

THE ROLE OF *CAROTENOID CLEAVAGE DIOXYGENASE 4* IN FLOWER COLOR  
OF THE ALLOPOLYPLOID *BRASSICA NAPUS*

A Thesis

Presented to

The Faculty of California Polytechnic State University,  
San Luis Obispo

In Partial Fulfillment

of the Requirements to the Degree

Master of Science in Biology

by

Leanne Denice Fogg

August, 2014

© 2014

Leanne Denice Fogg

ALL RIGHTS RESERVED

## COMMITTEE MEMBERSHIP

TITLE: The role of *CAROTENOID CLEAVAGE DIOXYGENASE 4* in flower color of the allopolyploid *Brassica napus*

AUTHOR: Leanne Denice Fogg

DATE SUBMITTED: August, 2014

COMMITTEE CHAIR: Ed Himelblau, PhD  
Associate Professor of Biological Sciences

COMMITTEE MEMBER: Michael Black, PhD  
Professor of Biological Sciences

COMMITTEE MEMBER: Ken Hillers, PhD  
Associate Professor of Biological Sciences

## ABSTRACT

### The role of *CAROTENOID CLEAVAGE DIOXYGENASE 4* in flower color of the allopolyploid *Brassica napus*

Allopolyploids are formed by interspecific hybridization and whole genome duplication, with the resulting organism contains multiple distinct subgenomes in one nucleus. Subgenomic interactions result in massive genetic and epigenetic reconstruction, contributing to variable phenotypic traits noted in newly formed allopolyploids. To better understand these mechanisms in the context of molecular biology, evolution, and plant breeding, plant biologists study the model organism *Brassica napus* (farmed as canola or oilseed rape). With white-flowering and yellow-flowering progenitors, flower color phenotype of *B. napus* presents an opportunity to examine subgenomic interactions. *CAROTENOID CLEAVAGE DIOXYGENASE 4* (*CCD4*) is known to play a major role in determining flower color phenotype of carotenoid-synthesizing angiosperms. Here, we investigate the genetic and epigenetic role of *CCD4* orthologs and their role in flower color phenotype of *B. napus*.

Keywords: polyploid, petal color, ortholog, genetics, *CAROTENOID CLEAVAGE DIOXYGENASE 4*, *CCD4*, carotenoids

## ACKNOWLEDGMENTS

I have infinite appreciation for the help, encouragement, and guidance provided by the faculty, staff, and students of the Biological Sciences Department at California Polytechnic State University, San Luis Obispo, CA. Dr. Ed Himelblau deserves exceptional recognition for the remarkable support and contributions he has made to make this work possible. In addition, I am thankful for the resources, facilities, and the College Based funds provided by California Polytechnic State University, San Luis Obispo, CA.

Access to the *Brassica oleracea* sequence, an essential component of this work, was made possible by Patrick Edger at University of Missouri.

I am forever grateful for the love and emotional support from family and friends I have received throughout my academic career.

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
CHAPTER	
1. INTRODUCTION .....	1
1.1. Formation of Allopolyploids.....	1
1.2. Evolution and Gene Fate of <i>Brassica napus</i> .....	2
1.3. Resynthesized <i>Brassica napus</i> and Genomic Reconfiguration.....	3
1.4. Flower Coloration Determined by Carotenoid Pigments.....	4
2. MATERIALS AND METHODS.....	6
2.1. Plant Materials.....	6
2.2. Growth Conditions and Selfed Lineages.....	6
2.3. Petal Color Analysis.....	7
2.4. Genotypic Analysis.....	8
2.4.1. <i>Bo.CCD4.a</i> & <i>Br.CCD4.a</i> Orthologs.....	8
2.4.2. <i>Bo.CCD4.b</i> & <i>Br.CCD4.b</i> Orthologs.....	9
2.5. Expression Analysis of <i>CCD4</i> Orthologs.....	10
3. RESULTS.....	11
3.1. Characterization of Petal Color in <i>B. napus</i> .....	11
3.2. Petal Color in Segregating Populations of <i>B. napus</i> .....	14
3.3. Phylogeny of <i>CCD4</i> Gene Family.....	16
3.4. Genotype of <i>CCD4</i> loci.....	17
3.5. <i>CCD4</i> Expression.....	19

4. DISCUSSION.....	20
4.1. Characterization of Petal Color in <i>B. napus</i> .....	20
4.2. Petal Color in Segregating populations of <i>B. napus</i> .....	21
4.3. Phylogeny of <i>CCD4</i> Gene Family.....	21
4.4. Genotype of <i>CCD4</i> Loci.....	22
4.5. Expression of <i>CCD4</i> orthologs.....	22
5. CONCLUSION.....	24
REFERENCES.....	25
APPENDIX.....	31

## LIST OF TABLES

	Page
Table 1. <i>CCD4</i> genotypes of <i>B. napus</i> parents.....	18
Table 2. Progeny <i>CCD4</i> genotypes.....	18

