang et al., 2016), which supports the hypothesis that polyploidy, instead of causing genomic problems, may rescue otherwise incompatible hybrid genomes (Sora et al., 2016). Hybridization has been shown to cause nonadditive gene expression, over- or underdominance, unequal allelic or monoallelic expression, organ-specific silencing, loss of maternal and paternal imprinting, and some parent-of-origin effects (reviewed by Adams, 2007) – all changes frequently found in allopolyploids (Adams, 2007) which seem to be an effect of hybridity, not polyploidy itself. The changes in allopolyploids are also rapid, occurring in a homoploid hybrid parent or in the first generations of polyploidy (Levy and Feldman, 2004).

Examinations of autopolyploids reveal what genetic changes polyploidy itself may cause. Autopolyploids are often so similar to their diploid progenitors that they are difficult to find, and their prevalence may be greatly underestimated (Soltis et al., 2007). There are only few instances of documented epigenetic instability in autopolyploids (Comai, 2005). Autopolyploids sometimes rapidly eliminate DNA, but they do not

always do so and they show less genome restructuring than allopolyploids (Parisod et al., 2010). There is little evidence for long-term genome reorganization in autopolyploids or for large effects on gene expression, despite the fact that theory suggests neofunctionalization or subfunctionalization would be consequences of genome duplication (Parisod et al., 2010). One study comparing diploids and autopolyploids found no difference in the proteomes (Albertin et al., 2005), though protein analysis is less sensitive than mRNA analysis (Comai, 2005). One of the biggest long-term changes in autopolyploids is diploidization – the transition from polysomic to disomic inheritance. Using computer simulations, Le Comber et al. (2010) found that genetic drift and homologue pairing fidelity is all that is required for polysomic inheritance to transition to disomic inheritance in autopolyploids. Therefore, the effects of genome doubling itself may be minimal. Both the advantages and disadvantages of allopolyploids may be due mostly to hybridity, with genome doubling simply a mechanism to restore fertility to interspecific hybrids.

Odd-ploidy plants such as triploids and pentaploids have their own unique changes. Their fertility can vary greatly by species from seedless to a few seeds to fertile, though their fertility remains lower than diploids (Comai, 2005). Some amphibians and plants have long-term odd-ploidy, with specialized meiosis that produces haploid sperm and diploid eggs or vice versa (Comai, 2005). There is also evidence of an odd-ploidy response in plants. One study in wheat found that for 10% of the genes examined, gene expression follows the expected pattern of increase in even-ploidy individuals but is the inverse or deviates in haploid and triploid individuals (Guo et al., 1996). Aneuploidy, common in neopolyploids of all types, can cause epigenetic effects by affecting dosage or

by exposing unpaired chromatin to remodelling mechanisms (Comai, 2005). These meiotically unpaired chromosomes are susceptible to silencing, which may contribute to the odd-ploidy response (Comai, 2005).

The evolutionary advantages of polyploidy

Polyploidy was originally believed to be detrimental (Stebbins, 1971), and indeed it brings with it many problems. Minority cytotype disadvantage, when individuals with a rare cytotype reproduce less successfully because of their incompatibility with the dominant cytotype, is one of the biggest problems (Oswald and Nuismer, 2007; Parisod et al., 2010), though the fact that there is more gene flow between cytotypes than previously believed may mitigate their isolation (Soltis et al., 2010). Since minority cytotype disadvantage is caused by the rare cytotype being flooded with pollen from the more common cytotype, resulting in low-fitness hybrids, its effects can also be mitigated by prezygotic reproductive barriers such as different flowering times or spatial structure influencing pollinator behavior that would decrease the chances of cross-cytotype pollination (Husband and Schemske, 2000). Selfing, perenniality, asexual reproduction, and gene flow between cytotypes may all help neopolyploids to overcome their minority cytotype disadvantage (Oswald and Nuismer, 2007; Parisod et al., 2010). However, to become established neopolyploids they would need more: a competitive advantage, ecological divergence, favorable random chance, or possibly recurrent polyploidy (Parisod et al., 2010). The prevalence of polyploidy suggests that there is some edge that allows polyploids to thrive and implies an evolutionary flexibility that contradicts Stebbins's (1971) suggestion that gene redundancy in polyploids would cause species to

be static (Comai, 2005). In contrast, Meyers et al. (2006) has proposed a null model of polyploidy, that it is simply a ratchet and lineages gradually increase in ploidy level because they are unable to decrease. This would make the process neutral with no evolutionary advantage. However, some research has found that polyploids decrease in chromosome count during diploidization (Mandáková et al., 2017; Mandáková and Lysak, 2018).

Modern research has found several possible advantages to polyploidy that may explain how neopolyploids are able to establish themselves and thrive. Oswald and Nuismer (2007) built mathematical models of host-pathogen relationships in plants and their results suggested that polyploids may be more resistant to pathogens, an idea first proposed by Levin (1983). There is little empirical evidence for this theory in plants (Oswald and Nuismer, 2007), but there is evidence in frogs that polyploidy may be a response to a parasite, resistance to which can arise from interspecific hybridization (Jackson and Tinsley, 2003). Polyploids may have an advantage during climate-driven environmental change and in recently disturbed landscapes; in the case of autopolyploids, these advantages would depend on genic redundancy and polysomic inheritance (Parisod et al., 2010). Gene redundancy in polyploids can mask deleterious alleles even at the gametophytic stage, protects against the dangers of deleterious recessive alleles and genotoxicity during inbreeding, and allows for subfunctionalization or neofunctionalization (Comai, 2005). Autopolyploids would have a 50% decrease in inbreeding depression due to polysomic inheritance (Parisod et al., 2010). In addition, polyploidy can disrupt self-incompatibility systems (Miller and Venable, 2000), which may give individuals an advantage when mates are scarce (Comai, 2005). Allopolyploids

may be successful due to heterosis, with their genomes doubled to restore fertility (Parisod et al., 2010). Allopolyploids are able to experience the advantages of heterosis longer because they maintain heterozygosity longer than diploid hybrids due to forced pairing of homologous chromosomes (Comai, 2005). Nonadditive gene expression, small RNAs, and epigenetic regulation in homoploid hybrids and allopolyploids contribute to heterosis (Chen, 2010).

These factors may contribute to the invasiveness of polyploids. Pandit et al. (2011) found that polyploids are 20% more likely than diploids to be invasive and that doubling chromosome number makes invasiveness 12% more likely. Polyploids are also less likely to be endangered (Pandit et al., 2011). Te Beest et al. (2012) examined the reasons polyploidy may correlate with invasiveness and determined that it may be due to polyploidy itself – for reasons such as higher original fitness, greater opportunity for adaptation due to a larger gene pool, or allowing for asexual reproduction – or due to polyploidy restoring fertility after hybridization, which can increase invasiveness as well (Ellstrand and Schierenbeck, 2000).

Despite these possible advantages, all neopolyploids go through a bottleneck of instability and reduced fertility before being able to compete successfully (Comai, 2005). There are also serious problems caused by polyploidy: increases in DNA volume that change interactions between chromatin and nuclear envelope proteins; dosage effects; increased aneuploidy due to difficulty in mitosis or meiosis (30-40% of the offspring of autopolyploid maize is aneuploid); intergenomic recombination in allopolyploids that can lead to problems as well as adaptations; and epigenetic instability in allopolyploids (reviewed by Comai, 2005). These instabilities and changes, however, provide material

for selection that may give polyploids advantages in the long-term (Comai, 2005; Levy and Feldman, 2004).

Experimental system

The triangle of U offers an ideal study system for polyploidy

To examine some of the questions on the prevalence and effect of polyploidy, the study described in this thesis used the species in the "triangle of U." The triangle of U is a diagram indicating the relationships among six *Brassica* species: *B. nigra, B. carinata, B. napus, B. juncea, B. oleracea,* and *B. rapa* (Figure 2). These relationships were reported by U in 1935, building on the work of Morinaga (1929). U determined that *B. carinata* (BBCC), *B. juncea* (AABB), and *B. napus* (AACC) were all allotetraploids formed by *B. nigra* (BB) and *B. oleracea* (CC), *B. rapa* (AA) and *B. nigra* (BB), and *B. oleracea* (CC) and *B. rapa* (AA), respectively. *Sinapis arvensis* (SrSr), also used in the present study, is closely related to the B genome or *B. nigra*.

Neoallopolyploids are known to form between *B. carinata* and *S. arvensis* via the one-step polyploidization pathway (Cheung et al., 2015) and the two-step pathway (Martin et al., unpublished data) allowing for the characteristics and formation rates of these two types of neo-allopolyploids to be compared. Because the three allopolyploids in the triangle of U are well-characterized they have been used in previous polyploid research (Albertin et al., 2006; Cheung et al., 2015; Kreiner et al., 2017; Kron and Husband, 2015; Liu et al., 2014; Mizushima, 1950; Sora et al., 2016; Szadkowski et al., 2011; Zhang et al., 2016), both as parents of neoallopolyploids and as tetraploids themselves, which allows this research to build on existing polyploidy research. The

system also allows for the hypothesis that allopolyploid formation will be higher when closely related genomes are involved (Levin, 2013) because *B. carinata* and *B. juncea* share the B genome and therefore are predicted to have higher formation rates with *S. arvensis*, while *B. napus* does not have the B genome and therefore would be predicted to have a lower allopolyploid formation rate with *S. arvensis*. Additionally, several triangle of U species have been sequenced (*B. rapa, B. napus, B. juncea,* and *B. oleracea*), which will facilitate more advanced molecular genetic research in the future.

The species in the triangle of U are also important crop species in Canada, as detailed below. Their close relative, *S. arvensis*, is a widespread weed and Agriculture and Agri-Food Canada and the Canadian Food Inspection Agency are interested in the possibility of hybridization between the triangle of U crops and *S. arvensis*, especially lines of crops incorporating genes for herbicide resistance. In particular, *B. carinata* is being developed as a novel oilseed crop for industrial uses, including biofuels (Marillia et al., 2014), but it is the least studied of the Brassica species with little work evaluating reproductive compatibility (FitzJohn et al., 2007). This research contributes to their investigation by studying whether spontaneous allopolyploidy increases the chances of a crop-weed hybrid population becoming established and providing a route for transgenes to escape into wild populations.

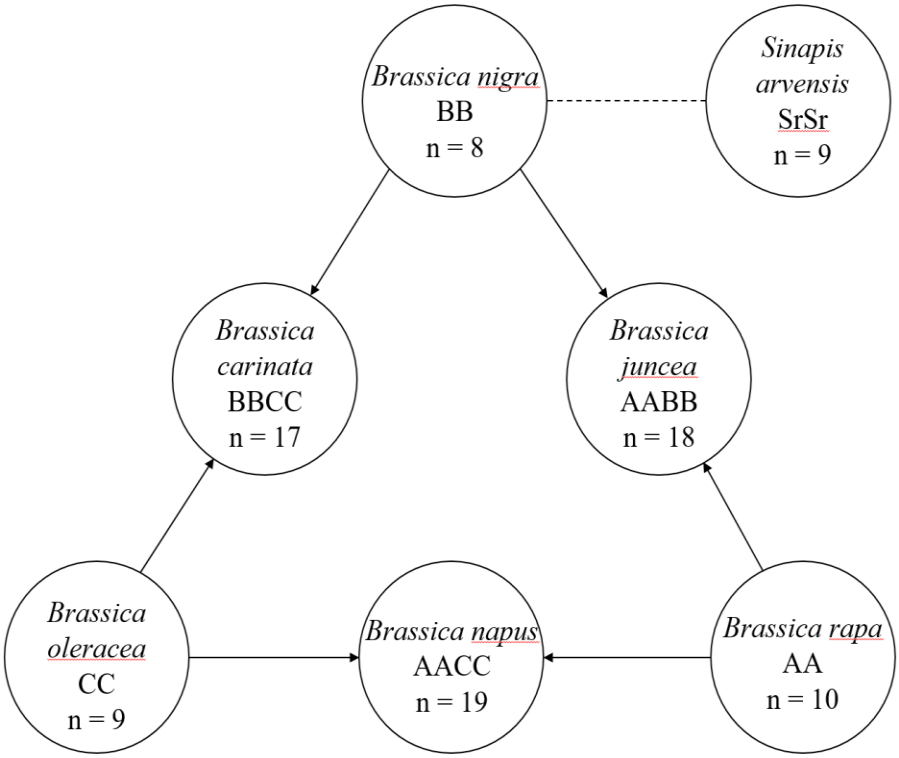


Figure 2: The triangle of U. S. arvensis is most similar to the B genome. Modified from U (1935).

Brassica carinata (Ethiopian mustard; BBCC, 2n=34)

Brassica carinata Braun, commonly known as Ethiopian mustard, is an oilseed crop being considered for cultivation in Canada. It is an allotetraploid with disomic inheritance (Mizushima, 1950), the result of interspecific hybridization between the diploid species *Brassica nigra* (L.) Koch and *Brassica oleracea* L. (Edwards *et al.*, 2007; FitzJohn *et al.*, 2007). The species is a self-compatible annual with no primary seed dormancy. *B. carinata* has been grown as an oilseed and vegetable crop in Ethiopia and India (FitzJohn *et al.*, 2007), but as demand for vegetable-based oils increases it is being

considered for wider cultivation because of its drought-resistance, shatter-resistance, and resistance to several pests (Marillia *et al.*, 2014).

Brassica juncea (Indian mustard; AABB, 2n=36)

Brassica juncea (L.) Czern is understood to be a natural allotetraploid formed by *B. napus* and *B. rapa.* In Canada, *B. juncea* is mostly grown for condiment mustard in the prairie provinces (Government of Canada, 2012), though it has also been developed as a canola-quality edible oil (Potts *et al.*, 1999). It is a self-compatible annual crop that does not shatter readily. It is more tolerant of heat and drought stress than the more common canola species, *B. napus* and *B. rapa* (Woods *et al.*, 1991).

Brassica napus (Canola; AACC, 2n=38)

Brassica napus L. is one of the two canola species grown in Canada, the other being *B. rapa*. It is an ancient crop that originated as an allotetraploid of *B. rapa* and *B. oleracea* and has been grown in India, Europe, and China for thousands of years. In Canada, it is mainly grown in the prairie provinces, usually south of the region where *B. rapa* is grown because *B. napus* requires a longer growing season (Government of Canada, 2012). It is a self-pollinating annual crop.

Sinapis arvensis (Wild mustard; SrSr, 2n=18), a persistent weed

Sinapis arvensis L., commonly known as wild mustard, is a common weed in Canada. *Sinapis arvensis* a self-incompatible annual species. It has a persistent seedbank, competitive growth, high fecundity and indeterminate growth (Warwick *et al.*, 2000).

Before the widespread use of herbicides, it was considered the worst weed in the cultivated land of the prairies, and it has already evolved some herbicide resistance (Mulligan and Bailey, 1975). *Sinapis arvensis* is found in all Canadian provinces (Warwick *et al.*, 2000). Although no interspecific hybrids have been confirmed in nature, *S. arvensis* has been intentionally hybridized with some members of the subtribe Brassicinae (including *B. napus, B. juncea,* and *B. nigra*), mostly through embryo or ovule rescue (Warwick *et al.*, 2000; FitzJohn *et al.*, 2007).

Research Objectives

The focus of this thesis is characterizing how the formation pathway of Brassicaceae allopolyploids affects progeny probability of establishment, including progeny fertility, morphology, and formation rates. This will contribute to the under-studied question of polyploid formation rates and the effect of allopolyploid formation pathway (Tayalé and Parisod, 2013). These questions are difficult to answer by observing natural populations because even when polyploids are found there is no way to know how many polyploids may have formed and not established. By examining formation rates and subsequent fertility, it will be possible to determine whether polyploids form frequently and establish rarely or form rarely and establish frequently, which will have consequences for how polyploidy is considered in the context of plant evolution. In addition, this thesis compares the impact of hybridization and polyploidization on the morphology of allopolyploids formed through different pathways, another important question (Soltis et al., 2010; Tayalé and Parisod, 2013).

There were three main questions:

1) What pathway is most frequent for allopolyploid formation: the one-step pathway or the two-step pathway through a homoploid hybrid bridge? Does this vary by parental species?

2) Are there differences in the fertility of polyploids formed through these two pathways? How does their fertility compare to parental fertility?

3) Are there morphological differences in allopolyploids produced by each pathway? Does hybridity or polyploidy contribute more to their phenotypes?

The hypothesis is that one-step allopolyploids will have higher formation rates than two-step allopolyploids because the low fertility of homoploid hybrids will result in few two-step allopolyploids. This is expected regardless of parental species, though *Brassica* species more closely related to *S. arvensis* are expected to more easily produce hybrids. The two-step allopolyploids that are produced are expected to have higher initial fertility because incompatibilities between the genomes will have been filtered out by the success of the homoploid hybrid parent. Their initial fertility is expected to be lower than parental fertility. One-step and two-step allopolyploids are expected to have similar morphology, and hybridity is expected to contribute more to their phenotypic divergence than polyploidy because it introduces more new genetic material and causes greater changes in gene expression.

2. Materials and Methods

Plant growth procedure

Parental seeds of *S. arvensis* were collected in the field and *Brassica* species parental seeds were obtained from seed banks (Table 1). Accessions of Ethiopian mustard were chosen by Cheung et al. (2015) to include the widest variety of geographic origin. Two accessions each of *B. juncea* and *B. napus* were used, chosen to provide the greatest genetic diversity available through Plant Gene Resources of Canada (PGRC). The accessions chosen for *B. juncea* were Varuna and Cutlass (Srivastava et al., 2001) and for *B. napus* were the spring cultivars Westar and Global (Lombard et al., 2000). Seed generated by the homoploid hybrids produced in hand crosses between Ethiopian and wild mustard (Cheung et al., 2015), and from insect mediated crosses between the species (Martin et al., in preparation) were also used.

Seeds were sown 0.25 cm deep in soil, peat, and sand (1:2:1 by volume) in 72 cell flats and put in a glasshouse with a 16 hour photoperiod, 26° C days, and 18° C nights. After flow cytometry testing at the rosette stage, plants of interest were transplanted into 10 cm pots.

Table 1: Source of seeds.

ID Number	Species	Source	Place of Origin
1696	*B. napus* (Westar)	PGRC (CN 42942)	Saskatchewan, Canada
1697	*B. napus* (Global)	PGRC (CN 46333)	Ontario, Canada
1698	*B. juncea* (Cutlass)	PGRC (CN 46238)	Saskatchewan, Canada

1699	*B. juncea* (Varuna)	PGRC (CN 105188)	India
8717	*B. carinata*	PGRC (CN 101648)	Pakistan
8736	*B. carinata*	PGRC (CN 101665)	UC Davis
636	*B. carinata*	B and T world seeds	Paguignan, France
3385	*S. arvensis*	Wild	Ottawa, ON, Canada
1686	*S. arvensis*	Wild	Embrun, ON, Canada
8180	*S. arvensis*	Wild	Fez, Morocco
9241	*S. arvensis*	Wild	Theodore, SK, Canada

Hand pollinations

Approximately 1000 flowers each of *B. napus* and *B. juncea* were pollinated by hand with *S. arvensis* pollen. In addition, allopolyploids were hand crossed with their parents as well as with their homoploid and polyploid siblings to examine their reproductive isolation (10-20 flowers per cross). Flowers were emasculated and then pollinated the next day with pollen from another individual. For *S. arvensis*, the individual selected for the backcross was from a different accession than the *S. arvensis* ancestor of the allopolyploid to avoid issues of self-incompatibility. As controls, five flowers were emasculated only, five were emasculated and self-pollinated, and five unmanipulated pods were collected.

Insect-mediated pollinations

Polyploid or homoploid hybrid individuals of the same lineage and generation were placed in a tent with bluebottle flies (*Calliphora vomitoria*, Forked Tree Ranch, Idaho, USA). The flies were received in larvae form and incubated at 27 ºC for two to three days to induce hatching before being released into the tents in the greenhouse. New flies were released weekly until the plants finished flowering or they had been flowering for four months.

Flow cytometry

Flow cytometry was done on leaf material using the method described in Martin et al. (2017) to identify the DNA content of individuals. Flow cytometry was run on pollen of the colchicine-transformed *B. carinata* autotetraploids using the procedure developed by Kron and Husband (2012) to determine which flowers were tetraploid and which were diploid, in order to collect tetraploid seed.

For leaves, tissue from the newest leaf on plants about two weeks old was wrapped in a moist paper towel and placed on ice. About 0.25 cm^2 of the leaf tissue was chopped with a razor in 750 µL of Galbraith buffer (Doležel and Bartoš, 2005). For pollen, six anthers were collected, two each from three flowers, and vortexed in 500 µL of Galbraith buffer to release the pollen. The chopped leaf tissue was filtered with a 30 µm filter (Sysmex Partec, Japan) and collected in a falcon tube. The pollen was filtered through a 100 µm filter to catch large pieces and a 20 µm filter to catch the pollen. The pollen on the 20 µm filter was burst with a pestle and then washed with 250 µL of

Galbraith buffer three times.

For both leaf and pollen samples, the nuclei were treated with 50 µL of RNAase (1 mg/ml) and 250 µL of propidium iodide (0.1 mg/ml) on ice in the dark for half an hour. Samples were then run through a Gallios flow cytometer (Gallios 1.2, Beckman Coulter Inc., California, USA). Samples were run at 480 V for 230 s on medium speed, with the exception of *B. carinata* autotetraploids, which were run at 430 V due to their higher DNA content. The expected DNA contents for parents and hybrids are shown in Supplementary Table 1.

In the first round of screening for potential hybrids and polyploids, leaf samples from two plants were chopped together and run in the same tube, with *Raphanus sativus* L. "Saxa" used as an external standard. If hybrids or polyploids were identified, tissue from those plants was run on three different days with an internal standard, as recommended by Dolezel *et al.* (Doležel et al., 2007). If a large number of hybrids with similar DNA content were identified, only a subset were run in triplicate. *R. sativus* (1.29 pg, (Cheung et al., 2015)) or *Camelina sativa* (1.54 pg, (Martin et al., 2017)) was used as the internal standard for *Brassica* parentals and hybrids, and *Glycine max* (L.) Merr. "Polanka" (2.5 pg, (Doležel et al., 1992)) was used for *S. arvensis*, to avoid overlap between the standard and the sample peak.

Autopolyploid induction

B. carinata autopolyploids were generated by treating *B. carinata* seeds with colchicine. Specifically, 42 seeds were sterilized in 70% ethanol for 1 minute and in 50% bleach for 5 minutes, and then treated with a 0.05% colchicine solution for 24h, rinsed

and planted into trays. Once the seedlings started to bolt, conversion was tested using flow cytometry. A conversion rate of approximately 12% was expected (Martin and Husband, 2012) with recovery of up to 5 autopolyploids anticipated. Converted individuals were allowed to self-pollinate. Because partial conversion of the meristematic tissue is possible, converted individuals were retested at two week intervals to determine if diploid tissue had overtaken tetraploid tissue. When flowering, pollen was tested on two different branches of each converted individual using flow cytometry (described above) to be certain which flowers were tetraploid. The seeds that developed from tetraploid flowers were grown and verified as tetraploid using flow cytometry.