## LIST OF FIGURES

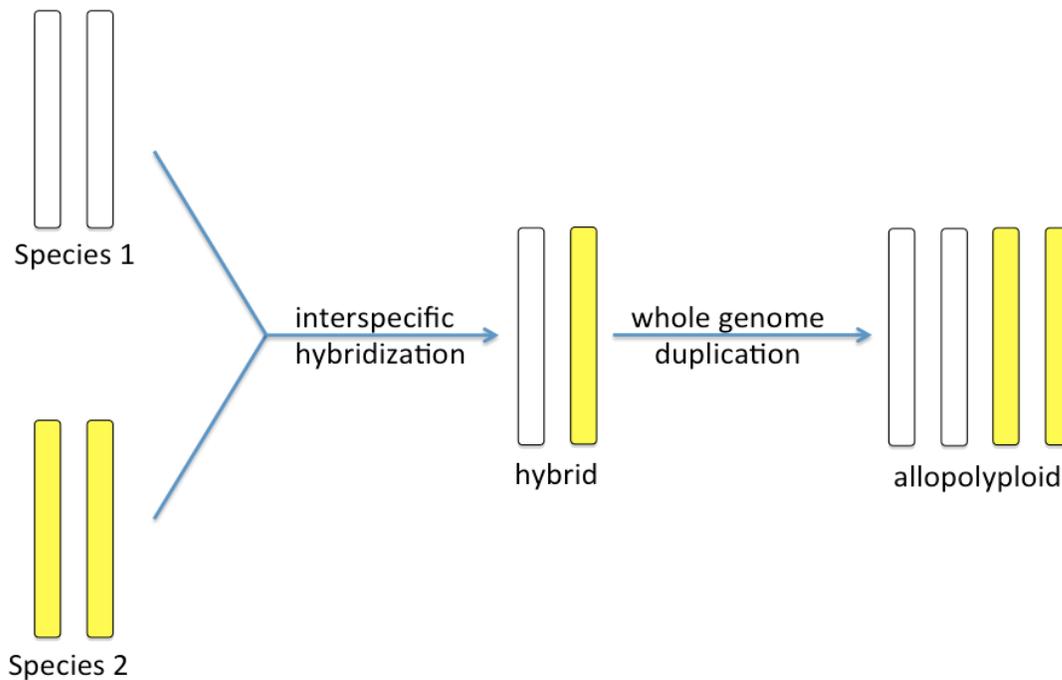
	Page
Figure 1. A possible mechanism of allopolyploid formation.....	2
Figure 2. Evolution of <i>CAROTENOID CLEAVAGE DIOXYGENASE (CCD4)</i> paralogs in <i>Brassica napus</i> .....	5
Figure 3. <i>Brassica napus</i> Selfed Lineage Diagram.....	7
Figure 4. Stages of flower development.....	12
Figure 5. Petal yellowness throughout flower development in <i>B. napus</i> and its progenitors.....	13
Figure 6. Yellowness of stage 6 <i>Brassica napus</i> flowers in comparison to progenitor species.....	15
Figure 7. Yellowness of <i>Brassica napus</i> lineage-specific progeny.....	16
Figure 8. Phylogeny of <i>CCD4</i> gene family.....	17
Figure 9. <i>CCD4</i> within species out-paralog expression <i>B. rapa</i> and <i>B. oleracea</i> assessed in various plant tissues.....	19

## 1. INTRODUCTION

### 1.1. Formation of Allopolyploids

Many plant species, including agronomically important plants like wheat and cotton, show evidence of polyploidy in their evolution (Leitch and Bennett, 1997; Pontes, 2004; Ozkan 2001; Kim 2008; Lim 2008; Schmutz, 2010; Leitch and Leitch, 2014).

Polyploids contain duplicate copies of genomes from the same species (autopolyploids) or from different species (allopolyploids) (Comai, 2005; Chen, 2007). While autopolyploid events involve increasing the copy number of a single genome, allopolyploid formations involves the union and multiplication of two or more distinct genomes (Leitch and Bennett, 1997; Wendel, 2000; Leitch and Leitch 2008). Generally, organisms from different species do not produce viable hybrid offspring (Widmer et al., 2008). However, species-specific fertilization mechanisms can be overcome if hybridization is followed concomitant to whole genome duplication (WGD) (Figure 1). Pre-zygotic (aneuploid gametes) or post-zygotic (non-disjunction after fertilization in germ line tissue) duplication events provide genomic stability for successful interspecific hybridization between two related, yet distinct species (Leitch & Leitch, 2008). Allopolyploid offspring contain multiple copies of each progenitor genome, resulting in their reproductive isolation from either progenitor species (Comai, 2005; Chen, 2007).



**Figure 1: A possible mechanism of allopolyploid formation.** Two species can form an allopolyploid through interspecific hybridization and whole genome duplication. A single chromosome represents the haploid ( $n$ ) genome from species 1 (white) and species 2 (yellow). The fertilization of haploid gametes from species 1 and species 2 results in an unstable hybrid. Whole genome duplication of the hybrid genome results in an allopolyploid, which contains the complete diploid genomes of both progenitor species.

## 1.2. Evolution and Gene Fate of *Brassica napus*

Approximately 20 million years ago (mya) *Brassica* diverged from *Arabidopsis* genera (Yang et al., 1999). This was followed by a genome triplication event in the *Brassica* lineage approximately 16 mya, resulting in a hexaploid *Brassica* genus (Yang et al, 1999; Rana et al., 2004). Slight genomic changes throughout the hexaploid *Brassica* genus resulted in the diversification of *Brassica* over evolutionary time, giving

rise to various species (Lysak et al, 2005). Allopolyploidization between *Brassica* plant taxa reunites distinct progenitor genomes, resulting in new species such as *Brassica napus* (U, 1935; Becker, 1995; Song et al., 1995; Yang et al., 1999).