Characterization of fertility and morphology

Morphology and male and female fertility of homoploid, allopolyploid, autopolyploid, and parental lines were characterized.

The 23 morphological traits that were measured were chosen based on traits that are considered to frequently distinguish polyploids in other systems, such as larger flowers and stomata and delayed development (Materson, 1994; Ramsey and Schemske, 2002; Sax and Sax, 1937, see Table 2). They include flower measurements, stomata size, pollen size, and developmental markers such as bloom date and height. The age at measurement was also recorded.

The fertility of the allohexapolyploids from both pathways was characterized across three generations for material from *B. carinata* crossed with *S. arvensis*. Pollen fertility was scored by staining two to six anthers with one drop of acetocarmine, counting 500 to 1000 grains and scoring them as viable or nonviable under a Leica

DFC450 microscope. Seed production was determined for each individual by dividing total seed weight by the average of three 100 seed weights, and compared to parentals grown in the same environment. Seed size was also measured, as the cross sectional spherical area viewed under a Leica M205 C microscope.

Table 2: Data collected for morphological analysis.

Category	Trait	Significance
Developmental	Plant height at flowering (cm)	Delayed development is believed to be linked to polyploidy (Ramsey and Schemske, 2002)
	Age at measurement (days)	Included to verify that observed differences are not due to different ages of the plant
	Days to germinate (days)	Delayed development is believed to be linked to polyploidy (Ramsey and Schemske, 2002)
	Days to flower (days)	Delayed development is believed to be linked to polyploidy (Ramsey and Schemske, 2002)
Flower	Petal width (mm)	Increased flower size is believed to be linked to polyploidy (Ramsey and Schemske, 2002)
	Petal length (mm)	
	Claw length (mm)	
	Sepal length (mm)	
	Anther length (mm)	

	Gyneocium length (mm)	
	Beak length (mm)	
	Percent of petal that is claw (%)	
	Petal width over height	
	Purple dots on anthers	Some species studied have purple dots on anthers and some do not
	Difference between anther and gynoecium (mm)	Anthers tend to be longer than gynoeciums in selfing species (*B. carinata*) but shorter in self-incompatible species (*S. arvensis*)
	Percent of gynoecium that is beak (%)	The species studied have different gynoecium proportions
	Flower length (mm)	Increased flower size is believed to be linked to polyploidy (Ramsey and Schemske, 2002)
	Flower display size (mm)	Increased flower size is believed to be linked to polyploidy (Ramsey and Schemske, 2002)
	Pedicle length (mm)	The species studied have different pedicle lengths
Stem	Trichome density per	Some species studied have trichomes and

	cm^2	some do not
	Trichome length (mm)	
	Stem angle	The species studied have different stem angles
	Purple joints	Some species studied have purple joints and some do not
Leaf	Stomata length (um)	Increased stomata size is believed to be linked to polyploidy (Materson, 1994; Sax and Sax, 1937)

Allopolyploid formation rate estimation

Because no studies were found that compared formation rates through different allopolyploid formation pathways, the following original method was developed. Polyploidization rates were calculated per ovule challenged. To calculate the number of parental ovules challenged to produce the one-step allopolyploids ($Ovules_P$), the number of hand-pollinated flowers (F) was multiplied by the average number of seeds per self-pollinated flower ($\overline{Seed/F_S}$), which was assumed to be the number of ovules per flower challenged by hand pollination.

Equation 1:

$$Ovules_P = F \times \overline{Seed/F_S}$$

The one-step allopolyploidization rate (R_{1-step}) was then the ratio of the number of one-step allopolyploids (A_{1-step}) to $Ovules_P$.

Equation 2:

$$R_{1-step} = {}^{A_{1-step}}\!/_{Ovules_P}$$

The two-step allopolyploidization rate (R_{2-step}) was estimated as the product of two values corresponding to the two steps: the homoploid hybridization rate (R_H) and the polyploidization rate of the homoploid hybrids ($R_{2-step}|H$).

Equation 3:

$$R_{2-step} = R_H \times (R_{2-step}|H)$$

The homoploid hybridization rate was calculated as the number of homoploid hybrids (H) divided by the number of parental ovules challenged ($Ovules_P$).

Equation 4:

$$R_H = {}^{H}\!/_{Ovules_P}$$

Similarly, the polyploidization rate of the homoploid hybrids was calculated as the number of two-step allopolyploids (A_{2-step}) divided by the number of homoploid hybrid ovules challenged ($Ovules_H$).

Equation 5:

$$R_{2-step}|H = {}^{A_{2-step}}\!/_{Ovules_H}$$

The number of homoploid hybrid ovules challenged could not be calculated in the same way as the number of parental ovules challenged because the homoploid hybrids were cross-pollinated with insects and not by hand, so the number of flowers was unknown. Additionally, the fertility of the homoploid hybrids was too low for the number of seeds per self-pollinated flower to be a reliable indicator of the number of ovules per flower. Therefore, the seed sets of parental plants were averaged to obtain a hypothetical number of ovules per homoploid hybrid ($\overline{Seed\ Set_P}$). To account for the lower fertility of

homoploid hybrids, the male fertility of homoploid hybrids as measured by pollen viability was used to decrease the expected number of ovules per plant. This assumed male and female fertility were equivalent. Finally, this estimate of ovules per homoploid was multiplied by the number of homoploids (H).

Equation 6:

$$Ovules_H = \overline{Seed\ Set}_P \times \frac{\overline{Pollen\ Viability}_H}{\overline{Pollen\ Viability}_P} \times H$$

Statistical analysis

All statistical analyses were done in R (version 3.4.0) with the following packages installed: plotrix and hmsic for graphing; MASS for LDA analysis; agricolae for summary of Tukey results; nlme, lsmeans, and multicompView for GLS analysis; and car and nlme for ANOVAs. Because many ANOVAs were calculated when analysing the morphological data, the p-values were adjusted to control for multiple comparisons using the R function p.adjust with the False Discovery Rate method.

The R package flowPloidy (Smith et al., 2018) was used to analyze the results of the flow cytometry. Only samples with more than 1000 events in each of the sample peak and the standard peak were included in the calculation.

3. Results

One-step allopolyploid formation rates

The rate of one-step allopolyploid formation between *B. carinata* and *S. arvensis* was 1 allopolyploid per 11,700 ovules challenged (Cheung et al, 2015). *B. juncea* and *B. napus* did not produce any one-step allopolyploids when crossed with *S. arvensis* (Table 3).

The homoploid hybrid production rate was 729 homoploids from 11,300 ovules challenged in the *B. carinata* crosses (Cheung et al, 2015), 27 homoploids from 12,000 ovules challenged for the *B. juncea* crosses, and 2 homoploids from 23,800 ovules challenged for the *B. napus* crosses (Supplementary Table 2). The number of hand crosses conducted for each cross were 997, 984, and 1022 respectively. This gave a power to detect hybridization of 3 in 10,000 for *B. carinata* and 4 in 10,000 for *B. juncea,* both much lower than the rate detected, and 1 in 10,000 for *B. napus*, which was approximately the rate detected (95% confidence interval, Jhala et al., 2011). For both *B. juncea* and *B. napus*, all homoploid hybrids came from only one of the two parental accessions crossed. Cheung et al (2015) similarly found that the *B. carinata* accessions differed in hybridization rate. It is possible that these hybridization and one-step formation rates are slightly overestimated if inbreeding depression resulted in the termination of some ovules in the self-pollinated controls. However, the effect of inbreeding depression on seed set in these species has not been well quantified. Future studies should include an intra-species cross to examine this possibility.

Two-step allopolyploid formation rates

The rate of two-step allopolyploid formation was estimated by sowing seeds from the homoploid hybrids (Supplementary Tables 3, 4). A rate of 1 two-step allopolyploid per 25,600 ovules challenged was found for the *B. carinata* lineage and a rate of 1 allopolyploid per 2,380,000 ovules for the *B. napus* lineage (Table 3). The *B. juncea* homoploids did not produce any seeds, therefore no two-step allopolyploids were produced (Table 3).

Table 3: One-step and two-step allopolyploid formation rates.

S. arvensis (male) x	F1 Homoploids : Ovules (R_H)	One-step : Ovules (R_{1-step})	Two-step : Ovules | H ($R_{2-step}|H$)	Two-step : Ovules (R_{2-step})
B. carinata	1:14	1:11,700	1:1,600	1:25,600
B. juncea	1:400	NA	NA	NA
B. napus	1:11,900	NA	1:200	1:2,380,000

Unexpected DNA contents of allopolyploid offspring

Polyploidy is often described as a unidirectional process and polyploids are expected to have polyploid offspring (Meyers and Levin, 2006). However, the one-step and two-step polyploids found here produced offspring with widely varying 2C DNA contents, indicating aneuploidy.

In the *B. carinata* one-step and two-step lineages, offspring with both lower and higher than expected DNA contents were found (Figures 3-6). The variation increased

over the three generations measured (Figures 4-6). Individuals with lower DNA contents were most prevalent among the offspring of the one-step allopolyploids resulting in a significant difference in average DNA content between the one and two-step lineages by generation three (Figure 6). In the third generation, 10 of the offspring in the one-step line had lost 1 to 12 chromosomes, if a chromosome is estimated to have 0.084 pg of DNA (*B. carinata* total 2C DNA content divided by number of chromosomes). The S genome of *S. arvensis* (SS) has 9 chromosomes and the B and C genomes of *B. carinata* (BBCC) have 8 and 9 chromosomes respectively, so a loss of 12 chromosomes could represent the loss of a whole genome or more. In the same generation of the one-steps 11 individuals have a 2C DNA content that matches the expectation for an allopolyploid and 5 have gained up to 3 chromosomes (Figure 4).

In contrast, the first generation of *B. napus*-maternal two-step allopolyploids included individuals with much higher than expected DNA contents possibly representing up to 29 additional chromosomes (Figure 3). The *B. napus* x *S. arvensis* homoploid hybrids which produced the two-step allopolyploids also had a higher DNA content than expected (Figure 3).

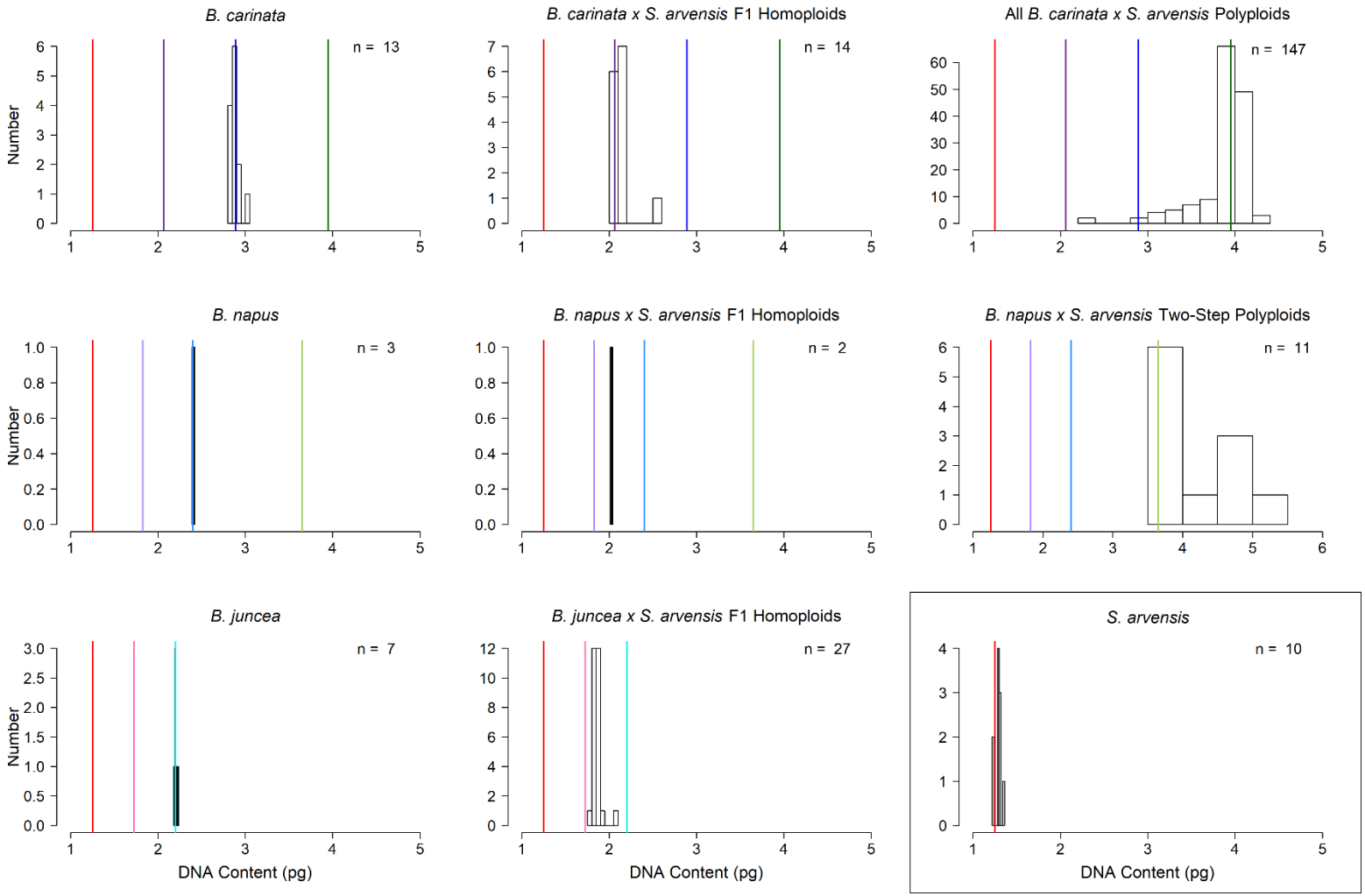


Figure 3: Histograms of parentals, homoploid hybrids and allopolyploids of all three crosses. The expected values are shown in colored lines (maternal parent blue, paternal parent red, homoploid hybrid pink or purple, allopolyploid green).

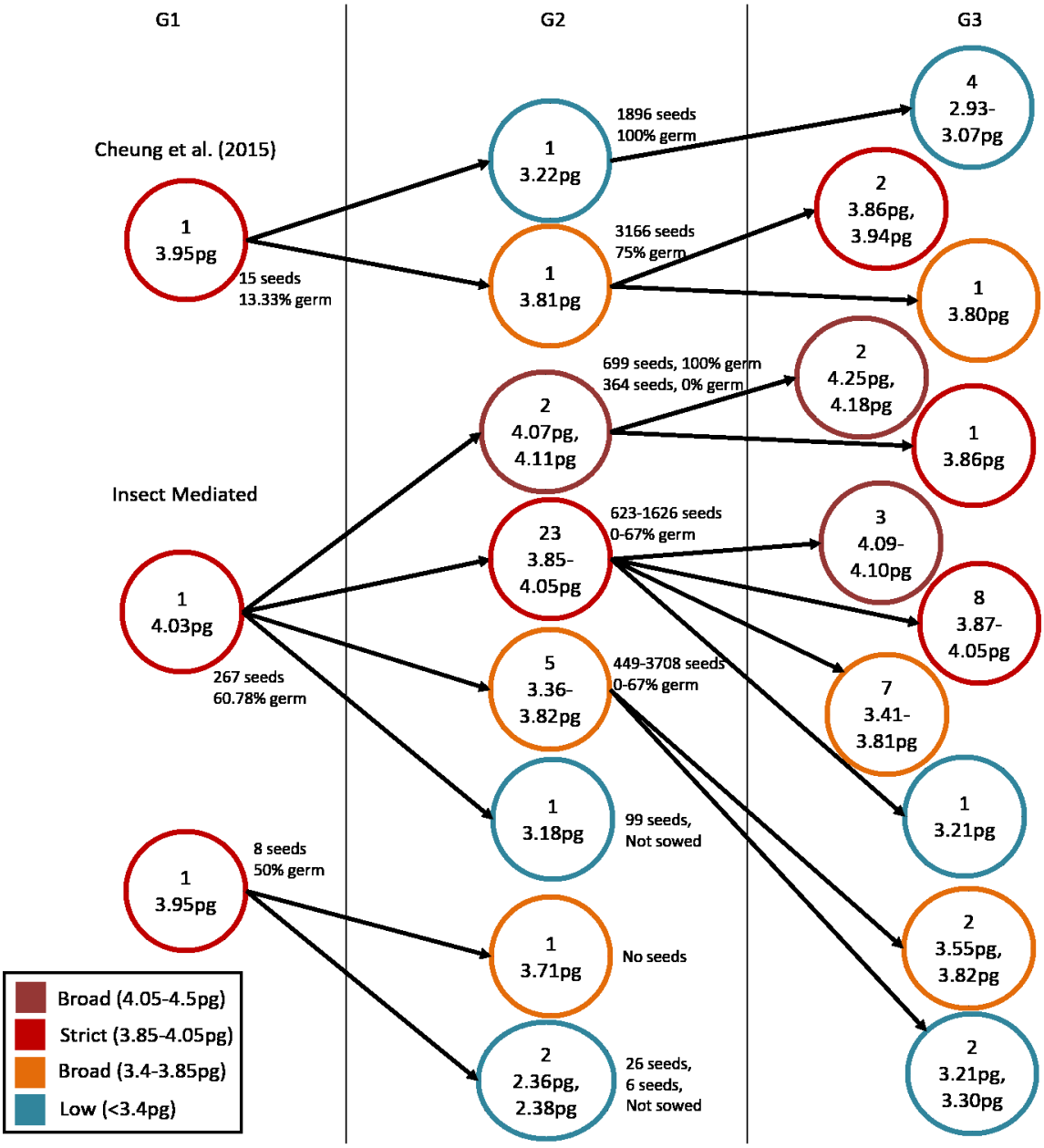


Figure 4: The DNA contents of three generations of one-step allopolyploids (B. carinata x S. arvensis). Arrows connect maternal parents to their offspring.

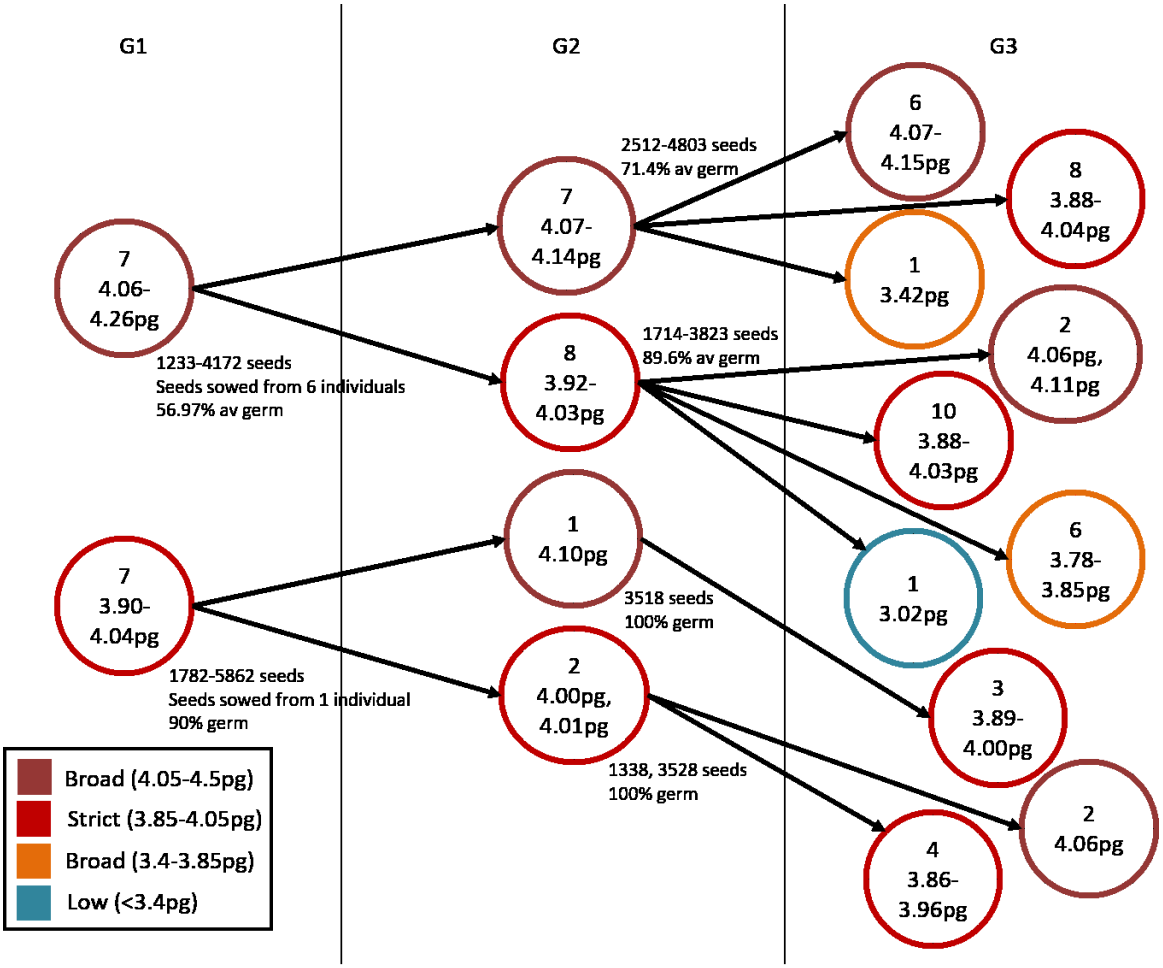
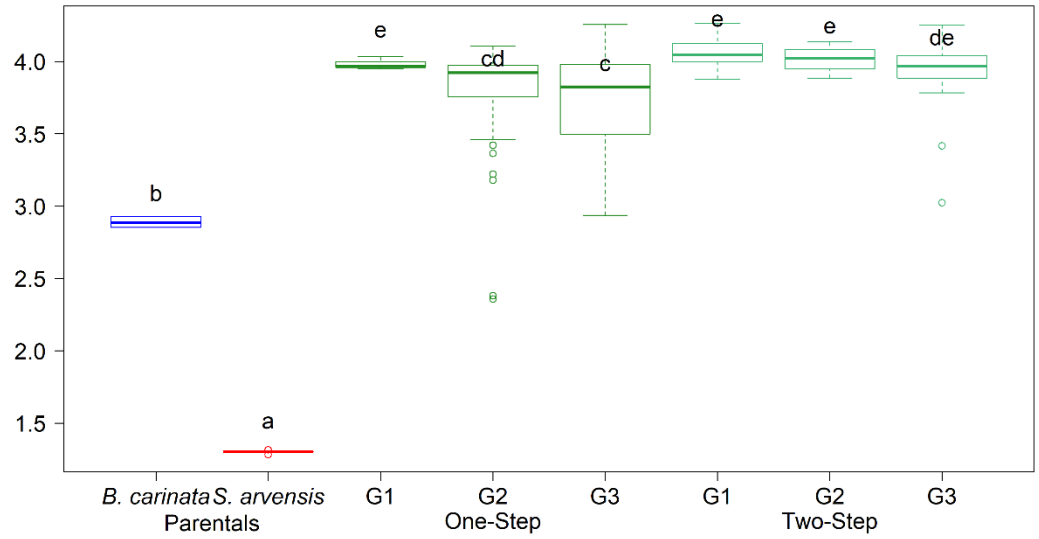