The interspecific hybridization and WGD of *Brassica rapa* and *Brassica oleracea* results in the *Brassica napus* allopolyploid. *Brassica napus* plants contain 19 pairs of homologous chromosomes – 9 pairs of chromosomes from genome “C” (*B. oleracea*), and 10 pairs of chromosomes from genome “A” (*B. rapa*) (Parkin et al., 1995). Although newly formed *B. napus* plants contain the complete diploid genomes of progenitor species, subsequent genetic changes can alter genome structure and therefore affect the way in which traits are manifested in the allopolyploid.

### 1.3. Resynthesized *Brassica napus* and Genomic Reconfiguration

As a result of subgenome interactions, current *B. napus* plants contain altered gene doses and hybrid chromosomes (Lagercrantz, 1998). In order to study the genetic mechanisms which gives rise to the current allopolyploid, scientists use a “resynthesized” form of *B. napus*, which contains the entire un-hybridized diploid genomes of both progenitor species (Gaeta and Piers, 2010). Scientists use this form to better understand the genetic and epigenetic reconstruction processes which occur post-hybridization.

Resynthesized *B. napus* undergo massive genomic reconstruction during early generations, resulting in offspring with altered genome structure (McClintock, 1983; Wendel, 2000; Wang et al., 2004; Chen, 2007; Gaeta and Pires, 2010). The high degree of homeology between subgenomes can affect chromosome pairing and subsequent recombination events during meiosis. Homeologous pairing and

recombination can result in the loss of a locus from a subgenome or changes in dosage between the two subgenomes (e.g. CCAA changing to CCCA) (Gaeta *et al.*, 2007).

A process referred to as “diploidization” is also responsible for the alteration of gene copy numbers. Diploidization is a process, which reduces gene redundancy, reverting the polyploid back to a diploid form (Wolfe, 2001). Paralogs (genes within the same genome related by a duplication event) are lost due to the evolution of pseudogenes (essential loss of function) or, less commonly, by the evolution of new gene functions (Wendel, 2000; Adams, 2007; Buggs *et al.*, 2012). Altered gene dosage and genomic structure can result in the evolution and variation of phenotypic traits.

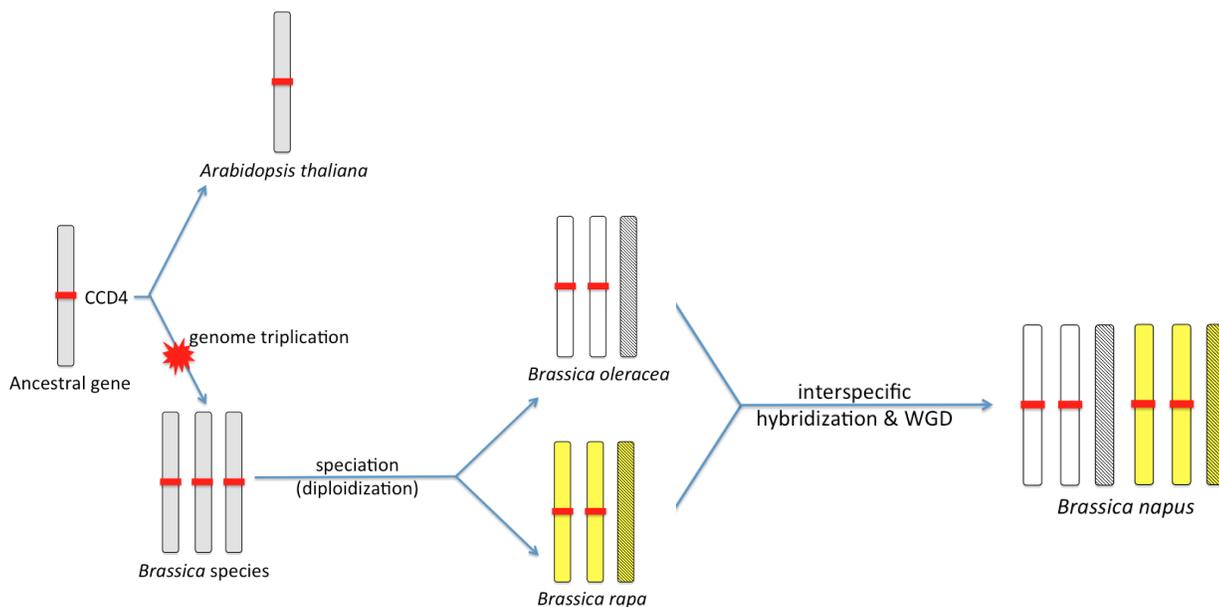
#### 1.4. Flower Coloration Determined by Carotenoid Pigments

The flower color of *Brassica napus* presents an opportunity to study the interaction between subgenomes in an allopolyploid. Flower color variation was evident in the selfed progeny from a single *B. napus* parent, ranging from white to bright yellow. This variation is thought to be genetically based since all progeny were observed under controlled environmental conditions.

Plants of the *Brassica* genera tend to have either white or yellow petals. This coloration is largely due to the presence or absence of carotenoid pigments - isoprenoid compounds responsible for the yellow to red coloration in many plant tissues (Zhu *et al.*, 2010). The accumulation of carotenoid pigments in chromoplasts of petal cells results in the yellow coloration of flower petals. However, when the enzyme CAROTENOID CLEAVAGE DIOXYGENASE 4 (CCD4) degrades carotenoids, the color intensity decreases, resulting in pale yellow or even white plant tissue (e.g. petals) (Auldridge *et al.*, 2006; Yamamizo *et al.*, 2010; Adami *et al.*, 2013; Gonzalez-Jorge *et al.*, 2013). In

chrysanthemum, functional alleles of *CCD4* are associated with white flowers and loss-of-function alleles of *CCD4* are associated with yellow and red flowers (Ohmiya, 2013).

The *B. rapa* and *B. oleracea* genomes have recently been sequenced and both appear to contain two paralogous copies of *CCD4* (Chang et al, 2011). Therefore, resynthesized *B. napus* is expected to have four orthologous *CCD4* loci (Figure 2). The interactions between all orthologs are likely to have a large effect on the flower color phenotype of the allopolyploid. Prior studies have demonstrated a strong genetic role of one *CCD4* ortholog for determining flower color in *Brassica oleracea* (Quezada et al., unpublished). Here we investigate subgenomic contributions/interactions of *CCD4* orthologs in *B. napus* and its relationship to petal color.



**Figure 2: Evolution of CAROTENOID CLEAVAGE DIOXYGENASE (*CCD4*) paralogs in *Brassica napus*.** Chromosomes and noted loci (red marks) depicted represent the number of *CCD4* paralogs in haploid cells for a given species/genera. A genome triplication event generated three *CCD4* paralogs in *Brassica* species. Over evolutionary time, genome reorganization in *Brassica* species resulted in the functional loss of *CCD4* paralogs (pictured by crossed out chromosomes). Newly synthesized *Brassica napus* individuals contain four functional *CCD4* paralogs (two from each progenitor genomes - *B. rapa* and *B. oleracea*).

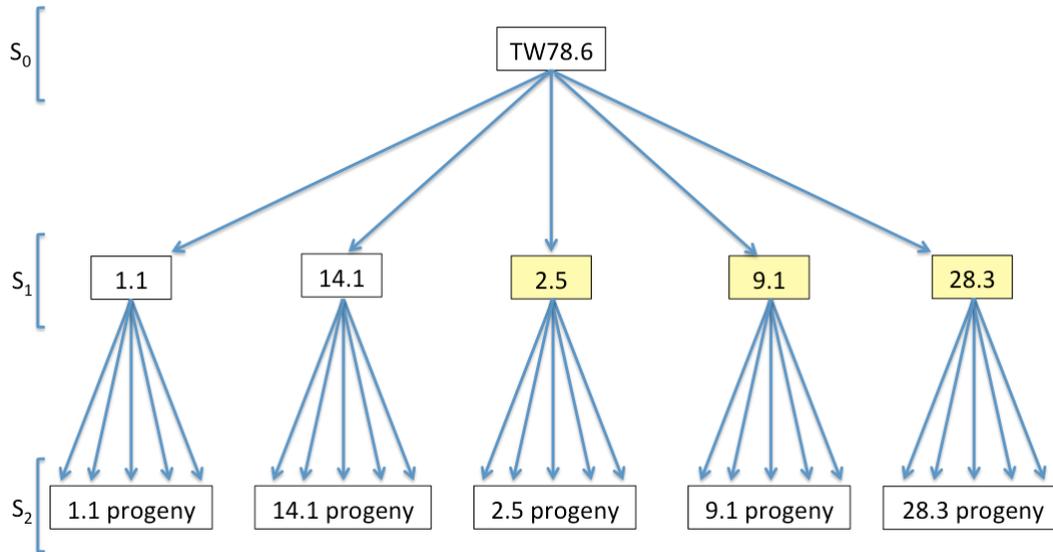
## 2. MATERIALS AND METHODS

### 2.1. Plant Materials

Resynthesized *Brassica napus* individuals were prepared through the interspecific hybridization of *Brassica oleracea* (TO1000; egg donor; genome “C”) and *Brassica rapa* (IMB218; pollen donor; genome “A”) as described by Lukens et al. (2006). Colchicine treatment induced whole genome duplication of the CA hybrids, resulting in uniformly homozygous “CCAA” *B. napus* individuals. Each individual gave rise to its own lineage via controlled self-pollination. Selfed offspring of TW78.6 parent showed notable variation in flower color (Figure 3). A selfed generation ( $S_2$ ) was grown from select  $S_1$  individuals for genotypic and phenotypic analyses.

### 2.2. Growth Conditions and Selfed Lineages

Seeds were planted in Miracle Gro Premium Potting Mix® soil and topped with vermiculite at California Polytechnic State University, San Luis Obispo, CA, USA. Plants were watered and treated with a pesticide (Adept, OHP Inc., Mainland, PA, USA) and were exposed to an 12 hour photoperiod (6AM – 10PM) under white fluorescent light. Approximately 3 weeks post-germination, individuals were transplanted into larger pots and either remained in the plant room or relocated to the greenhouse at California Polytechnic State University, San Luis Obispo, CA, USA. All plants were self-pollinated as previously described by Gaeta et al. (2007). Viable seeds were collected from each plant. Seeds from a select group of individuals were planted for the  $S_2$  generation.



**Figure 3: *Brassica napus* selfed lineage diagram.** The selfed progeny of a single parent (TW78.6) showed variation in flower color (flower color phenotype represented by white/yellow coloration). Selfed progeny of five  $S_1$  individuals were grown up for genotypic and phenotypic analyses ( $S_2$  individuals).