Figure 5: The DNA contents of three generations of two-step allopolyploids (B. carinata x S. arvensis). Arrows connect maternal parents to their offspring.

a)



b)

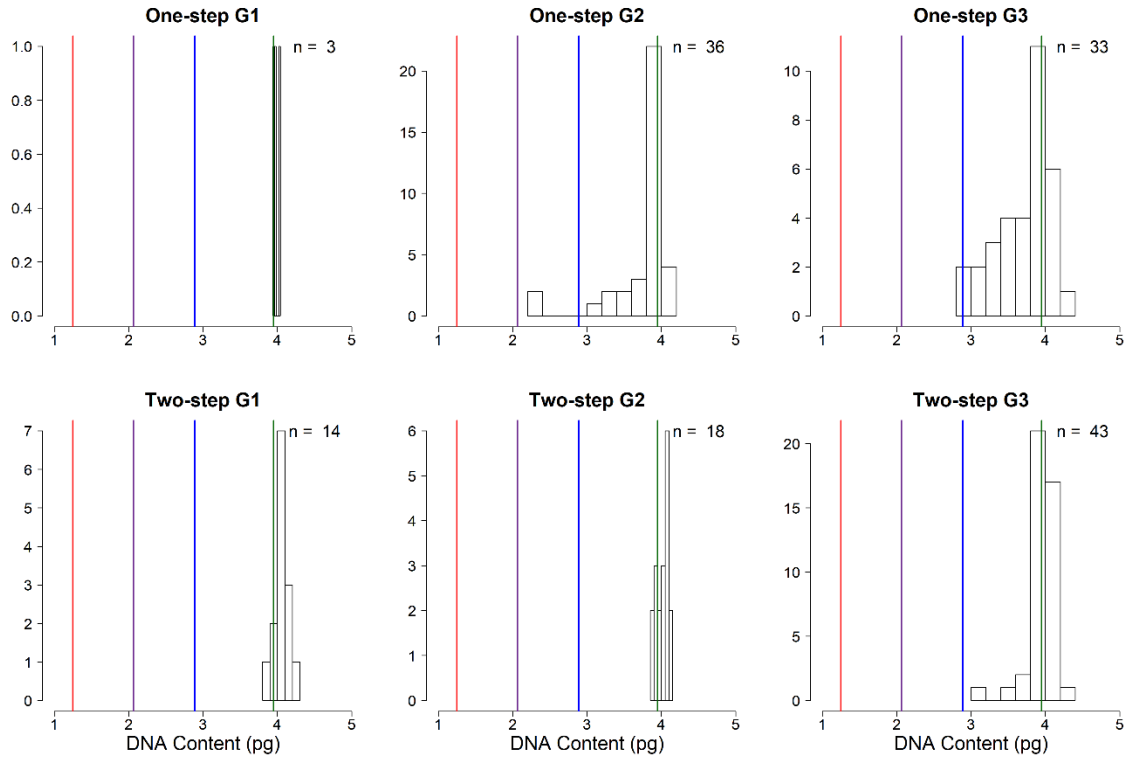


Figure 6: The 2C DNA content of B. carinata-maternal one-step and two-step allopolyploids over three generations. A) Significance values are based on a generalized

least square (GLS) model with a factor allowing for each group to have a different variance. Levels that share a letter are not statistically different. B) Vertical lines show expected values for B. carinata (blue), S. arvensis (red), homoploid hybrids (purple) and allopolyploids (green).

Reproductive isolation of polyploids

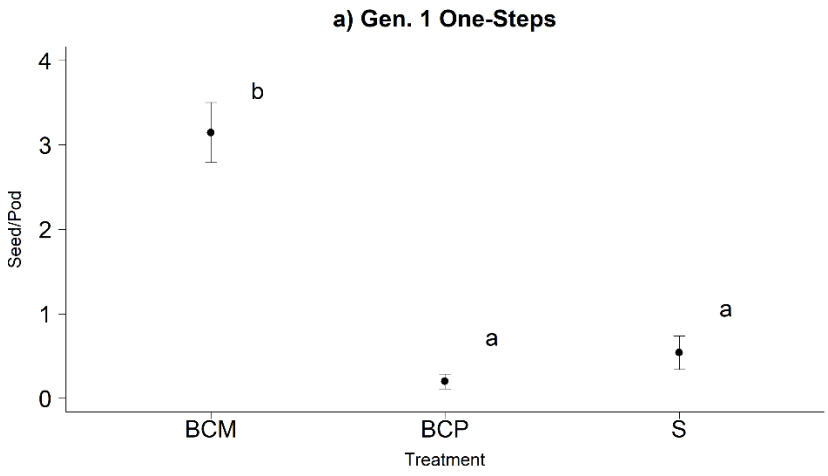
Though allopolyploids are expected to be reproductively isolated from their parents, hand crosses between *B. carinata*-maternal allopolyploids and their parentals produced seed. In the first generation, the one-step allopolypoids produced significantly more seeds with *B. carinata* (BCM) than with either *S. arvensis* (BCP) or themselves (S), and in the second generation they produced equivalent amounts of seeds when selfed (S) or when crossed with *B. carinata* (BCM) (Figure 7a and 7b). In contrast, the two-step allopolyploids did not produce a significant amount of seed with either parent (BCM or BCP) or with F2 homoploid hybrids (SIBD) (Figure 7c). The exception was one two-step individual which produced approximately three times more seed with *B. carinata* (BCM) and with F2 homoploid hybrids (SIBD) than with itself (S) and did not produce seed with any other cross (Figure 7d). A generalized least squares (GLS) model indicated that this individual produced significantly more seed with *B. carinata* and the F2 homoploids than with itself, but the more conservative Kruskal-Wallis test did not indicate significance, likely because only five flowers were self-pollinated which gives limited statistical power. This individual, which had the highest 2C DNA content of its generation, was the only two-step allopolyploid with homoploid siblings and was the only two-step allopolyploid to have 0% germination of the seeds it produced when crossed by insects with the other two-steps.

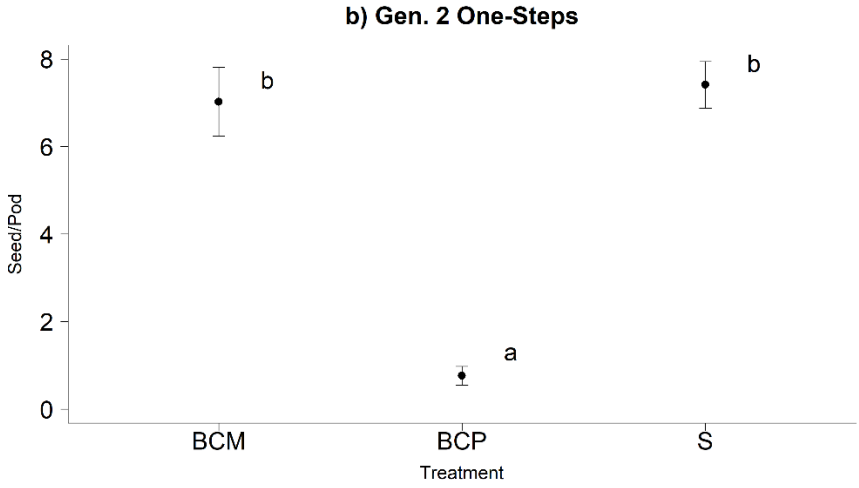
In the two-steps, selfed flowers (S) produced nearly twice as much seed as

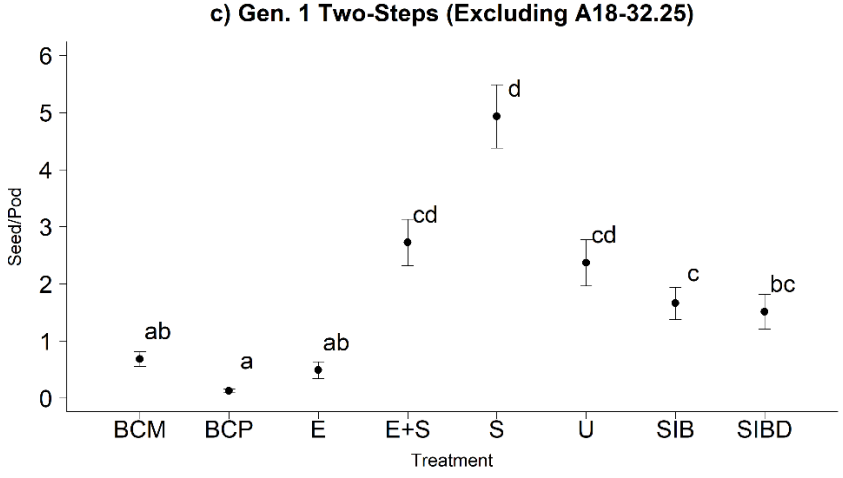
unmanipulated flowers (U), a difference which was significant in a GLS model, suggesting that the plants were not able to efficiently pollinate themselves despite self-compatibility. This may be due to orientation of the anthers. *B. carinata*, which is self-compatible, has anthers that face the stigma but in *S. arvensis*, which is self-incompatible, the anthers face away (Figure 8a and 8b). Some two-step allopolyploids had anthers slightly twisted away from the stigma (Figure 8c-f).

The autopolyploids also produced seed with their respective maternal parents. The *S. arvensis* autopolyploids only produced seed with *S. arvensis* (BCM) (Figure 8a). This was significantly different than the crosses with *B. carinata* (BCP) but was not significantly different than the other crosses due to limited power. These autopolyploids appeared self-incompatible (S, E+S). The *B. carinata* colchicine-generated autopolyploids produced significantly more seed with *B. carinata* than in the emasculation treatment (E) or when crossed with siblings (SIB1 and SIB2) and the plants were self-compatible (E+S) (Figure 8b). When crossed with autopolyploid siblings, one sibling sired seeds (SIB2) but the other did not (SIB1), however no significant difference was found between the two siblings due to low power (Figure 8b).

Hand crosses were also performed on the two individuals in the second generation of the one-step allopolyploids that had 2C DNA contents around 2.3 pg, lower than the 2C DNA content expected for *B. carinata* (2.89 pg) but still higher than the expectation for a homoploid hybrid (2.07 pg). These produced significantly more seed with *S. arvensis* (BCP) than any other treatment (Figure 9).


a) Gen. 1 One-Steps
Seed/Pod
0
1
2
3
4
b
a
a
BCM
BCP
S
Treatment


b) Gen. 2 One-Steps
Seed/Pod
0
2
4
6
8
b
a
b
BCM
BCP
S
Treatment


c) Gen. 1 Two-Steps (Excluding A18-32.25)
Seed/Pod
0
1
2
3
4
5
6
ab
a
ab
cd
d
cd
c
bc
BCM
BCP
E
E+S
S
U
SIB
SIBD
Treatment

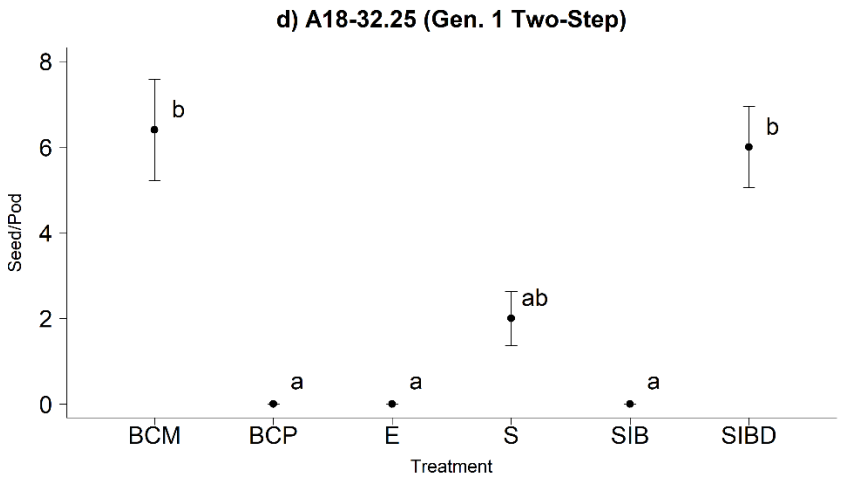
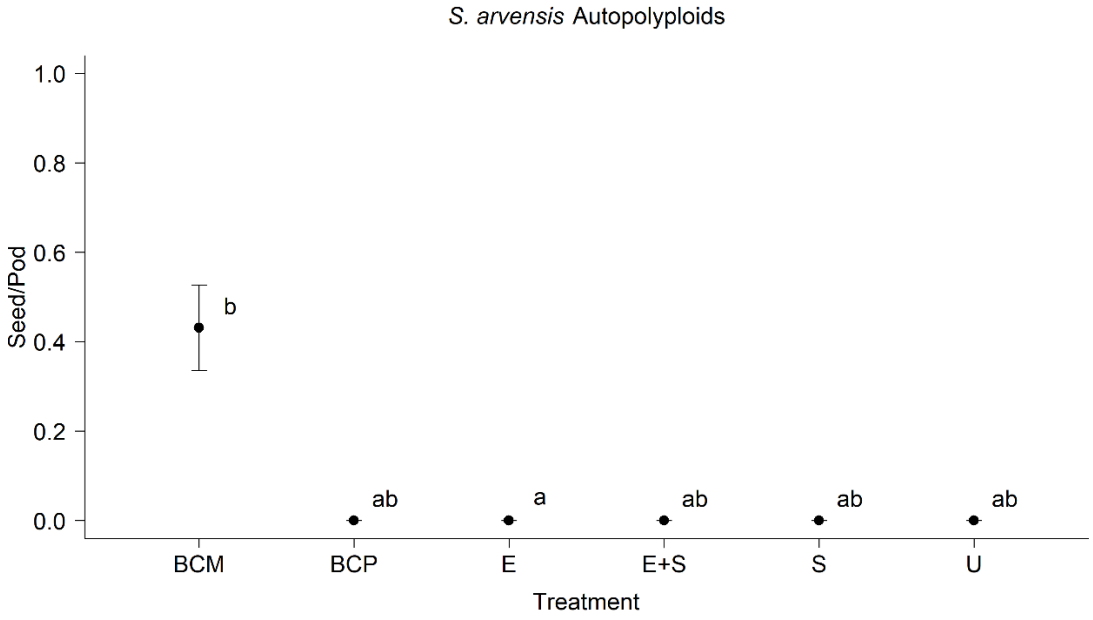


Figure 7: Seeds produced by hand crosses between allopolyploids and parentals. Significance letters are based on Kruskal-Wallis tests. Levels that share a letter are not statistically different. BCM: backcross to B. carinata (~10 flowers per individual). BCP: backcross to S. arvensis *(~10 flowers per individual). E: emasculated (~5 flowers per individual). E+S: Emasculated and selfed (~5 flowers per individual). S: Selfed (~5 flowers per individual). SIB: Crossed with polyploid sibling (~5 flowers per individual). SIBD: Crossed with homoploid sibling (~5 flowers per individual). U: unmanipulated (~5 flowers per individual). (a) and (b) show data from two individuals each, (c) shows data from 13 individuals, and (d) shows data from one individual.*

a)



b)

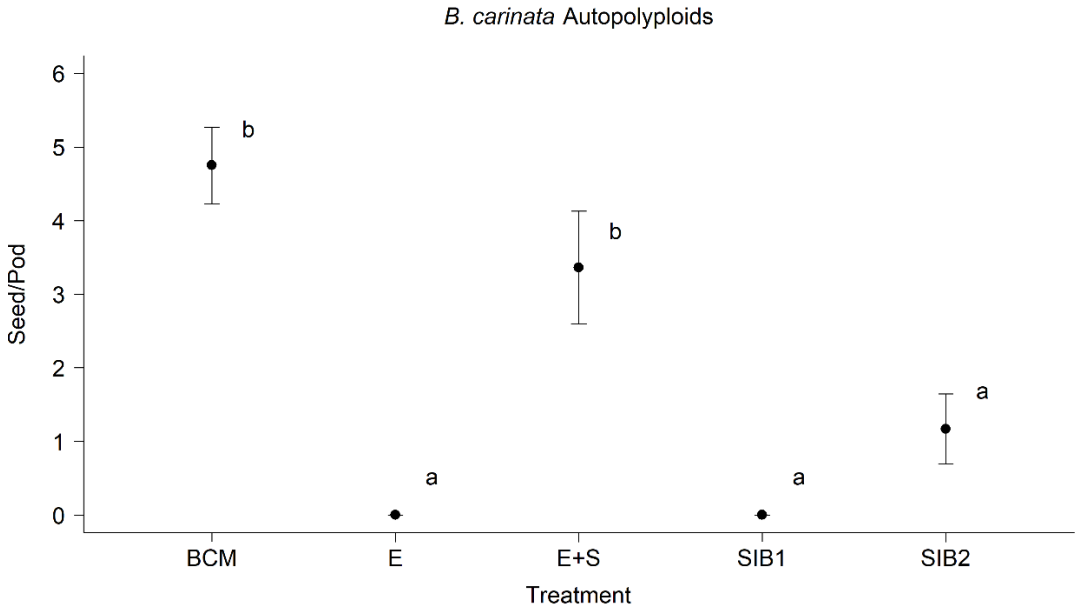


Figure 8: Seeds produced by hand crosses between autopolyploids and parentals. Significance letters are based on Kruskal-Wallis tests. Levels that share a letter are not statistically different. BCM: backcross to maternal parent (10-20 flowers per individual). BCP: backcross to B. carinata *(~20 flowers per individual). E: emasculated (~5 flowers per individual). E+S: Emasculated and selfed (~5 flowers per individual). S: Selfed (~5 flowers per individual). SIB1 or SIB2: Crossed with polyploid sibling (~5 flowers per individual). U: unmanipulated (~5 flowers per individual). (a) shows data from three individuals and (b) shows data from two.*

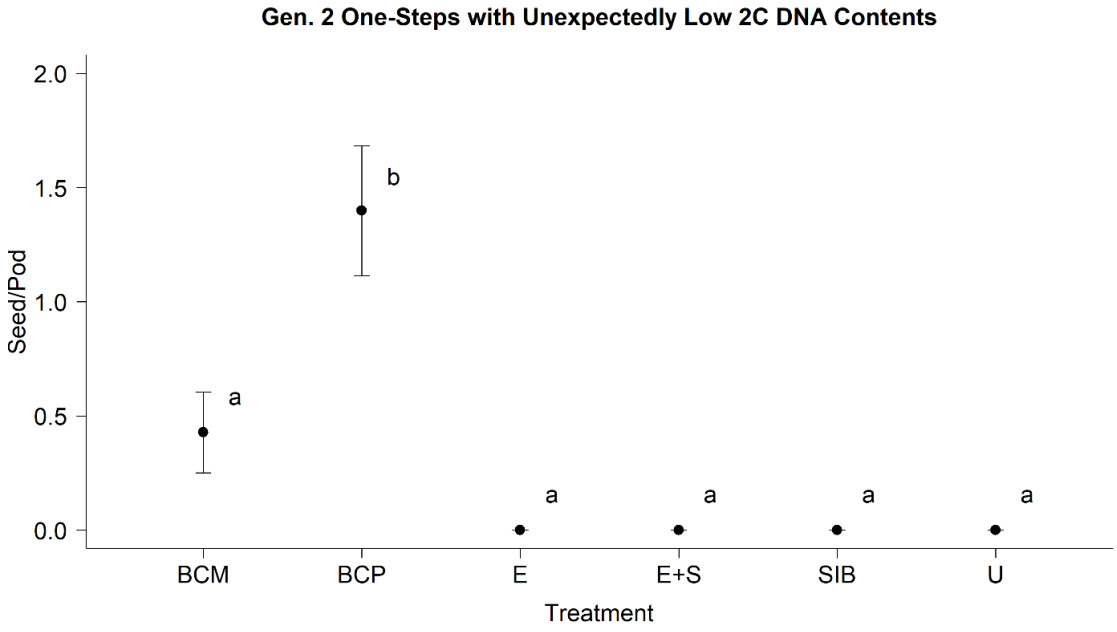


Figure 9: Seeds produced by hand crosses between the two second-generation one-step allopolyploids with unexpectedly low 2C DNA contents (~2.3 pg) and their parentals. Significance letters are based on Kruskal-Wallis tests. Levels that share a letter are not statistically different. BCM: backcross to B. carinata (~10 flo wers per individual). BCP: backcross to S. arvensis (~10 flowers per individual). E: emasculated (~5 flowers per individual). E+S: Emasculated and selfed (~5 flowers per individual). SIB: Crossed with the other unexpectedly low second-generation one-step (~5 flowers per individual). U: unmanipulated (~5 flowers per individual). Data from two individuals are included here.

Figure 10: Anther orientation in B. carinata (a), S. arvensis (b), first-generation two-step allopolyploids (c, d) and second-generation two-step allopolyploids (e, f).

Fertility in one-step and two-step allopolyploids over three generations

Not all descendants of the allopolyploids had the DNA content expected for allopolyploids, as defined as whole genome duplication. As a result, the descendants were categorized into allopolyploids and aneuploids before analysis. Strict allopolyploids were those whose DNA content was within one chromosome (estimated 0.084 pg) of the expected value (3.95 pg) ±0.02 pg to account for error. Therefore, only individuals with DNA content between 3.85 and 4.05 pg were considered strict allopolyploids (Figure 11). Broad allopolyploids were defined as any individual descended from an allopolyploid, regardless of DNA content, and therefore included a range of aneuploids.

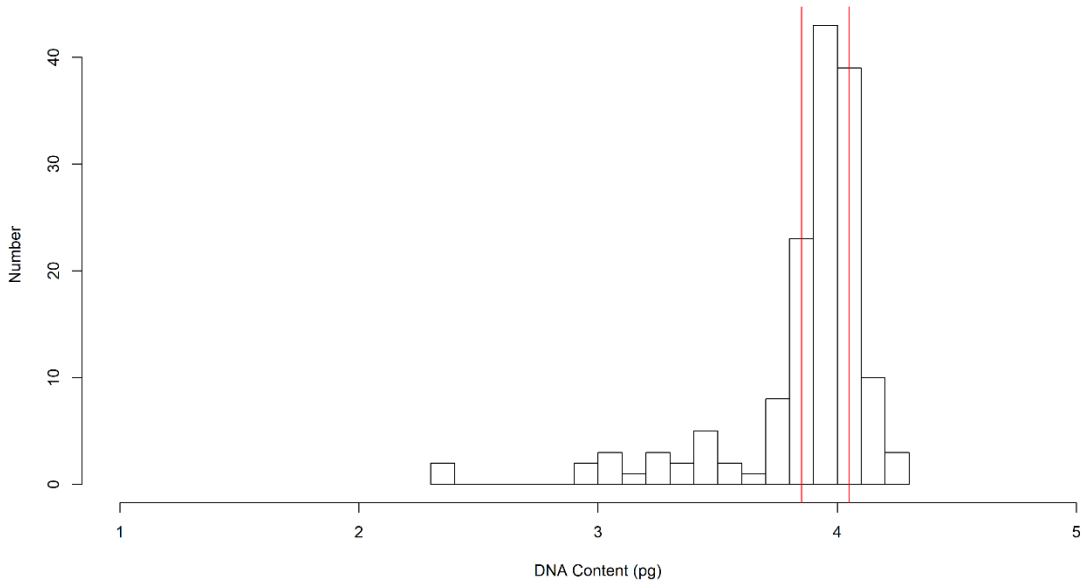


Figure 11: Histogram of all one-step and two-step B. carinata-maternal allopolyploids. The red lines indicate the boundaries of the strict definition of allopolyploidy.