### 2.3. Petal Color Analysis

Fully open blossoms flowers were collected from controls (TO1000, IMB218, resynthesized *B. napus*), select  $S_1$  individuals (1.1, 14.1, 2.5, 9.1, 28.3), and their progeny ( $S_2$ ) for petal analysis. Petals were taped down with scotch tape onto blue stationery and scanned (with a Canon ImageRunner Advance 6065) immediately after collection. Yellow “sticky-notes” from a single pad were used throughout petal collections as a control reference to ensure uniform color scans. The same Scotch® tape and blue stationery paper was used consistently throughout petal analysis. Yellowness of petals were assessed by obtaining a Y value with the eyedropper tool (CMYK values) on Adobe Photoshop® CS5 Extended (Version 12.1 x64). The Y value was noted for each scanned petal and averaged for each individual. The standard error was calculated as the standard error of the mean.

## 2.4. Genotypic Analysis

Non-cotyledon leaves were collected for DNA isolations. Genomic DNA was isolated from leaf tissue with the Qiagen® Plant Mini-Prep: DNA Extraction Kit (Qiagen; Valencia, CA, USA). DNA isolates were diluted to working concentrations and followed by genotyping procedures.

In order to genotype *B. napus* individuals for all *CCD4* orthologs, two procedures were developed using published *B. rapa* Chiifu-401 (Chen et al., 2001) and unpublished *B. oleracea* TO1000 sequences (Pat Edger, personal communication). The first procedure was designed to detect the presence of *Bo.CCD4.a* and *Br.CCD4.a* orthologs. dCAPS markers and Sall restriction enzyme was used to distinguish between PCR amplified orthologs. The second procedure targeted *Bo.CCD4.b* and *Br.CCD4.b* orthologs; PCR amplified products varied by size. Gel electrophoresis was used to distinguish between PCR amplified orthologs for both procedures.

### 2.4.1. *Bo.CCD4.a* & *Br.CCD4.a* Orthologs

A portion of the *Bo.CCD4.a* and *Br.CCD4.a* coding region was amplified with dCAPS markers, incorporating a Sall cut site into the *Br.CCD4.a* amplified product during polymerase chain reaction (PCR). DNA samples were prepared for PCR under the following conditions: 1.0U of GoTaq Flexi DNA Polymerase (Promega Corporation; Madison, WI, USA), 5X Green GoTaq Flexi Buffer, 25mM MgCl<sub>2</sub>, 2.5μM dNTPs, 10μM Sall dCAPS forward primer, 10μM Sall dCAPS reverse primer, 10ng of plant DNA, and diH<sub>2</sub>O to a 20μL final volume. The prepared solutions underwent PCR according to the Brassica 22 cycle PCR protocol: 94°C for 3 minutes; 8 cycles of 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 50 seconds; 22 cycles of 89°C for 30 seconds, 50°C for 30 seconds, 72°C for 50 seconds; 72°C for 3 minutes. Samples were stored at 4°C.

Following PCR, samples were digested with Sall enzyme (New England BioLabs Inc.) to differentiate between the *Bo.CCD4.a* and *Br.CCD4.a* orthologs. The digest samples were prepared under the following conditions: 2 $\mu$ L of PCR product, 4U of Sall, and diH<sub>2</sub>O to a 10 $\mu$ L final volume. For undigested control samples, the Sall enzyme was replaced with an equivalent volume of water. After digestion, a 363bp (cut) and 379bp (uncut) product corresponded to the presence of the *Br.CCD4.a* and *Bo.CCD4.a* ortholog, respectively.

Digested samples and all controls were diluted in 6X loading dye (Promega® Blue/Orange 6X Loading Dye) and diH<sub>2</sub>O, and loaded into a 2% agarose gel prepared with 1xTAE (40mM Tris, 20mM acetic acid, and 1mM EDTA). Bands were resolved on acrylamide gels.

The presence/absence of *Bo.CCD4.a* and *Bra013378* orthologs was noted for each individual (*Bo.CCD4.a* undigested band size: 379bp; *Br.CCD4.a* digested band size: 363bp).

#### 2.4.2. *Bo.CCD4.b* & *Br.CCD4.b* Orthologs

DNA samples were prepared for PCR under the previously described conditions. PCR amplified *Bo.CCD4.b* and *Br.CCD4.b* products contained DNA sequence located upstream each ortholog as well as a portion of the coding regions. PCR was performed under the same conditions as the previously described protocol, Brassica 22 cycle.

PCR amplified samples were diluted in 6X loading dye (Promega® Blue/Orange 6X Loading Dye) and diH<sub>2</sub>O, and loaded into a 2% agarose gel prepared with 1xTAE (40mM Tris, 20mM acetic acid, and 1mM EDTA). Bands were resolved on acrylamide gels. The presence/absence of *Bo.CCD4.b* and *Br.CCD4.b* orthologs was noted for

each individual (*Bo.CCD4.b* amplified product: 374bp; *Br.CCD4.b* amplified product: 354bp).

## 2.5. Expression Analysis of *CCD4* Orthologs

Plant tissue was harvested for RNA isolations from the following groups of plants: *B. rapa*, *B. oleracea*, *B. napus* white-flowering, and *B. napus* yellow-flowering individuals. A non-cotyledon leaf or inflorescence stock was selected from three individuals and pooled by group and tissue type. RNA isolations were performed with the Qiagen RNeasy® Plant Mini Kit (Qiagen; Valencia, CA, USA). A cDNA library was built from RNA isolates using New England BioLabs ProtoScript® II First Strand cDNA Synthesis Kit (New England BioLabs Inc.; Ipswich, MA, USA). cDNA libraries were synthesized as described by New England BioLabs Protoscript® II First Strand cDNA Synthesis Kit protocol with the addition of Oligo-dT primers [d(T)<sub>23</sub>VN].

RNA seq datasets are available for *B. oleracea* (Yu et al., 2014) and *B. rapa* (Tong et al., 2013). Expression of both *CCD4* orthologs from each species were determined for six tissues in the case of *B. oleracea* and 8 tissues in the case of *B. rapa*. Results were plotted using RStudio (Version 0.98.507, 2013).

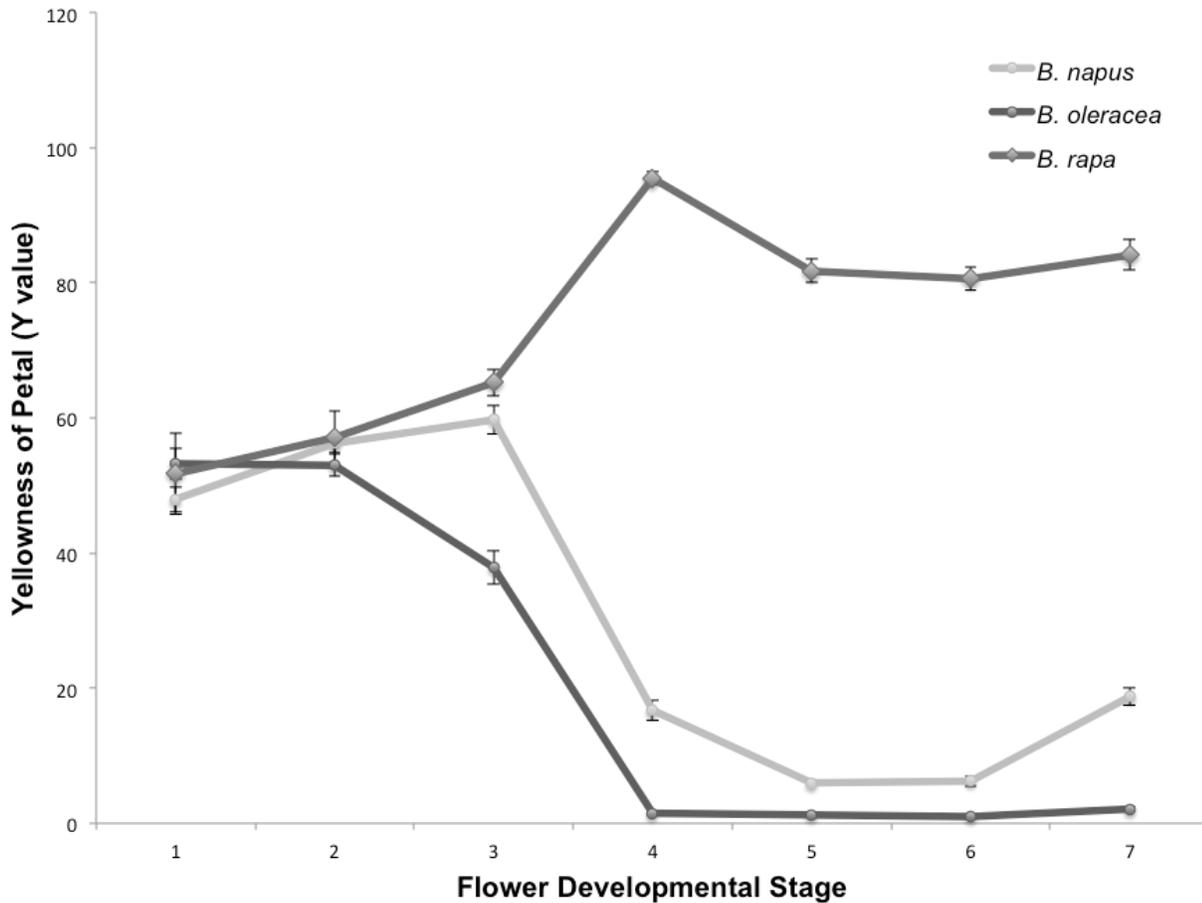
### 3. RESULTS

#### 3.1. Characterization of Petal Color in *B. napus*

Various stages of flower development were identified for each *Brassica* species to consider petal color variation during development (Figure 4). To compare progenitor and allopolyploid petal color phenotypes, petals were collected from *B. oleracea* (TO1000), *B. rapa* (IMB218), and resynthesized *B. napus* (TO1000 egg donor; IMB218; pollen donor) individuals and analyzed with Photoshop®. A time series of petal yellowness was generated to observe the accumulation and degradation of yellow pigmentation during flower development for each *Brassica* species (Figure 5). All *Brassica* species had similar initial carotenoid abundances (stage 1 and stage 2; non significant). Divergence of petal color began at stage 3, where *B. napus* and *B. rapa* were significantly more yellow than *B. oleracea* (F ratio 41.06;  $p < 0.0001$ ). By stage 4, all *Brassica* species exhibited distinct phenotypes: *B. rapa* was the most yellow, *B. oleracea* was the least yellow, while *B. napus* exhibited an intermediate yellow color phenotype (F ratio 3961.9;  $p < 0.0001$ ). *B. napus* and *B. oleracea* continued exhibiting significantly lower yellow values than *B. rapa* at stage 5 (F ratio 1488.9;  $p < 0.0001$ ). All *Brassica* species exhibited distinct yellowness by stage 6 (F ratio 1759.4;  $p < 0.0001$ ) and stage 7 (F ratio 845.8;  $p < 0.0001$ ).