In the first generation of both strict and broad allopolyploids, pollen viability was significantly higher in the two-steps than in the one-steps. However, the variation in pollen fertility increased in generation two and three for both one-steps and two-steps

resulting in no significant difference between the two lineages (Figure 12).

Seed counts were also significantly higher in the two-steps than the one-steps in the first generation. One-step allopolyploids saw a significant increase in seed count from the first to the second generation but remained lower than the second-generation two-steps, and in the third generation the seed count of the one-steps decreased to be statistically similar to the original level (Figure 12). The two-step allopolyploids maintained a seed count equivalent to that of *S. arvensis* in their first two generations, but like the one-steps showed a significant decrease in the third generation (Figure 12).

In the one-step allopolyploids, the individuals with the highest pollen viability and seed set were aneuploids rather than strict allopolyploids. No allopolyploids achieved equivalent pollen viability to the parentals except the third generation of the strict two-steps. Two-steps had equivalent seed set to *S. arvensis* in the first two generations (Figure 12).

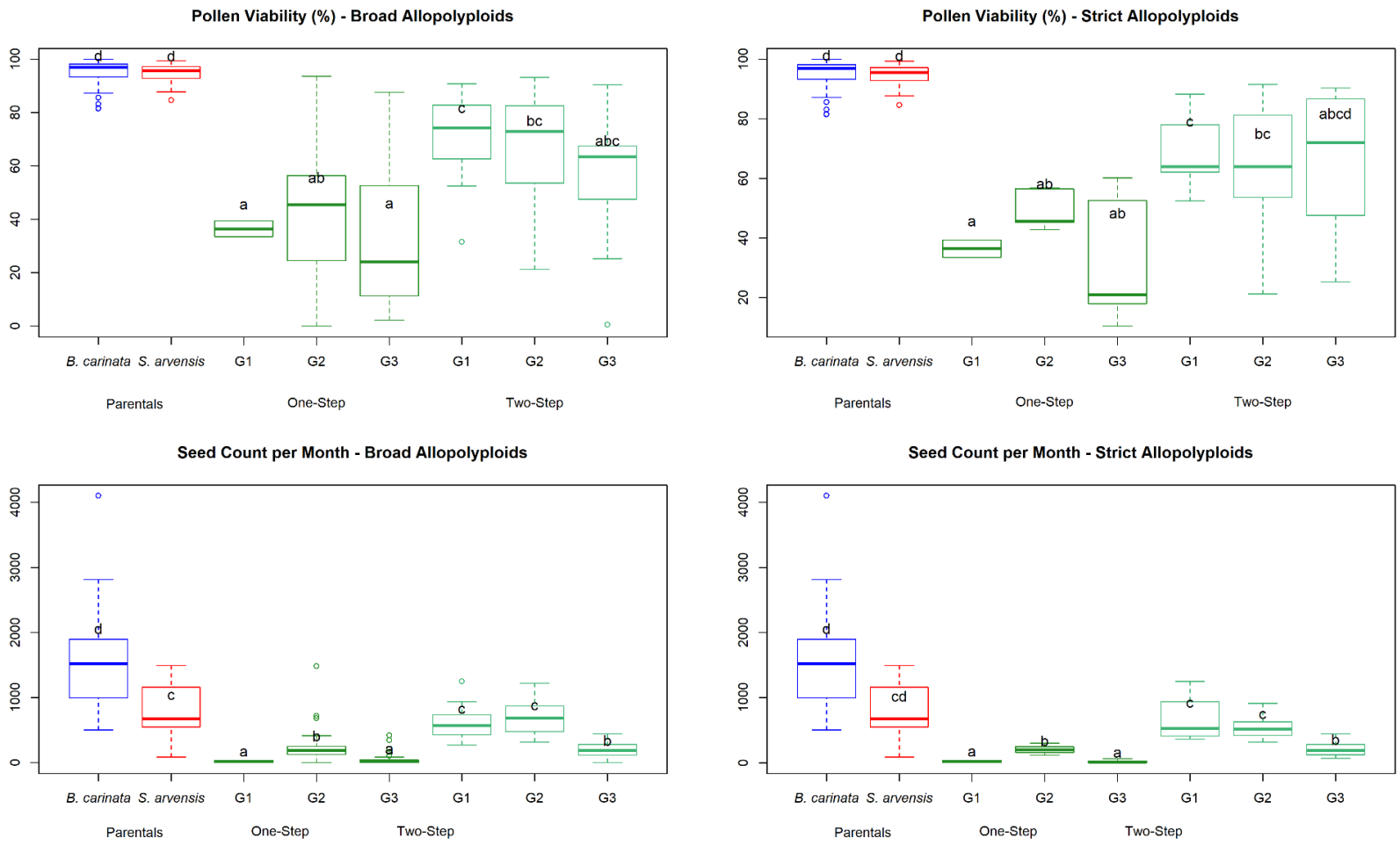


Figure 12: Pollen viability and seed production for three generations of one-step and two-step allopolyploids, compared to parentals. Significance indications are based on a generalized least square (GLS) model with a factor allowing for each group to have a different variance. Levels that share a letter are not statistically different.

Correlation between fertility and DNA content

To determine whether 2C DNA content correlated with fertility, as represented by pollen viability and seed set, all *B. carinata*-maternal allopolyploids (strictly or broadly defined) were analysed. The data were not normally distributed so Spearman's rank correlation was calculated and found to be significant and positive indicating that fertility showed an increase with DNA content. However, the data was noisy and several individuals with low 2C DNA contents had high fertility (Figure 13).

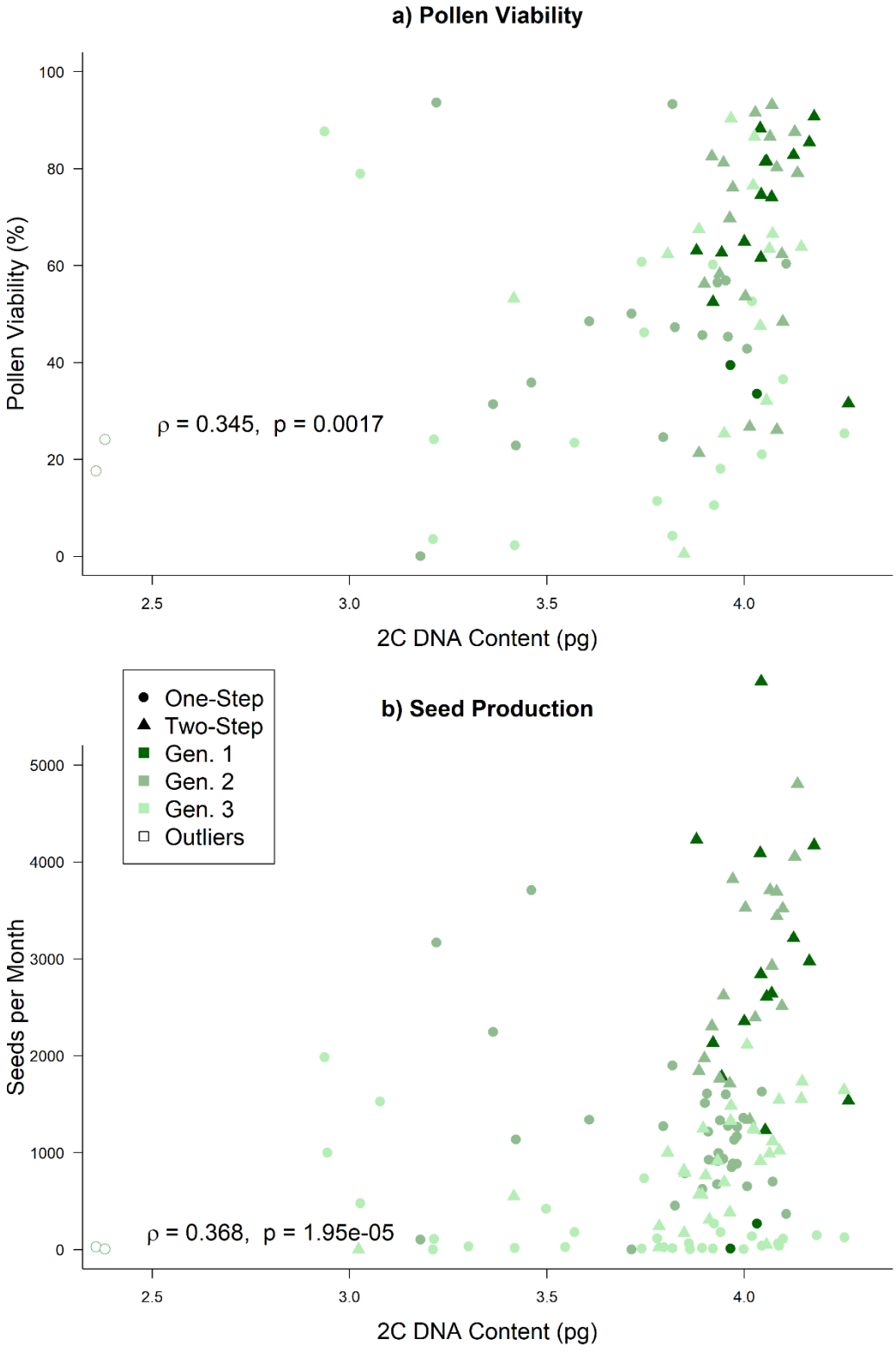


Figure 13: Correlations between 2C DNA content and fertility markers. Spearman rank coefficients are shown.

Morphological characteristics of different types of allopolyploids

ANOVAs indicated that strict allopolyploids had significant morphological differences from aneuploids descended from allopolyploids (for example, $p = 0.10$ for stomata size, $p < 0.01$ for days to flower, $p = 0.03$ for trichome density). Therefore, further analysis included only strict allopolyploids. Detailing the morphological differences between aneuploids and strict allopolyploids was beyond the scope of this thesis.

The allopolyploids were divided by type (one-step or two-step) and by generation and compared using a Linear Discriminate Analysis (LDA) (Figure 14). The first two axes of the LDA accounted for 93.9% of the variation in the data, with 72.0% on the first axis and 21.9% on the second. Beak length, the percent of the gynoecium which was beak, petal length, plant height at flowering, and trichome density accounted for the majority of the variation. A MANOVA was also run on the data ($F_{25,140} = 3.448$, $p < 0.001$). While the points overlapped, one-steps and two-steps were grouped separately and both types changed in a similar way over the three generations (Figure 14). ANOVAs were used to test each morphological trait individually for effects of type, generation, or an interaction between both, except for two binary traits which were tested using chi-squared (Table 4). Of the 23 morphological traits tested, nine were not affected by type or generation at all (stem angle, trichome length, stomata length, beak length, and all the markers of flower proportions), one was affected only by type (trichome density), seven only by generation (days to flower, purple joints, and some flower size measurements), and five were affected by both type and generation (pedicle length and the remaining flower size measurements) (Table 4). Only one was affected by type, generation, and an

interaction between the two (plant height at flowering) (Table 4).

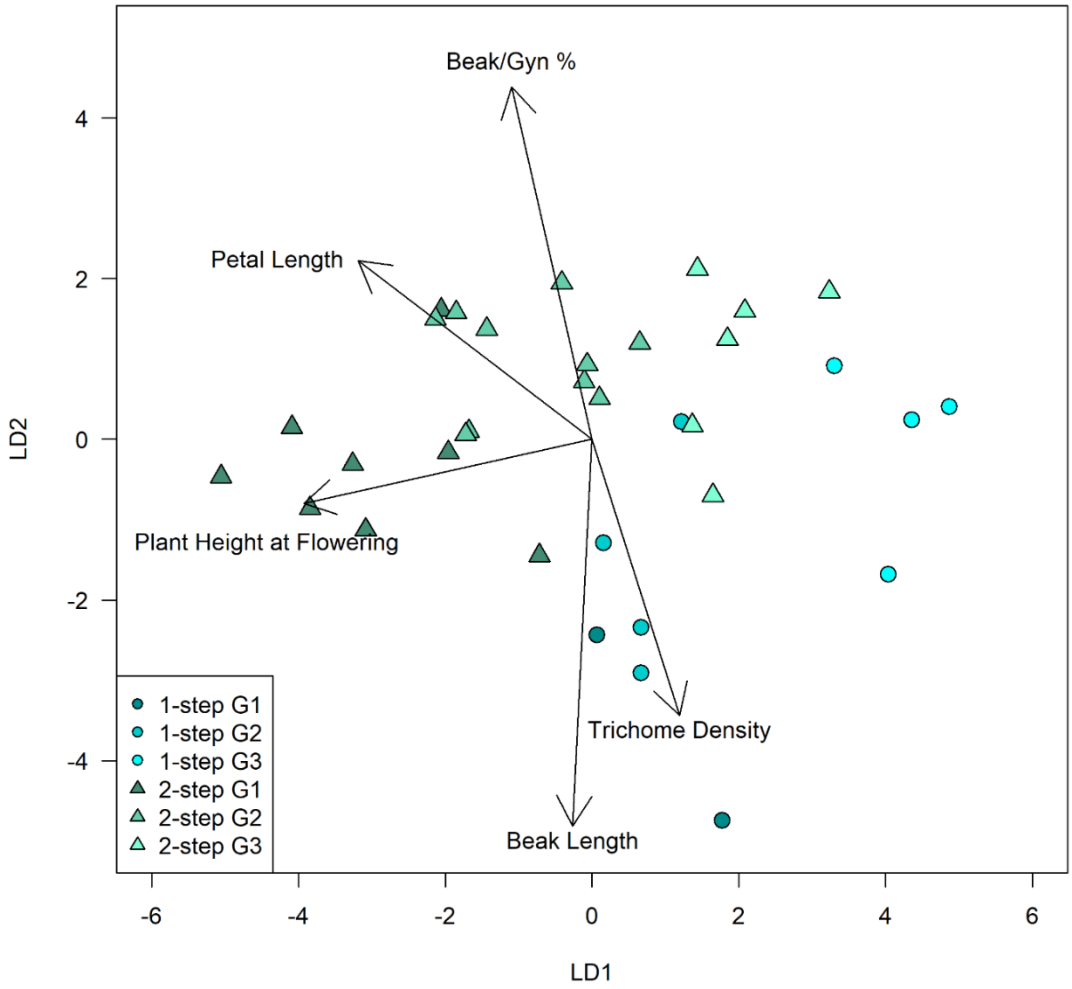


Figure 14: Differentiation between types and generations of B. carinata-maternal allopolyploids using linear discriminate analysis (LDA).

Table 4: ANOVAs and chi-squared tests examining the effect of step and generation on morphological traits. Significant p-values are bolded.

Trait	Type	Generation	Type-Gen. Interaction

	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p
Trichome density	1,31	18.23	**<0.01**	2,31	0.54	0.75	2,31	0.54	0.75
Days to flower	1,55	1.32	0.43	2,55	8.88	**<0.01**	2,55	1.98	0.29
Gynoecium length	1,29	1.91	0.34	2,29	17.98	**<0.01**	2,29	1.54	0.41
Flower length	1,30	1.12	0.48	2,30	3.99	**0.08**	2,30	1.227	0.48
Claw length	1,29	3.53	0.17	2,29	8.97	**<0.01**	2,29	0.44	0.78
Sepal length	1,29	5.82	0.07	2,29	7.61	**0.01**	2,29	2.79	0.18
Petal length	1,29	7.50	**0.04**	2,29	13.18	**<0.01**	2,29	0.01	0.99
Petal width	1,29	13.14	**<0.01**	2,29	9.088	**<0.01**	2,29	0.50	0.77
Flower display size	1,30	8.41	**0.03**	2,30	9.91	**<0.01**	2,30	0.09	0.95
Anther length	1,29	9.23	**0.02**	2,29	12.15	**<0.01**	2,29	0.41	0.78
Pedicle length	1,29	8.49	**0.03**	2,29	9.01	**<0.01**	2,29	1.33	0.46
Plant height at flowering	1,56	6.75	**0.05**	2,56	51.78	**<0.01**	2,56	5.27	**0.03**
Beak/Gyn. %	1,29	2.50	0.25	2,29	0.07	0.96	2,29	0.34	0.82
Claw/Petal %	1,29	0.36	0.72	2,29	0.43	0.78	2,29	1.24	0.48
Trichome length	1,30	2.54	0.25	2,30	1.05	0.54	2,30	1.20	0.49
Beak length	1,29	4.87	0.10	2,29	4.84	0.05	2,29	1.01	0.55
Stomata length	1,30	6.31	0.06	2,30	0.74	0.65	2,30	0.39	0.79
Days to germinate	1,72	<0.01	0.97	2,72	2.63	0.18	2,72	0.46	0.78
Stem angle	1,30	0.08	0.87	2,30	0.19	0.90	2,30	0.48	0.77
Petal W/H	1,29	1.36	0.43	2,29	0.34	0.82	2,29	0.59	0.73

Anther-Gynoecium	1,29	2.04	0.31	2,29	1.11	0.52	2,29	2.13	0.27
	df	χ^2	p	df	χ^2	p	df	χ^2	p
Purple joints	1	0.23	0.64	2	16.36	**<0.01**	NA		
Purple dots	1	1.42	0.23	2	13.20	**<0.01**			

The contribution of hybridity and polyploidy to allopolyploid morphology

To examine whether hybridity or polyploidy explained the variation in the allopolyploids, an LDA was used to compare the allopolyploids to parentals, homoploid hybrids, and autopolyploids (Figure 15). The first two axes of the LDA accounted for 88.2% of the variation in the data, with 65.5% on the first axis and 22.7% on the second. Trichome length, trichome density, petal length, the percent of petal that is claw, and plant height at flowering accounted for the majority of the variation. A MANOVA was also run on the data ($F_{55,485} = 8.3107$, $p < 0.001$). The parentals, *B. carinata* and *S. arvensis*, were clearly separated with their autopolyploids nearby and partially overlapping. The homoploid hybrids were located between the two parentals with the *S. arvensis*-maternal homoploids closer to *S. arvensis*. The allopolyploids were spread between *B. carinata* and the *B. carinata*-maternal homoploid hybrids, overlapping both. The two-step allopolyploids overlapped with *B. carinata* in the first generation but by the third generation they were closely overlapping the homoploids instead. The one-step allopolyploids showed less change over time, remaining close to the homoploids and not overlapping with *B. carinata*.

ANOVAs were used to compare the effect of hybridity and polyploidy on

individual morphological traits (Table 5-8). Overall, polyploidy explained variation in 14 traits, hybridity explained variation in 11 traits,and the hybridity-polyploidy interaction explained variation in 5 traits. Polyploidy principally affected flower proportions (beak/gyn %, petal width/height, etc.) and size (sepal length, petal width, etc.), stomata length, and some developmental markers (days to germinate). Hybridity principally affected trichome density, purple dots on anthers, some aspects of flower proportion (claw/petal %, flower display size), beak length, and plant height at flowering. Some of the traits affected by hybridity were traits inherited from only one parent; for instance, *S. arvensis* has trichomes while *B. carinata* does not, and *B. carinata* has purple dots on its anthers while *S. arvensis* does not. The interaction of hybridity and polyploidy in some cases affected stomata length, flower size (petal length, sepal length, etc.), and plant height at flowering. Variation in some traits, including stem angle, days to flower, and pedicle length, was not explained by either hybridity or polyploidy.

Overall, polyploidy explained the variation of fewer traits in one-steps than in two-steps (Table 7 and 8). The number of traits affected by polyploidy in both types decreased between generation one and three while the number of traits affected by hybridity increased (Table 6-8). The decreasing effect of polyploidy and increasing affect of hybridity aligns with the shift of the allopolyploids towards the homoploid hybrids in the LDA (Figure 15). The lesser effect of polyploidy on the one-steps may explain why the one-steps were never located as closely to the autopolyploids as the two-steps were (Figure 15) and it also may explain why they have smaller stomata, a trait associated with polyploidy, and denser trichomes, which is associated with hybridity in this system (Figures 16 and 17).

For specific traits of interest, a generalized least squares (GLS) model was used to identify differences between groups (Figure 16 to 17). Stomata length, as expected, increased in polyploids and was correlated with DNA content ($p < 0.01$) (Figure 16). However, this increase was only significant in the first generation of two-step allopolyploids and in *B. carinata* autopolyploids. Low sample size may have contributed to the lack of significance for the *S. arvensis* autopolyploids, but the one-step allopolyploids have a similar mean stomata length to *B. carinata*. Variation in trichome density and beak length was explained by hybridity (Figure 17). Homoploids and allopolyploids have intermediate trichome density to their parents and one-steps have more trichomes than two-steps, though again low sample size made it difficult to find significance. The trichome density of the *S. arvensis* autopolyploids is likely caused by their descent from *S. arvensis* and not from ploidy. Homoploids and allopolyploids had longer beaks than their parents, though in the third generation the beak length decreased (Figure 17). This pattern was found in most flower size markers. Petal width, for example, was greater in autopolyploids and allopolyploids, but decreased in the third generation, significantly in the two-steps (Figure 18). This decrease is unexpected as larger flower size is associated with polyploidy and petal width correlated with DNA content ($p < 0.01$). While flower size decreased in the third generation of allopolyploids, however, flower proportions such as petal width/height did not change (Figure 18).

Delayed development is also associated with polyploidy. However, polyploidy had no effect on days to flower and its effect on days to germinate is likely due to outliers (Figure 19). Plant height at flowering was not altered significantly by autopolyploidy (Figure 19). The plant height at flowering of allopolyploids was intermediate to their

parentals in most cases and decreased significantly over generations. This decrease is likely why its variation was explained by hybridity instead of polyploidy in the third generation of both types of allopolyploids.