**Figure 4: Stages of flower development.** Stages of flower development were identified for each *Brassica* species (top: *Brassica rapa*; middle: *Brassica napus*; bottom: *Brassica oleracea*). Stages of development are noted below a representative flower image, ranging from young bud (stage 1, left) to senescing flower (stage 7, right).



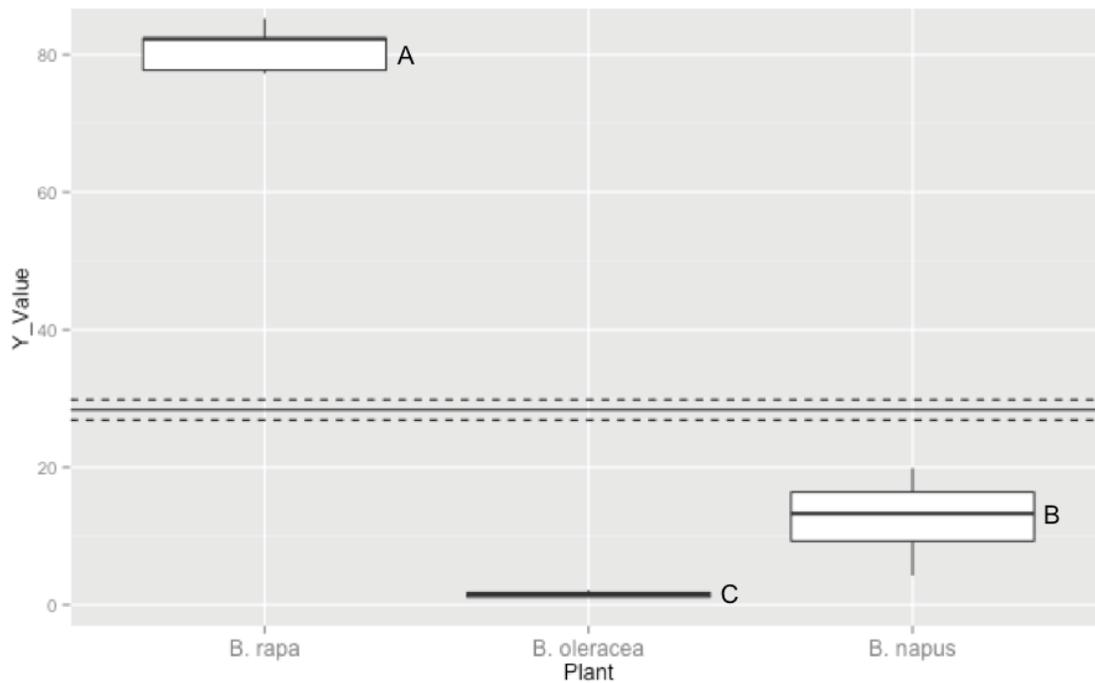
**Figure 5: Petal yellowness throughout flower development in *B. napus* and its progenitors.** Y values of petals (two flowers) were determined for each representative *Brassica* species individual, *B. rapa*, *B. oleracea*, and resynthesized *B. napus*, throughout flower development (stages 1 - 7). The standard error was calculated as the standard error of the mean.

Petal color of fully opened (stage 6) flowers from multiple individuals of each *Brassica* species was assessed with Photoshop® (Figure 6). *Brassica napus* exhibited distinct yellowness from either progenitor, but was more similar to the yellowness observed in *B. oleracea* individuals. Flower color phenotype in the allopolyploid appears to be non-additive - the midparent value (MPV) of petal yellowness from progenitor

species was significantly different than the yellowness observed in *B. napus* (F ratio 363.2;  $p < 0.0001$ ).