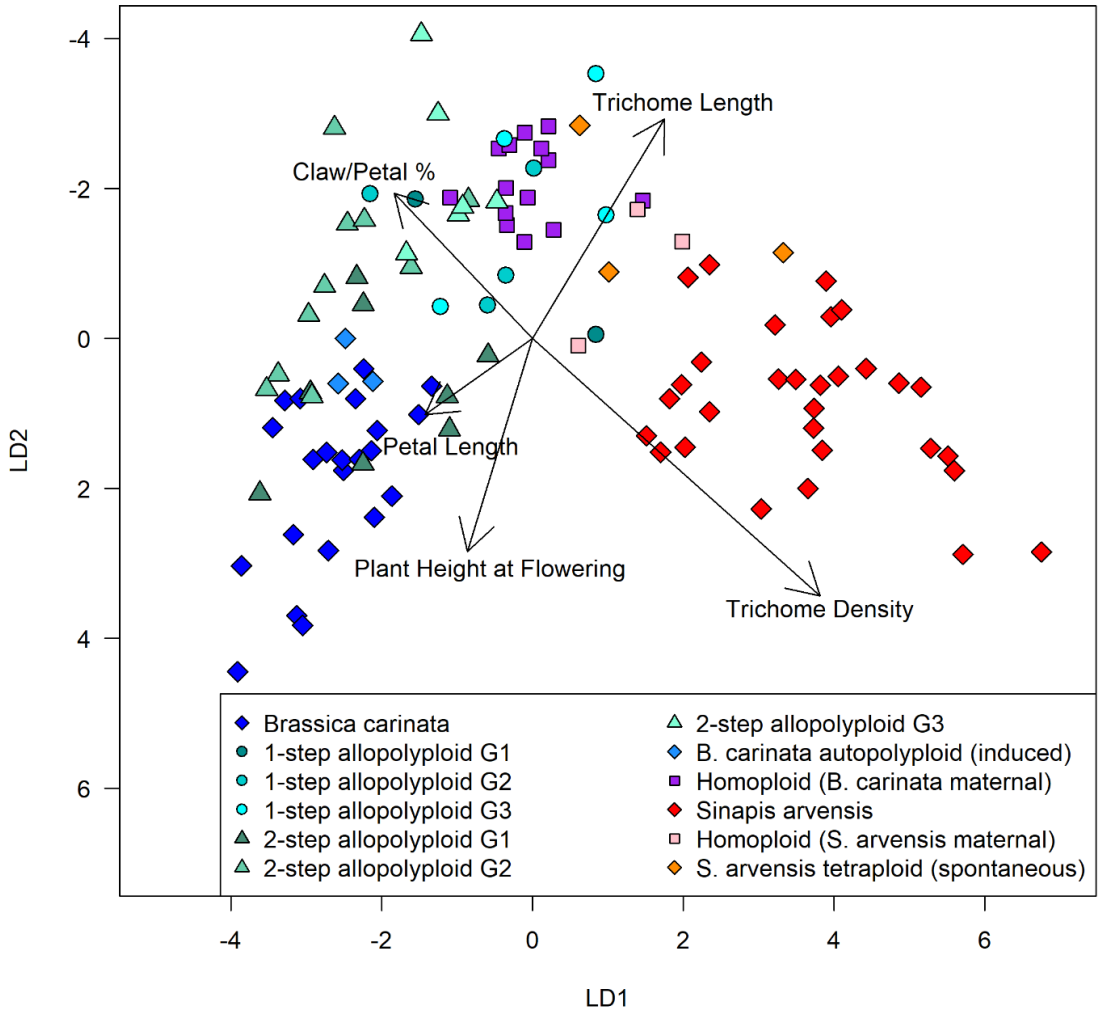


Figure 15: Differentiation of parentals, homoploids, allopolyploids, and autopolyploids using linear discriminate analysis (LDA).

Table 5: ANOVAs examining the effect of hybridity and polyploidy on morphological traits that were not affected by type or generation. Significant p-values are bolded.

Trait	Hybridity			Polyploidy			Interaction		
	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p
Petal W/H	1,115	5.45	0.07	1,115	8.087	**0.02**	1,115	0.56	0.62
Beak/Gyn. %	1,115	1.77	0.35	1,115	8.18	**0.02**	1,115	0.20	0.78
Days to germinate	1,155	3.09	0.18	1,155	7.31	**0.03**	1,155	0.32	0.74
Claw/Petal %	1,115	51.05	**<0.01**	1,115	15.303	**<0.01**	1,115	0.41	0.69
Anther-Gynoecium	1,69	0.78	0.55	1,69	5.98	0.06	1,69	4.12	0.12
Stem angle	1,116	0.69	0.57	1,116	0.88	0.52	1,116	0.06	0.88
Trichome length	1,116	5.08	0.07	1,116	3.92	0.12	1,116	2.36	0.26

Table 6: ANOVAs examining the effect of hybridity and polyploidy on morphological traits that were only affected by type. Significant p-values are bolded.

Trait	One-step allopolyploids									Two-step allopolyploids								
	Hybridity			Polyploidy			Interaction			Hybridity			Polyploidy			Interaction		
	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p
Stomata length	1,92	1.70	0.36	1,92	45.38	**<0.01**	1,92	7.31	**0.03**	1,104	25.56	**<0.01**	1,104	74.63	**<0.01**	1,104	0.65	0.59
Trichome density	1,93	8.81	**0.02**	1,93	0.18	0.78	1,93	3.01	0.19	1,104	23.22	**<0.01**	1,104	2.33	0.26	1,104	0.99	0.49

Table 7: ANOVAs examining the effect of hybridity and polyploidy on morphological traits that were only affected by generation. Significant p-values are bolded.

Trait	Generation 1									Generation 3								
	Hybridity			Polyploidy			Interaction			Hybridity			Polyploidy			Interaction		
	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p
Gynoecium length	1,90	11.30	**0.01**	1,90	24.43	**<0.01**	1,90	0.74	0.56	1,91	0.02	0.94	1,91	1.63	0.38	1,91	6.24	0.05
Flower length	1,90	8.61	**0.02**	1,90	13.79	**<0.01**	1,90	1.05	0.48	1,91	3.79	0.13	1,91	6.52	**0.05**	1,91	0.06	0.88
Claw length	1,90	3.90	0.13	1,90	11.48	**0.01**	1,90	6.21	0.05	1,91	0.01	0.97	1,91	1.53	0.40	1,91	0.20	0.78
Days to flower	1,84	3.94	0.13	1,84	3.27	0.17	1,84	3.47	0.16	1,110	0.13	0.82	1,110	0.02	0.93	1,110	0.01	0.96
	df	χ^2	p	df	χ^2	p	df	χ^2	p	df	χ^2	p	df	χ^2	p	df	χ^2	p

Purple joints	1	<0.01	1	1	0.06	0.81	NA	1	6.79	**0.01**	1	10.47	**<0.01**	NA
Purple dots	1	15.28	**<0.01**	1	2.76	0.10		1	5.47	**0.02**	1	<0.01	1	

Table 8: ANOVAs examining the effect of hybridity and polyploidy on morphological traits that were affected by both type and generation. Significant p-values are bolded.

Trait	Generation 1																	
	One-step									Two-step								
	Hybridity			Polyploidy			Interaction			Hybridity			Polyploidy			Interaction		
	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p
Beak length	1,82	11.77	**0.01**	1,82	0.71	0.57	1,82	0.08	0.87	1,88	17.71	**<0.01**	1,88	1.22	0.45	1,88	0.06	0.88
Flower display size	1,82	5.50	0.07	1,82	3.32	0.17	1,82	<0.01	0.97	1,88	1.30	0.43	1,88	10.47	**0.01**	1,88	1.09	0.48
Plant height at flowering	1,76	2.86	0.20	1,76	0.24	0.77	1,76	1.33	0.43	1,81	0.12	0.83	1,81	4.92	0.08	1,81	13.49	**<0.01**
Sepal Length	1,82	0.64	0.59	1,82	9.10	**0.02**	1,82	7.67	**0.03**	1,88	5.35	0.07	1,88	17.33	**<0.01**	1,88	5.95	0.05
Petal width	1,82	1.17	0.46	1,82	5.31	0.07	1,82	1.47	0.41	1,88	0.51	0.64	1,88	16.70	**<0.01**	1,88	4.73	0.09
Anther length	1,82	20.62	**<0.01**	1,82	9.43	**0.02**	1,82	5.85	0.06	1,88	5.01	0.08	1,88	34.04	**<0.01**	1,88	14.32	**<0.01**
Petal length	1,82	2.48	0.25	1,82	0.39	0.70	1,82	2.75	0.21	1,88	0.02	0.93	1,88	7.69	**0.03**	1,88	10.09	**0.01**
Pedicle length	1,82	0.19	0.78	1,82	1.52	0.40	1,82	0.75	0.56	1,88	0.57	0.62	1,88	0.52	0.64	1,88	2.12	0.29

Trait	Generation 3																	
	One-step									Two-step								
	Hybridity			Polyploidy			Interaction			Hybridity			Polyploidy			Interaction		
	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p	$df_{n,d}$	F	p
Beak length	1,85	4.10	0.12	1,85	1.81	0.34	1,85	6.26	0.05	1,86	9.65	**0.01**	1,86	0.01	0.96	1,86	0.99	0.49
Flower display size	1,85	13.24	**<0.01**	1,85	0.01	0.95	1,85	6.19	0.05	1,86	7.62	**0.03**	1,86	1.31	0.43	1,86	1.01	0.49
Plant height at flowering	1,84	9.69	**0.01**	1,84	2.88	0.21	1,84	<0.01	0.97	1,98	16.38	**<0.01**	1,98	3.23	0.17	1,98	0.02	0.93
Sepal Length	1,85	0.02	0.94	1,85	1.67	0.37	1,85	0.12	0.83	1,86	0.22	0.78	1,86	4.22	0.11	1,86	0.21	0.78
Petal width	1,85	4.07	0.12	1,85	0.77	0.55	1,85	1.33	0.43	1,86	0.82	0.54	1,86	5.57	0.06	1,86	0.25	0.77
Anther length	1,85	23.61	**<0.01**	1,85	6.86	**0.04**	1,85	0.71	0.57	1,86	18.62	**<0.01**	1,86	12.14	**0.01**	1,86	2.73	0.21
Petal length	1,85	6.30	0.05	1,85	0.39	0.70	1,85	0.09	0.86	1,86	3.91	0.13	1,86	0.04	0.91	1,86	0.42	0.69
Pedicle length	1,84	<0.01	0.97	1,84	2.97	0.19	1,84	0.07	0.88	1,86	0.21	0.78	1,86	1.26	0.44	1,86	1.13	0.47

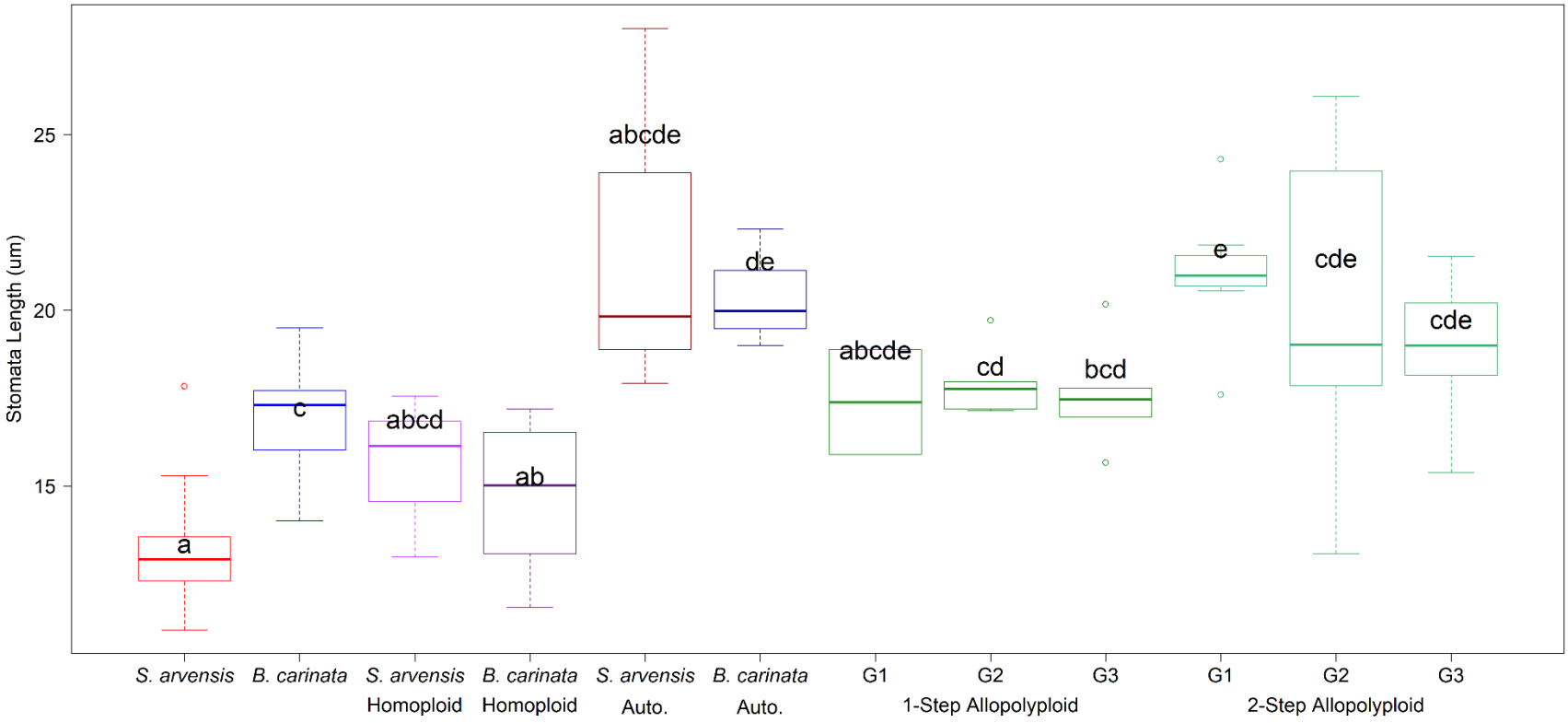


Figure 16: Stomata length of allopolyploids, autopolyploids, homoploids, and parentals, showing the effect of polyploidy. Significance letters are based on a generalized least square (GLS) model with a factor allowing for each group to have a different variance. Levels that share a letter are not statistically different.

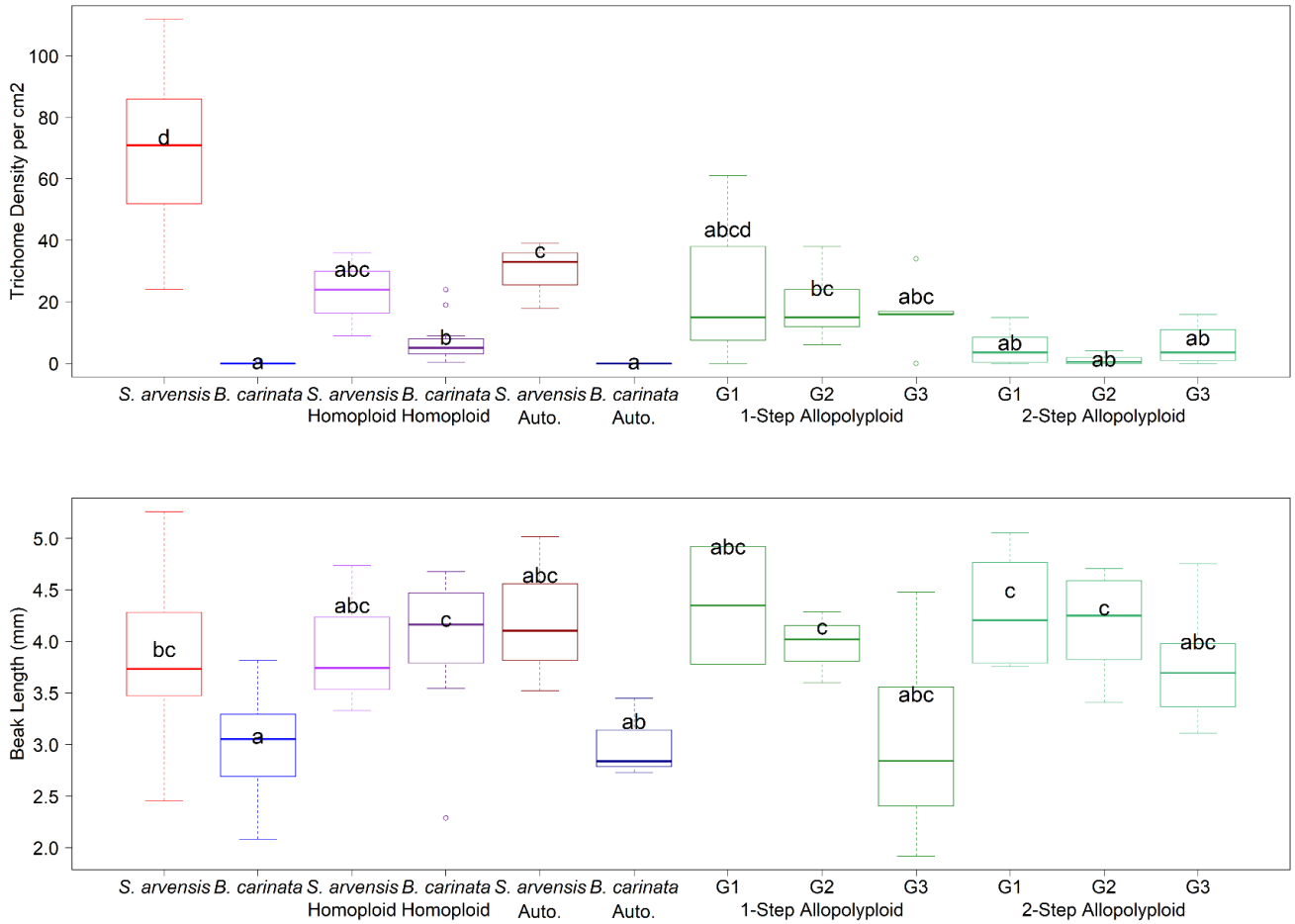


Figure 17: Trichome density and beak length of allopolyploids, autopolyploids, homoploids, and parentals, showing the effect of hybridity. Significance letters are based on a generalized least square (GLS) model with a factor allowing for each group to have a different variance. Levels that share a letter are not statistically different.

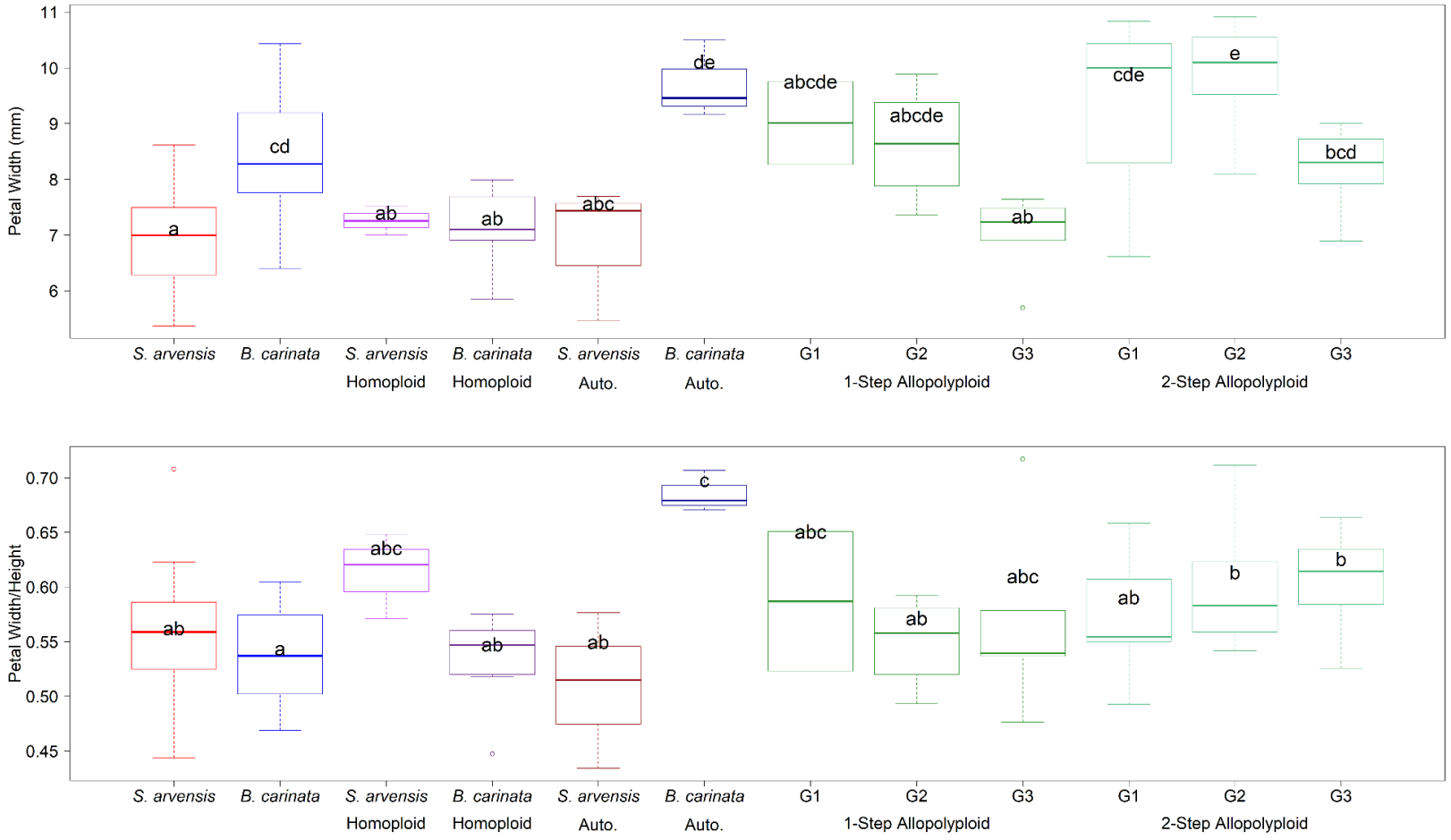


Figure 18: Petal width and petal width/height of allopolyploids, autopolyploids, homoploids, and parentals, showing the changing effect of polyploidy over generations. Significance letters are based on a generalized least square (GLS) model with a factor allowing for each group to have a different variance. Levels that share a letter are not statistically different.

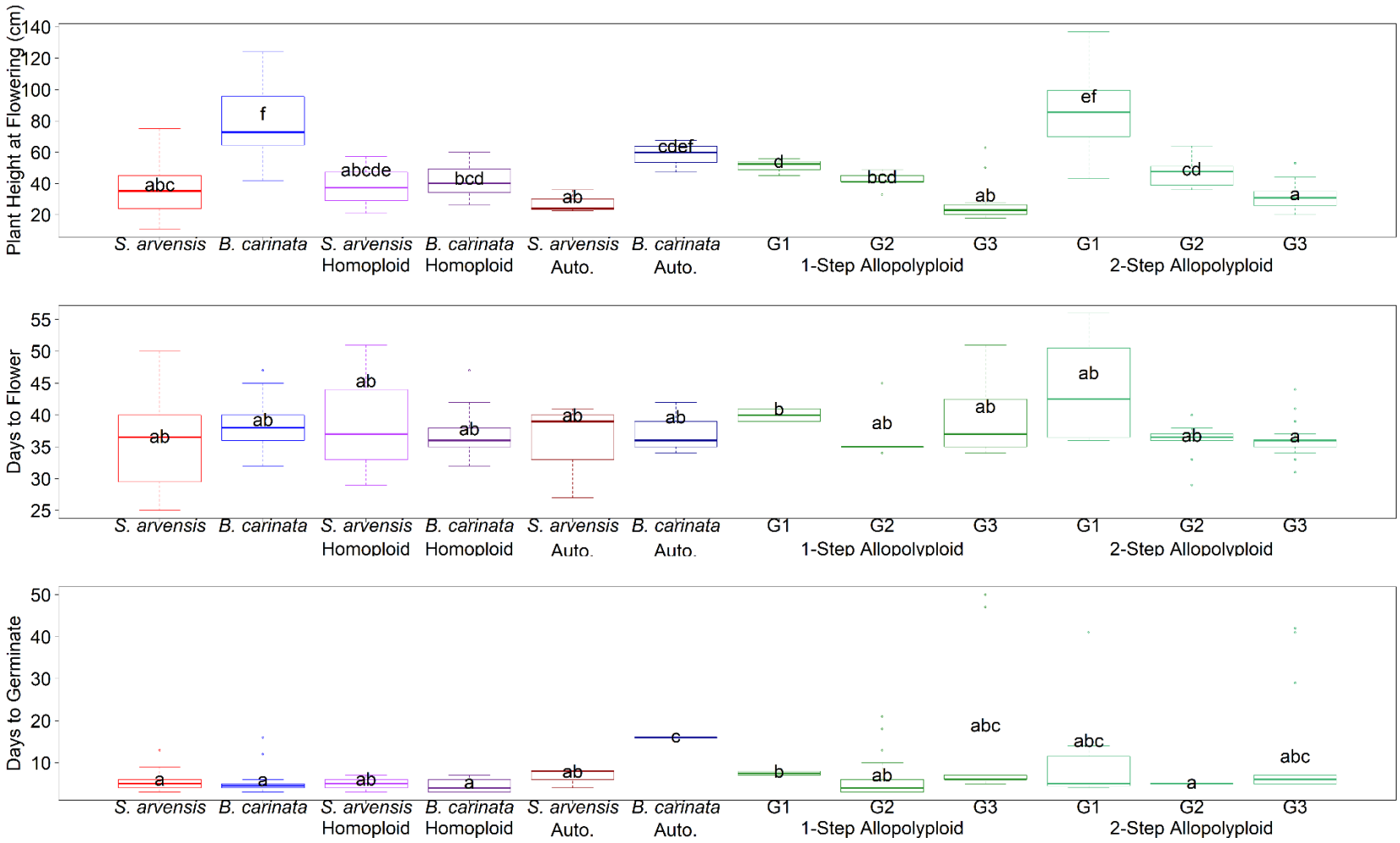


Figure 19: Developmental markers of allopolyploids, autopolyploids, homoploids, and parentals, showing little or no effect of polyploidy. Significance letters are based on a generalized least square (GLS) model with a factor allowing for each group to have a different variance. Levels that share a letter are not statistically different.

Pollen size

Pollen size correlated significantly with DNA content when *B. carinata*, *B. juncea*, *B. napus*, *S. arvensis*, their homoploid hybrids, the *S. arvensis* and the *B. carinata* autopolyploids, and the *B. carinata*-maternal allopolyploids were compared ($p < 0.01$). Pollen size also correlated significantly with DNA content when considering only allopolyploids and their descendants ($p = 0.01$). An ANOVA indicated that aneuploid descendants of allopolyploids had significantly smaller pollen than strict allopolyploids ($p < 0.01$). Considering only strict allopolyploids, there was no effect of type or generation, though likely too few individuals were included to find significance (average 3.2 individuals per group).

Hybridity, polyploidy, and their interaction all contributed to the variation in pollen size ($p < 0.01$ for all). Homoploid hybrids had large pollen equivalent to the pollen of polyploids, likely indicating a high rate of unreduced gametes (Figure 20). Unexpectedly, the pollen of *S. arvensis* was larger than the pollen of *B. carinata* despite the lower 2C DNA content of *S. arvensis*, and this held true for their respective autopolyploids and the homoploid for which they were the maternal parent (Figure 20).

While most individuals and groups had approximately normal distributions of pollen size, some individuals, especially homoploid hybrids, showed two peaks in histograms of their pollen size (Figure 21).

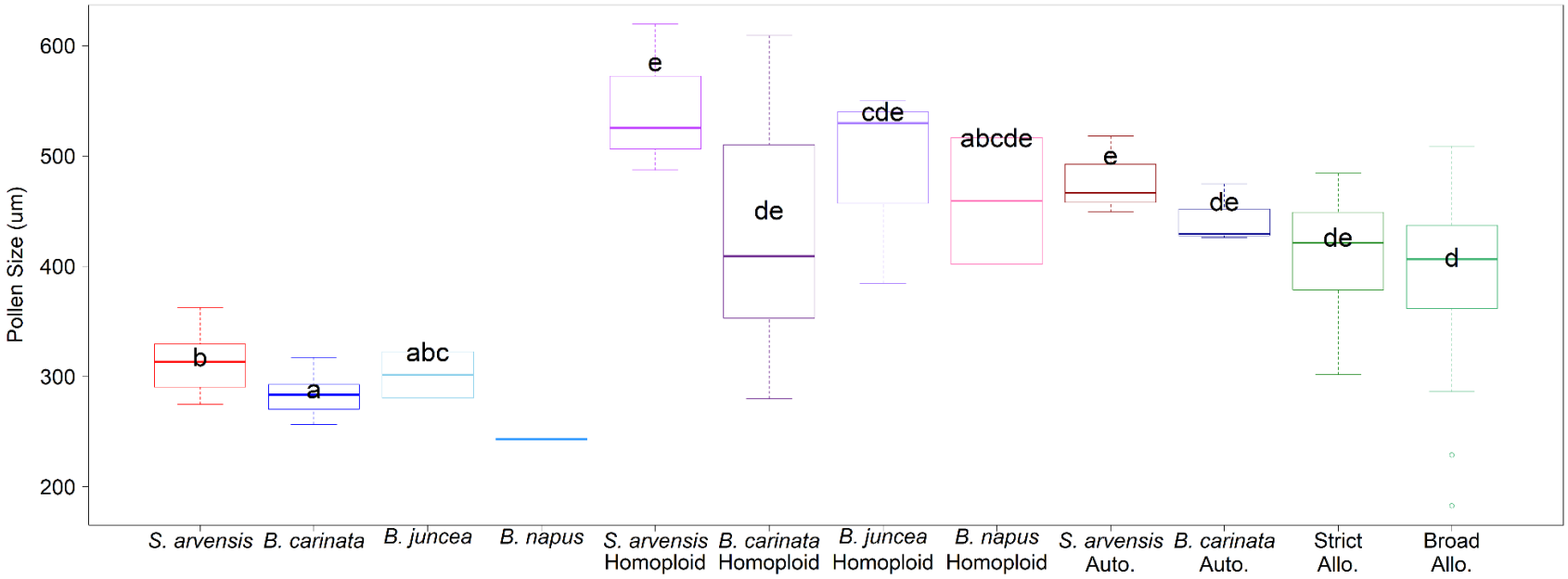


Figure 20: Pollen sizes of polyploids, homoploids, and parentals. Significance letters are based on a generalized least square (GLS) model with a factor allowing for each group to have a different variance. Levels that share a letter are not statistically different.

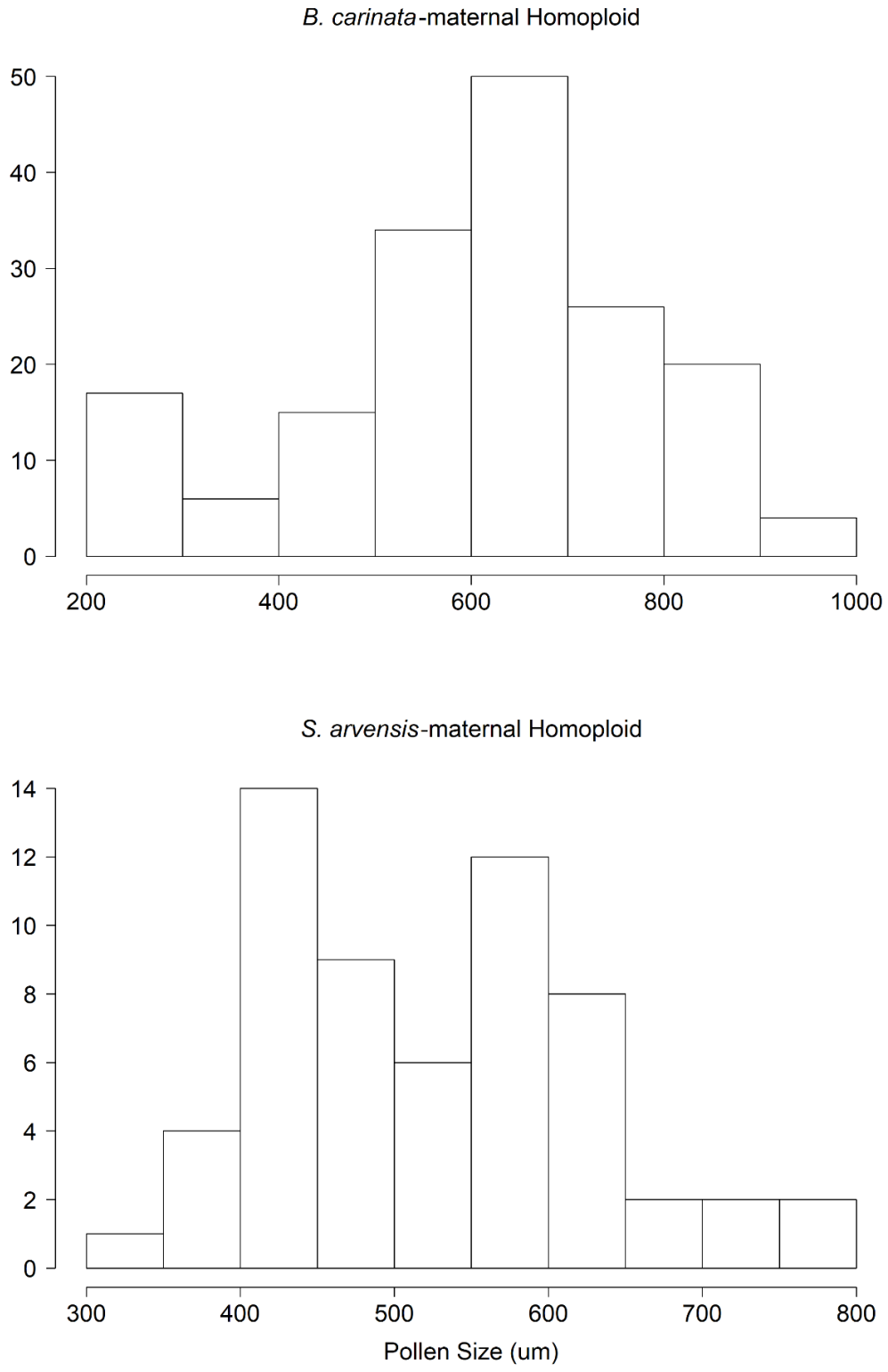


Figure 21: Examples of individuals for which histograms of pollen size had two peaks.

4. Discussion and Conclusions

Formation rates of one-step vs. two-step allopolyploids

Here *B. carinata* was the only species that produced both one-step and two-step allopolyploids with *S. arvensis*, and for this cross one-step allopolyploids formed twice as frequently as two-steps (Table 3). Since this study is the first to compare formation rates through different pathways, it is unknown if this pattern holds true in other crosses.

One-step allopolyploid formation rate is a function of the rate of unreduced gamete production by the parental plants while two-step formation rate is a function of hybrid formation rate by the parental plants and unreduced gamete production in those hybrids. Ramsey and Schemske (1998) suggested that the expected rate of one-step allopolyploid formation between two species without backcrossing could be estimated as the hybridization rate times $4.09 \text{ X } 10^{-2}$ based on average hybrid fertility and unreduced gamete production. Cheung et al. (2015) estimated the hybridization rate between *B. carinata* and *S. arvensis* was 6.43% suggesting that one-step allopolyploids should be observed at a rate of approximately 1 in 50 ovules pollinated. The hybridization rates between *B. junceae* and *B. napus* determined here were 0.25% and 0.008%, which would lead to estimates of one-step allopolyploid formation rates of approximately 1 in 10,000 and 1 in 300,000, respectively. While an insufficient number of crosses with *B. napus* or *B. juncea* as the maternal parent were done to exclude the possibility that one-step polyploids could form between these species and *S. arvensis*, the one-step allopolyploidy rate found here of approximately 1 in 10,000 is much lower than the calculation from Ramsey and Schemske (1998) would expect. There are several possible explanations for this lower rate: 1) the unreduced gamete formation rate suggested by Ramsey and

Schemske (1998) was higher than the rate in the parentals studied here, 2) the hybrid fertility in this system is lower than estimated by Ramsey and Schemske (1998), or 3) the majority of unreduced gametes united did not form viable seeds. Flow cytometry on the pollen of the individuals used as parents here would have provided greater insight into the first possible explanation. The average pollen fertility of the homoploid hybrids was estimated as less than 1% of the parental species, which is much lower than the estimate that hybrids would be about 6.25% as fertile as their parents (Ramsey and Schemske 1998). Finally, very few individuals that were the product of an unreduced gamete and a reduced gamete were observed, even though this combination should be occurring at a much higher frequency than the union of two unreduced gametes. This suggests that a large number of unreduced gametes produced by both *B. carinata* and *S. arvensis* do not go on to form viable progeny.

Previous work indicates that unreduced gamete production rate varies among individuals and species as well as with biotic and abiotic environmental conditions (Mason et al., 2011; Sora et al., 2016). For example, flow cytometry estimates indicated that on average approximately 0.6% of *S. arvensis*'s pollen and 1% of *B. napus*'s pollen were unreduced (Kreiner et al., 2017) although a substantial amount of variation among *S. arvensis* individuals was reported (0.06% to 2.17%) and a quarter of the pollen produced was unreduced for one outlier (27%) (Sora et al., 2016). Estimates from male gamete counts (Mason et al., 2011) were lower than flow cytometry-based estimates for five lines of *B. napus* (0 - 0.11%) and estimated unreduced male gamete production in *B. juncea* as 0.03% and in *B. carinata* as 0% based on observation of between 600 and 2000 male gametes. The higher rate of unreduced gametes in *B. napus* and *B. juncea* than *B.*

carinata is surprising given that only *B. carinata* produced one-step allopolyploids with *S. arvensis*. This may be due to variation within the species or it may indicate that unreduced gamete production is not the only factor that impacts one-step allopolyploid formation.

Once a homoploid hybrid formed, two-steps were ten times as frequent as one-steps for *B. carinata* and were a certainty for *B. napus*. This observation is in line with the observation by Ramsey and Schemske (1998) that in hybrid systems polyploid progeny are often the primary or only offspring produced. This is likely the result of the higher rate of unreduced gamete production in homoploid hybrids due to difficulties in meiosis with noncomplimentary chromosomes. The rate of unreduced gamete production has been estimated to be up to 50 times higher in hybrids than in parentals (Kreiner et al., 2017; Mason et al., 2011; Ramsey and Schemske, 1998; Sora et al., 2016; Zhang et al., 2010) and results in a higher rate of polyploidization (Zhang et al., 2016). Previous work on *B. napus* X *S. arvensis* hybrids generated by embryo rescue indicated that even after seven generations of backcrossing to *B. napus* unreduced gamete production rates were significantly higher in the backcrossed lineage than the parental lineages (Sora et al., 2016). Similarly, a *B. napus* X *S. arvensis* hybrid produced by ovule culture favored the production of gametes with more chromosomes than expected (Bing et al., 1995). Here, average pollen sizes for homoploids between all three *Brassica* species and *S. arvensis* were the size expected for unreduced gametes, equivalent to that of the polyploids (Figure 20). In a combined sample of anthers from these *B. carinata* x *S. arvensis* homoploid hybrids, flow cytometry determined that 94% of pollen was unreduced (Paul Kron, unpublished data) and pollen size data for several homoploids showed two peaks in

histograms (Figure 21), suggesting that they were producing reduced and unreduced pollen grains. Additionally, Mason et al. (2011) found that unreduced pollen was more viable than reduced pollen in hybrids between the triangle of U allopolyploids. Together this high rate of unreduced gamete formation in the homoploid hybrids and increased viability of their unreduced gametes likely explains the high rate of two-step allopolyploids in the *B. napus* and *B. carinata* lineages. In the case of *B. juncea,* two-step polyploids may not have been produced at the rate expected given the number of homoploid hybrids and expected unreduced gamete formation because of the high level of mildew on the hybrids or genome incompatibilities.

Overall, this study provides the first evidence that one-step allopolyploid formation is more frequent than two-step allopolyploid formation. As predicted by Ramsey and Schemske (1998), this is because formation of two-step allopolyploids is limited by the rate of hybrid formation.

Genomic distance affects hybridization rate

Brassica napus (AACC) had the lowest hybridization rate with *S. arvensis* (SS, similar to BB) as expected based on previous work (Bing et al., 1995; Mizushima, 1950). *B. napus* does not contain the B genome and is more genetically distant from *S. arvensis* than either *B. carinata* (BBCC) or *B. juncea* (AABB), so its low hybridization rate supports the observation that hybridization rate is dependent on genetic distance (Buggs et al., 2014; Levin, 2013; Tayalé and Parisod, 2013). The exclusive production of polyploid offspring of the *B. napus* X *S. arvensis* homoploid hybrids is in line with the suggestion that polyploidy can restore fertility to otherwise incompatible crosses (Sora et

al., 2016). As *B. napus* and *B. carinata* produced allopolyploids with *S. arvensis* while *B. juncea* did not, it is also possible that there is an aspect of the C genome that facilitates polyploidization or that mitigates the interactions between the B and S genomes.

However, hybridization and unreduced gamete production rates are genotype-dependent, so the performance of the accessions chosen here may not be indicative of the species overall (David et al., 2004; Soltis et al., 2010). For all three *Brassica* species, hybridization rate varied between accessions and one of the two *B. juncea* accessions performed poorly in the greenhouse overall. Bing et al. (1995) produced no hybrids in greenhouse crosses between four *S. arvensis* cultivars and six *B. napus* cultivars – including Westar, which was the maternal parent of the two *B. napus* x *S. arvensis* hybrids found here – though only 20-60 flowers were crossed per cultivar which may not have given enough power to detect hybridization. Lefol et al. (1996) found a hybridization rate of 1:8,300, similar to the rate found here (Table 3), between male-sterile *B. napus* grown in a garden with *S. arvensis*. However, Fitzjohn (2007) reviewed literature on crosses in Brassicaceae and found a median of 0% hybrid production for all three crosses in this study. More accessions need to be tested to verify these results are indicative of the species overall.

In conclusion, allopolyploid formation rates vary based on the genomes involved because parental genomic distance directly affects the fertility and formation rate of homoploid hybrids, which is a bottleneck in two-step allopolyploid formation rates. This is consisted with previous observations (Buggs et al., 2014; Levin, 2013; Tayalé and Parisod, 2013).

Two-step allopolyploids had higher fertility but fertility declined over generations

Seed and pollen fertility were both significantly higher for the two-step allopolyploids than the one steps in the first generation. This provides the first experimental evidence to support the hypothesis that passing through a hybrid intermediate may eliminate some incompatibilities. The fertility of both types of allopolyploid was lower than that of *B. carinata*, though seed set for the two-steps was not significantly different than that of *S. arvensis* in the first two generations. This is unexpectedly high as allopolyploids are expected to be less fit than their parental species.

For both one- and two-step allopolyploids, seed set and pollen viability decreased or stayed constant over the three generations (Figure 12). This bottleneck of low fertility in early generations is expected for allopolyploids (Comai, 2005; Lim et al., 2008): the resynthesized *B. napus* in Gaeta et al. (2007) did not recover fertility even after five generations, and in a common garden experiment with wild and cultivated radish it took ten years for homoploid hybrids to achieve parental fertility (Snow et al., 2010). Because fertility correlated with DNA content (Figure 13), selection would likely have favored strict allopolyploids and may have resulted in stabilization of the genome and increased fertility. However, selection pressure was limited in this experiment because the same number of seeds was sowed from each individual to form the next generation. More generations sown in a way that allows selection may provide the opportunity for stabilizing selection on meiotic behaviour and the elimination of the aneuploids (Lim et al., 2008), and may better indicate the probability of establishment for the allopolyploids. The results of the hand-crosses on two-step allopolyploids, which showed that flowers pollinated with self pollen produced more seed than unmanipulated flowers (Figure 7),

suggest that there may be morphological characteristics such as twisted anther filaments (Figure 10) that affect the ability of allopolyploids to efficiently self-pollinate in addition to meiotic challenges. Two-step allopolyploids, however, had significantly higher seed fertility than one-steps and fewer aneuploids (see below) indicating that polyploids that form through this pathway may have a higher chance of successful establishment.

Incomplete reproductive isolation between cytotypes

Polyploidization is often considered a mechanism of near-instant speciation (Abbott et al., 2013; Hegarty et al., 2006; Sikka, 1940) but gene flow is known to occur between cytotypes (Abbott et al., 2013; Parisod et al., 2010; Soltis et al., 2010) and can be a mechanism of allopolyploid formation (Ramsey and Schemske, 1998). Here, the one-step allopolyploids, the autopolyploids, and one two-step individual produced as much seed with pollen from their maternal parental as with their own. However, as this seed was not tested we do not know if seed from these crosses is viable. It is also unknown what types of gametes formed the seeds, and therefore whether the offspring are hexaploids or of an intermediate ploidy. The two-step allopolyploids appeared more reproductively isolated from their parentals than the one-steps (Figure 7). Because the one-steps produced so many offspring with lower than expected DNA contents (Figure 6), this raises the question of whether the one-step allopolyploid gametes that successfully formed seeds with parental pollen contained less DNA than expected.

The lack of complete reproductive isolation means that polyploids may have an impact on parental populations even if they do not establish by facilitating gene flow between species. If the backcrossed seed is viable, the fact that polyploids produced seed

as easily with parents as with themselves may also diminish their minority cytotype disadvantage, which is assumed to be one of the biggest obstacles to polyploid populations (Oswald and Nuismer, 2007; Parisod et al., 2010; Ramsey and Schemske, 1998). Backcrossing with parentals, especially parentals with a high rate of unreduced gamete production, may allow an isolated polyploid to persist until more polyploidization events can occur and may be a route to allopolyploid formation (Ramsey and Schemske, 1998), though this would depend on the ploidy of the backcrossed offspring. Alternatively, the reproductive isolation of the two-steps may be advantageous by facilitating divergence and avoiding low-fitness hybrids (Husband and Schemske, 2000; Oswald and Nuismer, 2007; Parisod et al., 2010).

Formation route affects morphology

ANOVAs examining the effect of formation pathway and generation indicated that one-step and two-step allopolyploids showed significant differences in trichrome density and floral characteristics such as petal size. Initially, one-step allopolyploids were more similar to the homoploid hybrids while two-step allopolyploids were more similar to the autopolyploids, though two-steps moved towards the homoploid hybrids by the third generation. Either the formation route or the specific genomic contents of the parentals involved in the one and two- step allopolyploids had a significant effect on the morphology of the allopolyploids.

The changes in the allopolyploids over the generations indicate that the allopolyploids have not yet stabilized, which is expected given that gene expression changes in allopolyploids vary for several generations after formation (Wang et al.,

2004). While the autopolyploids were morphologically similar to their parentals, and a few first-generation two-step allopolyploids were also grouped with *B. carinata* in the LDA (Figure 15), the allopolyploids were morphologically distinct from their parentals likely due to the large changes that occur in allopolyploid genomes from the first generation (Levy and Feldman, 2004; Song et al., 1995). The allopolyploids were more similar morphologically to *B. carinata* than to *S. arvensis*, which is consistent with previous observations that hybrids are more similar to their maternal parents (Warwick et al., 2003).

Effects of hybridity on morphology endure but effects of polyploidy fade

Hybridity and polyploidy both contributed to the phenotype of the allopolyploids. The effect of polyploidy seemed to decrease over the three generations, while the effect of hybridity increased. The significant and enduring effect of hybridity is consistent with research that established that the majority of gene expression changes associated with allopolyploidy are due to hybridization (Adams, 2007; Albertin et al., 2006; Hegarty et al., 2006). The immediate but decreasing effect of polyploidy observed here may not have been found in studies on established allopolyploids, and demonstrates that there is more to learn about the effects of hybridity and polyploidy on neoallopolyploids (Soltis et al., 2010; Tayalé and Parisod, 2013). The one-step and two-step allopolyploids were affected differently by hybridity and polyploidy, with less of the variation in one-step allopolyploids explained by polyploidy than in two-steps, suggesting that formation pathway should be considered when untangling the effects of hybridity and polyploidy.

Polyploidy is associated with increased stomata length, larger flowers, and

delayed development (Husband and Schemske, 2000; Materson, 1994; Ramsey and Schemske, 2002; Sax and Sax, 1937). Both stomata length and petal width correlated with DNA content in these samples, though GLS analysis did not find significant differences between most groups likely due to small sample sizes. The stomata size of the three autopolyploid *S. arvensis* was larger on average than the diploids, but highly variable resulting in no significant difference between the groups. There was less variability in the *B. carinata* autopolyploids, which resulted in a significant difference (Figure 16). However, the morphology of the colchicine-induced *B. carinata* autopolyploids should be considered with reservations since somatic doubling via colchicine is known to have different effects on meiotic behaviour than formation by unreduced gametes (Szadkowski et al., 2011) and the morphological effects of colchicine-induced doubling are unknown. Hybrids had stomata sizes that were intermediate to the parents, variable, and did not differ significantly from the majority of the one- and two-step allopolyploids across generations. The decrease in flower size in allopolyploids in the third generation was unexpected as polyploidy is often associated with larger flowers (Figure 18). This may have been caused by greenhouse conditions or may indicate that the initial increase in flower size in neopolyploids does not endure. Also in contrast to expectation, development did not seem to be delayed in the polyploids (Figure 19). Polyploidy did not explain the variation in days to flower and likely only explained variation in days to germinate due to outliers and the late germination of the *B. carinata* autopolyploids which, as mentioned above, may be affected by the colchicine-induced somatic doubling. The difference in flowering time associated with polyploidy in the literature may result from post-polyploidization selection for different flowering times

as a prezygotic reproductive barrier rather than a direct cause of polyploidy, since preventing hybrids with parental species may increase fitness (Husband and Schemske, 2000). Martin and Husband (2012) found that when exposed to selection for later flowering time neopolyploids had a higher evolutionary response than diploids or existing autopolyploids lineages, and attributed this rapid evolution to polyploidy itself. The minimal selection pressure in this experiment may explain why polyploidy did not result in delayed development.

Given these results, it appears that hybridization has a more significant and lasting effect than polyploidy on the morphology of these allopolyploids.

Dramatic DNA downsizing in allopolyploid offspring

Meiotic aberrations and resulting aneuploidy are common in both auto- and allo-neopolyploids (Lim et al., 2008; Parisod et al., 2010; Ramsey and Schemske, 2002). Homoploid hybrids are also known to produce offspring with a wide range of DNA contents (Zhang et al., 2016). However, DNA downsizing back to the level of the parentals in early generations without backcrossing is unexpected. The polyploid ratchet theory, which suggests that ploidy can only increase over short evolutionary timescales (Meyers and Levin, 2006), is often used as a null model in studies on the evolutionary advantage of polyploidy (Parisod et al., 2010; Scarpino et al., 2014). These studies assume that if there is an evolutionary advantage to polyploidy, polyploids should be more abundant than can be explained by the ratchet model. The decrease of ploidy in this system suggests that the ratchet model may not be an appropriate null model for every situation. Additionally, “polyploid drop” or “descending dysploidy”, where during the

process of diploidization there is a heritable decrease in chromosome number, is known to occur in polyploid systems, and higher ploidies, such as hexaploids, may show more dramatic drops than lower ploidies (Mandáková et al., 2017; Mandáková and Lysak, 2018). The variation in DNA content of the descendants of the allopolyploids found here may be the initial stages of this process. These studies, however, found descending dysploidy by looking across established mesopolyploid taxa, so the speed at which these changes begin following polyploidization has not yet been established. Our observations suggest it may begin immediately after polyploidization. Interestingly, while the range of DNA contents in the descendants of the *B. carinata*-maternal allopolyploids matched the expectation for descending dysploidy, the *B. napus*-maternal allopolyploids showed ascending dysploidy, which is less common (Mandáková and Lysak, 2018). It is possible that other researchers may have had DNA downsizing to this degree but did not realize it due to only verifying ploidy in the first generation. In Gaeta et al. (2007), for instance, loss of DNA markers was observed over five generations in 47 lineages of resynthesized *B. napus*, but whether the loss of markers was caused by the loss of chromosomes was not examined.

The amount of DNA lost in some descendants of the allopolyploids studied here may represent a whole genome or, in some cases, more. This raises the possibility that a specific genome may have been lost in its entirety. For instance, the loss of the entire *S. arvensis* genome would explain the decrease in DNA content from allopolyploid back to *B. carinata* levels (Figure 3). Molecular genetic analysis to determine which specific chromosomes have been retained would be required to verify this, but there is reason to believe polyploids may have mechanisms to distinguish between parental genomes.

There is often an immediate bias against the expression of the genome of one parent (Adams, 2007; Szadkowski et al., 2011; Tayalé and Parisod, 2013) and genomes can be spatially separated in allopolyploids (Bennett, 2004).

It is also possible that even those individuals with expected DNA contents may have lost or gained chromosomes or parts of chromosomes. In a study on *Tragopogon* allopolyploids, Lim et al. (2008) found only one aneuploid (missing one chromosome) but many highly fertile individuals had the correct number of chromosomes through trisomy and monosomy. They also determined that certain chromosomes were more likely to be involved in aneuploidy, with frequencies ranging from 0% to 40%. Rivero-Guerra (2008) also found trisomy and aneuploidy to be more common in tetraploid than diploid populations of *Santonlina pectinate*, though trisomy and aneuploidy did not affect pollen fertility. Rivero-Guerra (2008) found that in addition to the loss of whole chromosomes, polyploidy decreased the length of chromosome arms or caused the loss or gain of chromosome arms. Measuring DNA content with flow cytometry, as done in this study, would not reveal any of these changes.

The variation in DNA contents suggests there may be a few numbers of chromosomes that are functional and different lineages are following different paths to fitness. Figure 3 and 4 show that lineages that begin to lose chromosomes often continue to lose chromosomes, and lineages that gain chromosomes continue to gain. The increasing variation in DNA content over generations of both one-steps and two-steps (Figure 6) provides material for selection, potentially resulting in alternative successful chromosome conformations. In the one-step allopolyploids especially, the increased genomic instability is likely to be disadvantageous, but may suggest alternate routes to

fitness in the case of some high-fertility aneuploids. For instance, descending dysploidy, which may be occurring, is likely responsible for the wide range of base chromosome numbers often found in Brassicaceae genera (Mandáková et al., 2017; Mandáková and Lysak, 2018). While the results of a recognizable allopolyploidization event between given species are expected to be repeatable, there is evidence that polyploids may differentiate as is observed here. Werth and Windham (1991), for example, found geographically isolated populations of allotetraploids experienced silencing of the same gene in different parental genomes, resulting in hybrid sterility and allopatric speciation. Loss of chromosomes similarly causes reproductive isolation (Mandáková and Lysak, 2018). The significantly greater DNA downsizing in the one-steps compared to the two-steps suggests again that formation pathway needs to be considered when studying neopolyploids. Allopolyploidization may be responsible for generating a wider range of variation than might typically be recognized when examining established polyploids and their diploid progenitors.

Importance of rare individuals

Despite the correlation between DNA content and fertility (Figure 13), some aneuploids had surprisingly high fertility. These could represent alternate successful chromosome arrangements and these plants may be able to establish populations with alternate cytotypes if their high fertility persists. Rare individuals like these could also have an impact on parental populations. For example, the two offspring of one-step allopolyploids with extremely low DNA contents (~2.3 pg, less than the 2.9 pg expected for *B. carinata*) were the only member of the *B. carinata* lineage able to produce seed

with *S. arvensis* (Figure 9). These individuals could facilitate gene flow between *B. carinata* and *S. arvensis*. Similarly, only one two-step was able to backcross to *B. carinata*. Rare individuals also impact polyploid formation. A few homoploid hybrids with much higher than average pollen viability or much higher than average unreduced gamete production rates have outsized effects on polyploid production (Kreiner et al., 2017; Warwick et al., 2008; Zhang et al., 2010). When estimating the chances of establishment of polyploids, rare individuals with high fertility or a unique crossing ability should not be underestimated.

Conclusions

This is the first study to compare formation rates and resulting fertility and morphology of allopolyploids formed through different pathways. Allopolyploids were formed more frequently through the one-step pathway, but once homoploid hybrids formed two-steps were much more frequent. The hybridization rate, which greatly impacted allopolyploid formation through the two-step pathway, was dependent on genetic distance as predicted (Levin, 2013).

One-step allopolyploids were originally less fertile than two-steps (Figure 12) in line with the hypothesis that the formation of a semi-fertile hybrid would reduce genomic incompatibilities. This shows that while one-step allopolyploids form more frequently, two-step allopolyploids are initially more successful. With the exception of the pollen viability of the third-generation two-step allopolyploid, no allopolyploid achieved equal pollen fertility to their parentals (Figure 12). This is expected since all polyploids go through a bottleneck of decreased fertility (Comai, 2005; Lim et al., 2008). However,

seed fertility in the first two generations of the two-step allopolyploids surprisingly did not differ statistically from *S. arvensis*. The one-step allopolyploids were less reproductively isolated from their parentals than the two step allopolyploids (Figure 7) and also produced offspring with a wider and lower range of DNA contents (Figure 6). The causes of this may have also caused their lower fertility. Whether these allopolyploids could establish a population would depend on many factors, but the high fertility of a few individuals suggests that there may be conditions where some allopolyploids could establish. The wide variation in DNA contents provides abundant material for selection and the surprisingly high fertility and low reproductive isolation of some aneuploids suggests that even if strictly allopolyploid populations cannot establish, their aneuploid descendants might. The higher fertility of the two-steps and the lower reproductive isolation of the one-steps suggest that different conditions might favor the establishment of allopolyploids from different pathways.

The one-step and two-step allopolyploids were similar morphologically but not identical (Figure 14). Polyploidy explained less of the variation in the traits of the one-step allopolyploids, as expected by the greater similarity between one-step allopolyploids and homoploid hybrids than between two-step allopolyploids and homoploid hybrids. Both hybridity and polyploidy explained the variation in the morphology of the allopolyploids, but the effect of hybridity endured over three generations while the effect of polyploidy decreased.

This study is the first to show that there are significant differences in formation rate, fertility, reproductive isolation, and morphology of allopolyploids formed through the one-step or two-step pathway. More plant species need to be studied to determine if

these differences are unique to this study system or widespread. These results show that formation pathway should be taken into consideration in future research on neoallopolyploids.

Future directions

To draw general conclusions from this research, this study would need to be repeated on other individuals, accessions, and species. The allopolyploids studied here descended from a small base of parentals and homoploid hybrids, and the influence of their individual parents cannot be determined without more repetitions. Unreduced gamete production, for example, can vary not just by individual but by flower and even anther (Kreiner et al., 2017). Plant groups and species can vary widely (Husband et al., 2013), so other species need to be studied before general conclusions can be drawn about the impact of formation pathway.

In addition to the effect of the individual or accession, the effect of maternal lineage is a major factor that should be considered in future research. Cytoplasmic effects are large in plant evolution and reciprocal allopolyploids – with the maternal and paternal parents swapped – may have different morphology or be reproductively isolated from each other (Soltis et al., 2010). Resynthesized *B. napus* differed in the amount of homeologous crossovers in the first meiosis depending if *B. rapa* or *B. oleracea* was the maternal parent, which may have large effects on progeny (Szadkowski et al., 2010). Hybrids are often more morphologically similar to their maternal parent (Warwick et al., 2003) and unreduced gamete production depends on maternal line (Zhang et al., 2010). In this study, the homoploid hybrids between *B. carinata* and *S. arvensis* were more

morphologically similar to *B. carinata*. It is much more difficult to produce hybrids and allopolyploids between the species studied here if *S. arvensis* is the maternal parent due to its smaller genome size (Cheung et al., 2015), but maternal effects should be considered (Gaeta and Pires, 2010).

While it seems likely that the polyploid production rate is correlated with the unreduced gamete production rate, this system would be ideal to provide evidence for that assumption (Soltis et al., 2010). The rate of unreduced gamete production in the parentals and homoploid hybrids should be quantified using new flow cytometry techniques (Kron and Husband, 2015) and compared to polyploid formation rate to determine if they are directly correlated or if other factors impact polyploid formation beyond the presence of unreduced gametes. A question related to the unreduced gamete production rate is why no odd-ploidy individuals were found in this study. Triploids are common in systems with polyploidy and often involved with polyploid establishment (Husband, 2004). Allopolyploid formation through an odd-ploidy individual is the third formation pathway alongside the one-step and two-step pathways (Tayalé and Parisod, 2013). However, no odd-ploidy individuals were found here which suggests that even in the presence of unreduced gametes some polyploids may not be able to form.

Another surprising result of this research was the production of a wide range of aneuploids. This may be a sign that "polyploid drop" (Mandáková et al., 2017; Mandáková and Lysak, 2018) begins immediately after polyploidization. The effect of large aneuploidy compared to the effect of polyploidization is unknown (Soltis et al., 2010) and the aneuploids found here should be examined for fertility and morphology and compared to the strict allopolyploids to identify differences between them. The study

should be repeated to see whether the pattern of aneuploidy, where *B. carinata*-maternal allopolyploids lost chromosomes but *B. napus* allopolyploids gained them, is an effect of species, accession or individual. Future aneuploids should be hand crossed to their siblings of differing DNA contents to see if changes in chromosome numbers are causing reproductive isolation among the siblings. The seed produced by the autopolyploids in this experiment should also be tested to determine if the autopolyploids show the same instability in genome size. If they do not, that would suggest that the aneuploidy found in the allopolyploids is an effect of hybridity. Most importantly, molecular genetic analysis should be done on the aneuploids to determine which chromosomes have been lost: whether particular chromosomes are lost more easily as has been found before (Lim et al., 2008) or whether chromosomes from one parent are being lost preferentially.

A goal of this research was to determine if allopolyploids might provide a route for gene flow between the *Brassica* crops and the weed *S. arvensis*. Conducting experimental crosses as was done here is an essential first step in risk analysis (Armstrong et al., 2005). Hand crosses with the allopolyploids as the maternal parent generated seed, and this seed should be tested for viability and fertility. Additionally, hand crosses should be performed in the other direction with parental species as the maternal parent, especially with pollen from the aneuploids with DNA contents closer to the parentals. This would help to determine whether allopolyploids might allow genes to move to *S. arvensis* populations or volunteer *Brassica* crop populations, thus impacting the two species even if the polyploids die out in a few generations. Any chromosomes passed from an allopolyploid to backcrossed offspring are likely to contain introgressions from homoeologous crossover events (Gaeta et al., 2007; Gaeta and Pires, 2010; Klinger

et al., 1991; Lim et al., 2008; Szadkowski et al., 2010). Homoeologous rearrangements are likely selected against in the long term, so neopolyploids, which occur spontaneously and recurrently (Rivero-Guerra, 2008), would be the most likely to pass them on (Gaeta and Pires, 2010). Introgressions depend on allele, however, and some crop alleles may introgress more easily than others (Snow et al., 2010). Using techniques such as GISH (Genomic in situ Hybridization) to identify chromosomes participating in homoeologous rearrangements would contribute important information to a risk analysis for transgene escape. After hand crosses are completed, the parentals should be tested in field conditions to see if hybridization and polyploidization will occur without hand pollination.

If the allopolyploids can establish their own population, then they would provide a route for crop genes to escape. Future studies should test the allopolyploids and their parentals in common garden experiments and in field conditions to determine their relative fitness. Studies should also determine the evolutionary response rate of the allopolyploids, which may be faster than their parentals and thus give them an advantage in changing environments (Martin and Husband, 2012). Assessing the risk of transgene escape is important as transgenes have already escaped into wild populations of closely related species through hybridization, and can persist for more than five years (Warwick et al., 2008), as can non-transgenic crop alleles (Whitton et al., 1997). Allopolyploids in particular have also been shown to carry genes between crops and wild populations (David et al., 2004). Transgenes from *B. juncea* have passed to hybrids with *S. arvensis* in experimental tests, and despite a low hybridization rate, lack of backcrossing to *S. arvensis*, and a low chance of introgression, the risk of transgene escape was still high

due to the wide variety in fertility of the hybrids – by the fourth generation, the hybrids were vigorous and likely could compete with parentals (Warwick and Martin, 2013). One allopolyploid was found in this study and it had advantageous self-compatibility (Warwick and Martin, 2013). Hybridization can also increase weed aggressiveness through gene flow from non-transgenic crops if the crops provide traits that are beneficial to the wild populations (Ellstrand et al., 1999; Ellstrand and Hoffman, 1990). An important next step in assessing the risk of gene flow through polyploids in particular is determining which accessions of the *Brassica* crops are most likely to produce polyploids. Accessions that rarely produce polyploids can then be chosen for cultivation (David et al., 2004).

5. Supplementary Data

Supplementary Table 1: Theoretical DNA contents for parents and hybrids. Actual DNA contents for polyploids may be lower because the laser in the cytometer measures surface area of nuclei, which does not increase linearly with volume. The expected DNA content for a *B. carinata x S. arvensis* allopolyploid, therefore, is 3.95 pg not 4.2 pg.

Species or cross	Expected genome	n	Theoretical 2C DNA content (pg)
B. carinata	BBCC	17	2.9
S. arvensis	SS	9	1.3
B. napus	AACC	19	2.4
B. juncea	AABB	18	2.2
B. carinata autopolyploid	BBBBCCCC	34	5.8
S. arvensis autopolyploid	SSSS	18	2.6
B. carinata x S. arvensis homoploid hybrid	BCS	26	2.1
B. carinata x S. arvensis allopolyploid	BBCCSS	52	4.2
B. juncea x S. arvensis homoploid hybrid	ABS	27	1.7
B. juncea x S. arvensis allopolyploid	AABBSS	54	3.6
B. napus x S. arvensis homoploid hybrid	ACS	28	1.8
B. napus x S. arvensis allopolyploid	AACCSS	56	3.7

Supplementary Table 2. Flowers pollinated and seeds and hybrids produced from *B. juncea* and *B. napus* hand crosses with *S. arvensis*.

Maternal parent	Paternal parent (*S. arvensis*)	Flowers Pollinated	Seeds	Homoploid hybrids
B. juncea (Cutlass), Mother 1	8180	177	151	8
	1686	161	76	3
B. juncea (Cutlass), Mother 2	8180	87	43	4
	1686	114	93	3
B. juncea (Varuna), Mother 1	8180	110	8	0
	1686	108	0	0
B. juncea (Varuna), Mother 2	8180	92	13	0
	1686	103	6	0
B. juncea (Varuna), Mother 3	8180	18	9	0
	1686	11	0	0
B. juncea (Varuna), Mother 4	8180	0	0	0
	1686	3	0	0
B. napus (Westar), Mother 1	8180	126	41	0
	1686	121	25	2
B. napus (Westar), Mother 2	8180	126	16	0
	1686	129	21	0
B. napus (Global), Mother 1	8180	162	39	0
	1686	161	70	0

B. napus (Global), Mother 2	8180	105	20	0
	1686	92	27	0

Supplementary Table 3. Germination rates for seeds produced by Cheung *et al.*'s (2015) homoploid hybrids and DNA content of their offspring. The expected DNA content for a homoploid hybrid between these parentals is 2.1 pg and for an allopolyploid is 3.95 pg. *B. carinata* has a DNA content of 2.9 pg.

Maternal parent (Homoploid hybrid)	Paternal parent	Seeds produced	Seeds germinated	Germination rate (%)	DNA content of offspring (pg)
A2-15	Self	2	2	100	4.05, 4.18
A6-29	Self	1	0	0	NA
A10-7	Self	4	0	0	NA
A12-7	Self	1	0	0	NA
A18-23	Self	4	0	0	NA
A18-32	Self	67	62	92.5	1.95 to 2.98 (μ=2.65)
A18-37	Self	16	10 (1 died)	62.5	3.88 to 4.16 (μ=4.03)
A20-5	Self	17	4	23.53	1.96, 2.88, 2.94, 2.95
A22-27	Self	34	19	55.88	1.96 to 2.94

					(μ=2.62)
A22-37	Self	86	54	62.79	1.94 to 2.93 (μ=2.46)
A23-59	Self	10	0	0	NA
A27-21	Self	1	0	0	NA
A22-25	Self	1	0	0	NA
A1-2	Self	1	1	100	3.92
A30-233	Self	1	1 (died)	0	NA
A2-5	Self	1	0	0	NA
A24-18	Other homoploid	1	1	100	4.04

Supplementary Table 4: DNA content and germination rates of second-generation *B. napus* hybrids. The expected DNA content of an allopolyploid between these parentals is 3.7 pg.

Maternal parent (Homoploid hybrid)	Seeds produced	Seeds germinated	DNA content of offspring (pg)
NWC-100.24	8	4	3.77, 3.87, 4.11, 4.50
NWC-100.25	16	7	3.74, 3.77, 3.91, 3.95, 4.52, 4.92, 5.49

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W., Zinner, D., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Adams, K.L., 2007. Evolution of Duplicate Gene Expression in Polyploid and Hybrid Plants. J Hered 98, 136–141. https://doi.org/10.1093/jhered/esl061

Adams, K.L., Cronn, R., Percifield, R., Wendel, J.F., 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. PNAS 100, 4649–4654. https://doi.org/10.1073/pnas.0630618100

Albertin, W., Balliau, T., Brabant, P., Chèvre, A.-M., Eber, F., Malosse, C., Thiellement, H., 2006. Numerous and Rapid Nonstochastic Modifications of Gene Products in Newly Synthesized *Brassica napus* Allotetraploids. Genetics 173, 1101–1113. https://doi.org/10.1534/genetics.106.057554

Albertin, W., Brabant, P., Catrice, O., Eber, F., Jenczewski, E., Chèvre, A.-M., Thiellement, H., 2005. Autopolyploidy in cabbage (*Brassica oleracea* L.) does not alter significantly the proteomes of green tissues. Proteomics 5, 2131–2139. https://doi.org/10.1002/pmic.200401092

Armstrong, T.T., Fitzjohn, R.G., Newstrom, L.E., Wilton, A.D., Lee, W.G., 2005. Transgene escape: what potential for crop–wild hybridization? Molecular Ecology 14, 2111–2132. https://doi.org/10.1111/j.1365-294X.2005.02572.x

Bennett, M.D., 2004. Perspectives on polyploidy in plants – ancient and neo. Biol J Linn Soc 82, 411–423. https://doi.org/10.1111/j.1095-8312.2004.00328.x

Bing, D.J., Downey, R.K., Rakow, G.F.W., 1995. An evaluation of the potential of intergeneric gene transfer between *Brassica napus* and *Sinapis arvensis*. Plant Breeding 114, 481–484. https://doi.org/10.1111/j.1439-0523.1995.tb00840.x

Brownfield, L., Köhler, C., 2011. Unreduced gamete formation in plants: mechanisms and prospects. J Exp Bot 62, 1659–1668. https://doi.org/10.1093/jxb/erq371

Buggs, R.J.A., Wendel, J.F., Doyle, J.J., Soltis, D.E., Soltis, P.S., Coate, J.E., 2014. The legacy of diploid progenitors in allopolyploid gene expression patterns. Phil. Trans. R. Soc. B 369, 20130354. https://doi.org/10.1098/rstb.2013.0354

Chen, Z.J., 2010. Molecular mechanisms of polyploidy and hybrid vigor. Trends in Plant Science 15, 57–71. https://doi.org/10.1016/j.tplants.2009.12.003

Cheung, K.W., Razeq, F.M., Sauder, C.A., James, T., Martin, S.L., 2015. Bidirectional but asymmetrical sexual hybridization between *Brassica carinata* and *Sinapis arvensis* (Brassicaceae). J Plant Res 128, 469–480. https://doi.org/10.1007/s10265-015-0702-2

Comai, L., 2005. The advantages and disadvantages of being polyploid. Nat Rev Genet 6,

836–846. https://doi.org/10.1038/nrg1711
David, J.L., Benavente, E., Brès-Patry, C., Dusautoir, J.-C., Echaide, M., 2004. Are neopolyploids a likely route for a transgene walk to the wild? The Aegilops ovata × Triticum turgidum durum case. Biol J Linn Soc 82, 503–510. https://doi.org/10.1111/j.1095-8312.2004.00336.x
Doležel, J., Bartoš, J., 2005. Plant DNA Flow Cytometry and Estimation of Nuclear Genome Size. Ann Bot 95, 99–110. https://doi.org/10.1093/aob/mci005
Doležel, J., Greilhuber, J., Suda, J., 2007. Estimation of nuclear DNA content in plants using flow cytometry. Nat. Protocols 2, 2233–2244. https://doi.org/10.1038/nprot.2007.310
Doležel, J., Sgorbati, S., Lucretti, S., 1992. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. Physiologia Plantarum 85, 625–631. https://doi.org/10.1111/j.1399-3054.1992.tb04764.x
Ehrendorfer, F., 1980. Polyploidy and Distribution, in: Polyploidy, Basic Life Sciences. Springer, Boston, MA, pp. 45–60. https://doi.org/10.1007/978-1-4613-3069-1_3
Ellstrand, N.C., Hoffman, C.A., 1990. Hybridization as an Avenue of Escape for Engineered Genes. BioScience 40, 438–442. https://doi.org/10.2307/1311390
Ellstrand, N.C., Prentice, H.C., Hancock, J.F., 1999. Gene Flow and Introgression from Domesticated Plants into Their Wild Relatives. Annual Review of Ecology and Systematics 30, 539–563.
Ellstrand, N.C., Schierenbeck, K.A., 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? PNAS 97, 7043–7050. https://doi.org/10.1073/pnas.97.13.7043
FitzJohn, R.G., Armstrong, T.T., Newstrom-Lloyd, L.E., Wilton, A.D., Cochrane, M., 2007. Hybridisation within Brassica and allied genera: evaluation of potential for transgene escape. Euphytica 158, 209–230. https://doi.org/10.1007/s10681-007-9444-0
Gaeta, R.T., Pires, J.C., 2010. Homoeologous recombination in allopolyploids: the polyploid ratchet. New Phytologist 186, 18–28. https://doi.org/10.1111/j.1469-8137.2009.03089.x
Gaeta, R.T., Pires, J.C., Iniguez-Luy, F., Leon, E., Osborn, T.C., 2007. Genomic Changes in Resynthesized *Brassica napus* and Their Effect on Gene Expression and Phenotype. The Plant Cell 19, 3403–3417. https://doi.org/10.1105/tpc.107.054346
Grundt, H.H., Kjølner, S., Borgen, L., Rieseberg, L.H., Brochmann, C., 2006. High biological species diversity in the arctic flora. PNAS 103, 972–975. https://doi.org/10.1073/pnas.0510270103
Guo, M., Davis, D., Birchler, J.A., 1996. Dosage Effects on Gene Expression in a Maize Ploidy Series. Genetics 142, 1349–1355.
Hegarty, M.J., Barker, G.L., Wilson, I.D., Abbott, R.J., Edwards, K.J., Hiscock, S.J., 2006. Transcriptome Shock after Interspecific Hybridization in Senecio Is Ameliorated by Genome Duplication. Current Biology 16, 1652–1659. https://doi.org/10.1016/j.cub.2006.06.071
Husband, B.C., 2004. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. Biological Journal of the Linnean Society 82, 537–546. https://doi.org/10.1111/j.1095-8312.2004.00339.x
Husband, B.C., Schemske, D.W., 2000. Ecological mechanisms of reproductive isolation

between diploid and tetraploid *Chamerion angustifolium*. Journal of Ecology 88, 689–701. https://doi.org/10.1046/j.1365-2745.2000.00481.x
Husband, P.D.B.C., Baldwin, D.S.J., Suda, P.J., 2013. The Incidence of Polyploidy in Natural Plant Populations: Major Patterns and Evolutionary Processes, in: Greilhuber, J., Dolezel, J., Wendel, J.F. (Eds.), Plant Genome Diversity Volume 2. Springer Vienna, pp. 255–276. https://doi.org/10.1007/978-3-7091-1160-4_16
Jackson, J.A., Tinsley, R.C., 2003. Parasite infectivity to hybridising host species: a link between hybrid resistance and allopolyploid speciation? International Journal for Parasitology 33, 137–144. https://doi.org/10.1016/S0020-7519(02)00255-2
Klinger, T., Elam, D.R., Ellstrand, N.C., 1991. Radish as a Model System for the Study of Engineered Gene Escape Rates Via Crop-Weed Mating. Conservation Biology 5, 531–535. https://doi.org/10.1111/j.1523-1739.1991.tb00360.x
Kostoff, D., 1933. A contribution to the sterility and irregularities in the meiotic processes caused by virus diseases. Genetica 15, 103–114. https://doi.org/10.1007/BF01591853
Kostoff, D., Kendall, J., 1930. Irregular meiosis in *Datura ferox* caused by *Tetranychus telarius*. Genetica 12, 140–144. https://doi.org/10.1007/BF01508073
Kostoff, D., Kendall, J., 1929. Irregular meiosis in *Lycium halimifolium* Mill. Produced by gall mites (Eriophyes). Journ. of Gen. 21, 113–115. https://doi.org/10.1007/BF02983362
Kreiner, J.M., Kron, P., Husband, B.C., 2017. Frequency and maintenance of unreduced gametes in natural plant populations: associations with reproductive mode, life history and genome size. New Phytol n/a-n/a. https://doi.org/10.1111/nph.14423
Kron, P., Husband, B.C., 2015. Distinguishing 2N gamete nuclei from doublets in pollen using flow cytometry and pulse analysis. Cytometry 87, 943–957. https://doi.org/10.1002/cyto.a.22696
Kron, P., Husband, B.C., 2012. Using flow cytometry to estimate pollen DNA content: improved methodology and applications. Ann Bot 110, 1067–1078. https://doi.org/10.1093/aob/mcs167
Le Comber, S.C., Ainouche, M.L., Kovarik, A., Leitch, A.R., 2010. Making a functional diploid: from polysomic to disomic inheritance. New Phytologist 186, 113–122. https://doi.org/10.1111/j.1469-8137.2009.03117.x
Lefol, E., Danielou, V., Darmency, H., 1996. Predicting hybridization between transgenic oilseed rape and wild mustard. Field Crops Research, Multidisciplinary assessment of transgenic crop plants: Technical, Agronomic and Ecological Aspects 45, 153–161. https://doi.org/10.1016/0378-4290(95)00067-4
Levin, D.A., 2013. The timetable for allopolyploidy in flowering plants. Ann Bot 112, 1201–1208. https://doi.org/10.1093/aob/mct194
Levin, D.A., 1983. Polyploidy and Novelty in Flowering Plants. The American Naturalist 122, 1–25. https://doi.org/10.1086/284115
Levy, A.A., Feldman, M., 2004. Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization. Biol J Linn Soc 82, 607–613. https://doi.org/10.1111/j.1095-8312.2004.00346.x
Lim, K.Y., Soltis, D.E., Soltis, P.S., Tate, J., Matyasek, R., Srubarova, H., Kovarik, A., Pires, J.C., Xiong, Z., Leitch, A.R., 2008. Rapid Chromosome Evolution in Recently Formed Polyploids in *Tragopogon* (Asteraceae). PLOS ONE 3, e3353.

https://doi.org/10.1371/journal.pone.0003353
Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I.A.P., Zhao, M., Ma, J., Yu, J., Huang, S., Wang, Xiyin, Wang, Junyi, Lu, K., Fang, Z., Bancroft, I., Yang, T.-J., Hu, Q., Wang, Xinfa, Yue, Z., Li, H., Yang, Linfeng, Wu, J., Zhou, Q., Wang, W., King, G.J., Pires, J.C., Lu, C., Wu, Z., Sampath, P., Wang, Z., Guo, H., Pan, S., Yang, Limei, Min, J., Zhang, D., Jin, D., Li, W., Belcram, H., Tu, J., Guan, M., Qi, C., Du, D., Li, Jiana, Jiang, L., Batley, J., Sharpe, A.G., Park, B.-S., Ruperao, P., Cheng, F., Waminal, N.E., Huang, Yin, Dong, C., Wang, L., Li, Jingping, Hu, Z., Zhuang, M., Huang, Yi, Huang, J., Shi, J., Mei, D., Liu, J., Lee, T.-H., Wang, Jinpeng, Jin, H., Li, Z., Li, X., Zhang, J., Xiao, L., Zhou, Y., Liu, Z., Liu, X., Qin, R., Tang, X., Liu, W., Wang, Y., Zhang, Y., Lee, J., Kim, H.H., Denoeud, F., Xu, X., Liang, X., Hua, W., Wang, Xiaowu, Wang, Jun, Chalhoub, B., Paterson, A.H., 2014. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. Nature Communications 5, ncomms4930. https://doi.org/10.1038/ncomms4930
Lombard, V., Baril, C.P., Dubreuil, P., Blouet, F., Zhang, D., 2000. Genetic Relationships and Fingerprinting of Rapeseed Cultivars by AFLP: Consequences for Varietal Registration. Crop Science 40, 1417–1425. https://doi.org/10.2135/cropsci2000.4051417x
Mandáková, T., Li, Z., Barker, M.S., Lysak, M.A., 2017. Diverse genome organization following 13 independent mesopolyploid events in Brassicaceae contrasts with convergent patterns of gene retention. The Plant Journal 91, 3–21. https://doi.org/10.1111/tpj.13553
Mandáková, T., Lysak, M.A., 2018. Post-polyploid diploidization and diversification through dysploid changes. Current Opinion in Plant Biology, 42 Genome studies and molecular genetics 2018 42, 55–65. https://doi.org/10.1016/j.pbi.2018.03.001
Marillia, E.-F., Francis, T., Falk, K.C., Smith, M., Taylor, D.C., 2014. Palliser's promise: *Brassica carinata*, An emerging western Canadian crop for delivery of new bio-industrial oil feedstocks. Biocatalysis and Agricultural Biotechnology, Trait Introduction Methods and Innovation Platforms in Plant Biotechnology 3, 65–74. https://doi.org/10.1016/j.bcab.2013.09.012
Martin, S.L., Husband, B.C., 2012. Whole Genome Duplication Affects Evolvability of Flowering Time in an Autotetraploid Plant. PLOS ONE 7, e44784. https://doi.org/10.1371/journal.pone.0044784
Martin, S.L., Husband, B.C., 2009. Influence of phylogeny and ploidy on species ranges of North American angiosperms. Journal of Ecology 97, 913–922. https://doi.org/10.1111/j.1365-2745.2009.01543.x
Martin, S.L., Smith, T.W., James, T., Shalabi, F., Kron, P., Sauder, C.A., 2017. An update to the Canadian range, abundance, and ploidy of *Camelina* spp. (Brassicaceae) east of the Rocky Mountains. Botany 95, 405–417. https://doi.org/10.1139/cjb-2016-0070
Mason, A.S., Nelson, M.N., Yan, G., Cowling, W.A., 2011. Production of viable male unreduced gametes in *Brassica* interspecific hybrids is genotype specific and stimulated by cold temperatures. BMC Plant Biology 11, 103. https://doi.org/10.1186/1471-2229-11-103
Materson, J., 1994. Stomatal Size in Fossil Plants: Evidence for Polyploidy in Majority of

Angiosperms. Science.
Meyers, L.A., Levin, D.A., 2006. On the abundance of polyploids in flowering plants. Evolution 60, 1198–1206. https://doi.org/10.1554/05-629.1
Miller, J.S., Venable, D.L., 2000. Polyploidy and the Evolution of Gender Dimorphism in Plants. Science 289, 2335–2338. https://doi.org/10.1126/science.289.5488.2335
Mizushima, U., 1950. Karyogenetic studies of species and genus hybrids in the tribe Brassiceae of Cruciferae. Tohoku journal of agricultural research 1, 1–14.
Oswald, B.P., Nuismer, S.L., 2007. Neopolyploidy and pathogen resistance. Proceedings of the Royal Society of London B: Biological Sciences 274, 2393–2397. https://doi.org/10.1098/rspb.2007.0692
Pandit, M.K., Pocock, M.J.O., Kunin, W.E., 2011. Ploidy influences rarity and invasiveness in plants. Journal of Ecology 99, 1108–1115. https://doi.org/10.1111/j.1365-2745.2011.01838.x
Parisod, C., Holderegger, R., Brochmann, C., 2010. Evolutionary consequences of autopolyploidy. New Phytologist 186, 5–17. https://doi.org/10.1111/j.1469-8137.2009.03142.x
Ramsey, J., Schemske, D.W., 2002. Nepolyploidy in Flowering Plants. Annual Review of Ecology and Systematics 33, 589–639. https://doi.org/10.1146/annurev.ecolsys.33.010802.150437
Ramsey, J., Schemske, D.W., 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annu. Rev. Ecol. Syst. 29, 467–501. https://doi.org/10.1146/annurev.ecolsys.29.1.467
Rieseberg, L.H., 2001. Polyploid evolution: Keeping the peace at genomic reunions. Current Biology 11, R925–R928. https://doi.org/10.1016/S0960-9822(01)00556-5
Rivero-Guerra, A.O., 2008. Cytogenetics, geographical distribution, and pollen fertility of diploid and tetraploid cytotypes of Santolina pectinata Lag. (Asteraceae: Anthemideae). Bot J Linn Soc 156, 657–667. https://doi.org/10.1111/j.1095-8339.2007.00766.x
Sax, K., Sax, H.J., 1937. Stomata Size and Distribution in Diploid and Polyploid Plants. Journal of the Arnold Arboretum 18, 164–172.
Scarpino, S.V., Levin, D.A., Meyers, L.A., 2014. Polyploid Formation Shapes Flowering Plant Diversity. The American Naturalist 184, 456–465. https://doi.org/10.1086/677752
Sikka, S.M., 1940. Cytogenetics of *Brassica* hybrids and species. Journ. of Genetics 40, 441–474. https://doi.org/10.1007/BF03028231
Smith, T.W., Kron, P., Martin, S.L., 2018. flowPloidy: An R package for genome size and ploidy assessment of flow cytometry data. Applications in Plant Sciences 6, e01164. https://doi.org/10.1002/aps3.1164
Snow, A.A., Culley, T.M., Campbell, L.G., Sweeney, P.M., Hegde, S.G., Ellstrand, N.C., 2010. Long-term persistence of crop alleles in weedy populations of wild radish (*Raphanus raphanistrum*). New Phytologist 186, 537–548. https://doi.org/10.1111/j.1469-8137.2009.03172.x
Soltis, D.E., Buggs, R.J.A., Barbazuk, W.B., Schnable, P.S., Soltis, P.S., 2009. On the Origins of Species: Does Evolution Repeat Itself in Polyploid Populations of Independent Origin? Cold Spring Harb Symp Quant Biol sqb.2009.74.007. https://doi.org/10.1101/sqb.2009.74.007

Soltis, D.E., Buggs, R.J.A., Doyle, J.J., Soltis, P.S., 2010. What we still don't know about polyploidy. Taxon 59, 1387–1403.
Soltis, D.E., Soltis, P.S., Schemske, D.W., Hancock, J.F., Thompson, J.N., Husband, B.C., Judd, W.S., 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? Taxon 56, 13–30.
Song, K., Lu, P., Tang, K., Osborn, T.C., 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. PNAS 92, 7719–7723. https://doi.org/10.1073/pnas.92.17.7719
Song, K., Tang, K., Osborn, T.C., 1993. Development of synthetic *Brassica* amphidiploids by reciprocal hybridization and comparison to natural amphidiploids. Theoret. Appl. Genetics 86, 811–821. https://doi.org/10.1007/BF00212606
Sora, D., Kron, P., Husband, B.C., 2016. Genetic and environmental determinants of unreduced gamete production in *Brassica napus, Sinapis arvensis* and their hybrids. Heredity 117, 440–448. https://doi.org/10.1038/hdy.2016.69
Srivastava, A., Gupta, V., Pental, D., Pradhan, A.K., 2001. AFLP-based genetic diversity assessment amongst agronomically important natural and some newly synthesized lines of *Brassica juncea*. Theor Appl Genet 102, 193–199. https://doi.org/10.1007/s001220051635
Stebbins, G.L., 1971. Chromosomal evolution in higher plants. Chromosomal evolution in higher plants.
Szadkowski, E., Eber, F., Huteau, V., Lodé, M., Coriton, O., Jenczewski, E., Chèvre, A.M., 2011. Polyploid formation pathways have an impact on genetic rearrangements in resynthesized *Brassica napus*. New Phytologist 191, 884–894. https://doi.org/10.1111/j.1469-8137.2011.03729.x
Szadkowski, E., Eber, F., Huteau, V., Lodé, M., Huneau, C., Belcram, H., Coriton, O., Manzanares-Dauleux, M.J., Delourme, R., King, G.J., Chalhoub, B., Jenczewski, E., Chèvre, A.-M., 2010. The first meiosis of resynthesized *Brassica napus*, a genome blender. New Phytologist 186, 102–112. https://doi.org/10.1111/j.1469-8137.2010.03182.x
Tayalé, A., Parisod, C., 2013. Natural Pathways to Polyploidy in Plants and Consequences for Genome Reorganization. CGR 140, 79–96. https://doi.org/10.1159/000351318
te Beest, M., Roux, L., J, J., Richardson, D.M., Brysting, A.K., Suda, J., Kubešová, M., Pyšek, P., 2012. The more the better? The role of polyploidy in facilitating plant invasions. Ann Bot 109, 19–45. https://doi.org/10.1093/aob/mcr277
U, N., 1935. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Jpn J Bot 7, 389–452.
Wang, J., Tian, L., Madlung, A., Lee, H.-S., Chen, M., Lee, J.J., Watson, B., Kagochi, T., Comai, L., Chen, Z.J., 2004. Stochastic and Epigenetic Changes of Gene Expression in *Arabidopsis* Polyploids. Genetics 167, 1961–1973. https://doi.org/10.1534/genetics.104.027896
Warwick, S.I., Légère, A., Simard, M.-J., James, T., 2008. Do escaped transgenes persist in nature? The case of an herbicide resistance transgene in a weedy *Brassica rapa* population. Molecular Ecology 17, 1387–1395. https://doi.org/10.1111/j.1365-294X.2007.03567.x

Warwick, S.I., Martin, S.L., 2013. Gene flow from transgenic oilseed *Brassica juncea* (L.) Czern. into weedy *Sinapis arvensis* L. (wild mustard). Plant Breed 132, 688–693. https://doi.org/10.1111/pbr.12098

Warwick, S.I., Simard, M.-J., Légère, A., Beckie, H.J., Braun, L., Zhu, B., Mason, P., Séguin-Swartz, G., Stewart, C.N., 2003. Hybridization between transgenic *Brassica napus* L. and its wild relatives: *Brassica rapa* L., *Raphanus raphanistrum* L., *Sinapis arvensis* L., and *Erucastrum gallicum* (Willd.) O.E. Schulz. Theor Appl Genet 107, 528–539. https://doi.org/10.1007/s00122-003-1278-0

Werth, C.R., Windham, M.D., 1991. A Model for Divergent, Allopatric Speciation of Polyploid Pteridophytes Resulting from Silencing of Duplicate-Gene Expression. The American Naturalist 137, 515–526. https://doi.org/10.1086/285180

Whitton, J., Wolf, D.E., Arias, D.M., Snow, A.A., Rieseberg, L.H., 1997. The persistence of cultivar alleles in wild populations of sunflowers five generations after hybridization. Theor Appl Genet 95, 33–40. https://doi.org/10.1007/s001220050529

Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B., Rieseberg, L.H., 2009. The frequency of polyploid speciation in vascular plants. PNAS 106, 13875–13879. https://doi.org/10.1073/pnas.0811575106

Zhang, G.Q., Tang, G.X., Song, W.J., Zhou, W.J., 2004. Resynthesizing *Brassica napus* from interspecific hybridization between *Brassica rapa* and *B. oleracea* through ovary culture. Euphytica 140, 181–187. https://doi.org/10.1007/s10681-004-3034-1

Zhang, L.-Q., Liu, D.-C., Zheng, Y.-L., Yan, Z.-H., Dai, S.-F., Li, Y.-F., Jiang, Q., Ye, Y.-Q., Yen, Y., 2010. Frequent occurrence of unreduced gametes in Triticum turgidum–Aegilops tauschii hybrids. Euphytica 172, 285–294. https://doi.org/10.1007/s10681-009-0081-7

Zhang, X., Liu, T., Li, X., Duan, M., Wang, J., Qiu, Y., Wang, H., Song, J., Shen, D., 2016. Interspecific hybridization, polyploidization, and backcross of *Brassica oleracea var. alboglabra* with *B. rapa* var. *purpurea* morphologically recapitulate the evolution of *Brassica* vegetables. Scientific Reports 6, srep18618. https://doi.org/10.1038/srep18618