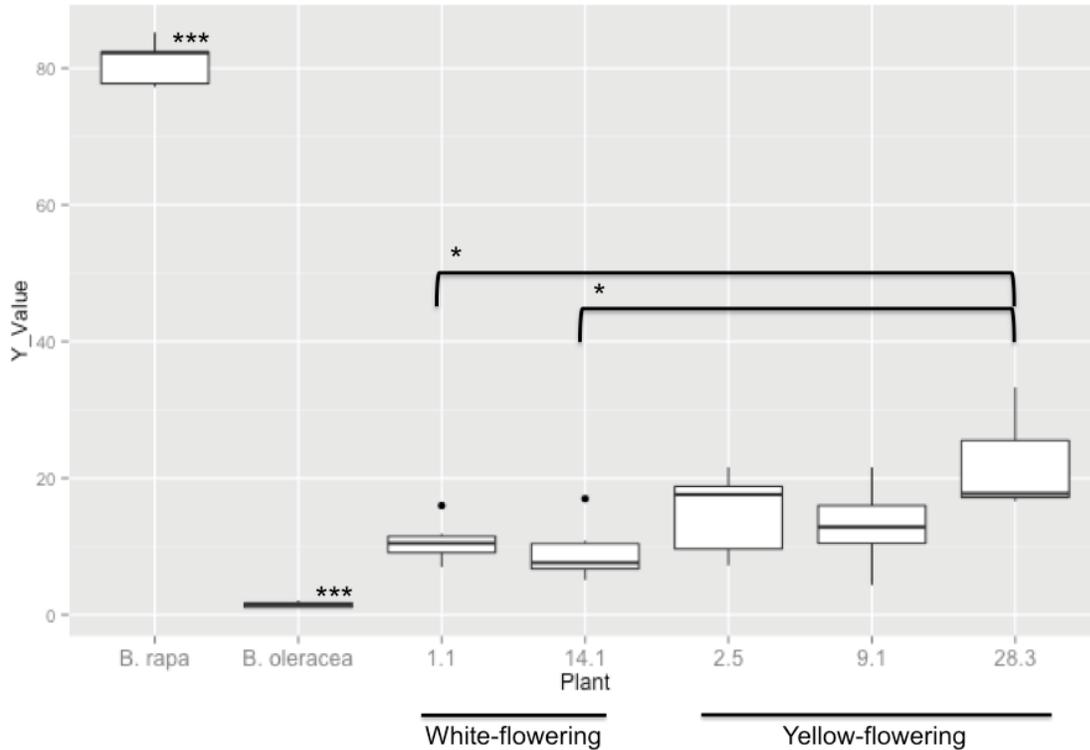
### 3.2. Petal Color in Segregating Populations of *B. napus*

Previously, yellow-flowered plants were identified among a populations of sibling plants derived from a single *B. napus* parent (Wang and Himmelblau, unpublished). Selfed seed from both yellow-flowered and white-flowered individuals were grown. The self-pollination process was repeated for a single individual, generating selfed lineages 1.1, 14.1, 2.5, 9.1, and 28.3. Petal color was determined for viable progeny (i.e. progeny that survived to flowering).



**Figure 6: Yellowness of stage 6 *Brassica napus* flowers in comparison to progenitor species.**

The Y values of stage 6 petals from five to nine individuals were determined for each *Brassica* species box plot. Each *Brassica* species exhibited distinct yellow-colored phenotypes (F ratio 934.1956;  $p < 0.0001$ ). Groups not connected by same letter are significantly different ( $p < 0.0001$ ). The midparent value (MVP) represents the expected additive phenotype from progenitor species (solid horizontal line; standard error of the estimate is represented by dashed horizontal lines). *Brassica napus* phenotype is significantly different than the MVP, supporting the non-additive genetic basis of flower color phenotype (effect contrast test; F ratio 363.2;  $p < 0.0001$ ). The standard error was calculated as the standard error of the mean for each *Brassica* species.

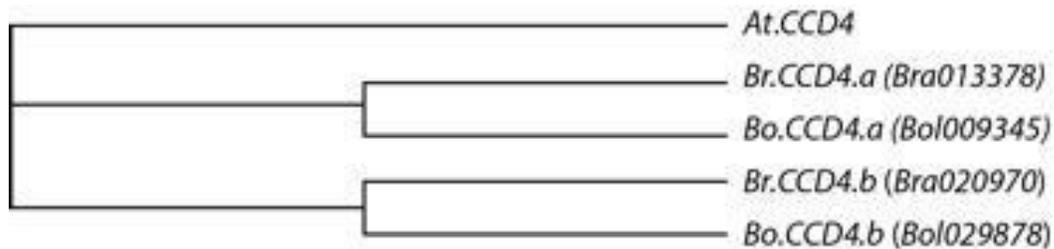


**Figure 7: Yellowness of *Brassica napus* lineage-specific progeny.** The Y values of stage 6 petals from three to nine individuals were determined for *Brassica* species controls (*B. rapa* and *B. oleracea*) box plots. Two stage 6 flowers were collected from lineage progeny to determine lineage-specific box plots of Y values. Progenitors exhibited significantly different petal color phenotypes from all *B. napus* lineage progeny ( $p < 0.0001$ ). Yellow-flowering 28.3 lineage had significantly higher Y values than 14.1 ( $p = 0.0062$ ) and 1.1 ( $p = 0.0288$ ) white-flowering lineages (Tukey-Kramer Post Hoc). The standard error was calculated as the standard error of the mean for each *Brassica* species.

### 3.3. Phylogeny of *CCD4* Gene Family

A local phylogeny was generated for *CCD4* gene family members using published and unpublished sequences (Figure 8). The gene family includes two *CCD4* out-paralogs in each *B. rapa* (*Br.CCD4.a*, *Br.CCD4.b*) and *B. oleracea* (*Bo.CCD4.a*, *Bo.CCD4.b*) species. Orthologs, specifically ohnologs (paralogs by WGD), found in both *B. rapa* and *B. oleracea* cluster together, supporting their origin by WGD. A single

*Arabidopsis thaliana* ortholog (*At.CCD4*) clusters as an outgroup, most distantly related to the *Brassica CCD4* gene family members found in *B. rapa* and *B. oleracea*.



**Figure 8: Phylogeny of *CCD4* gene family.** Local phylogeny of *CCD4* gene family members created with in-frame amino acid sequences (from published and unpublished sequence data) using ClustalW2 (European Molecular Biology Laboratory - The European Bioinformatics Laboratory, Cambridgeshire, UK). Between subgenome out-paralogs cluster together; *A. thaliana CCD4* ortholog is an outgroup.

### 3.4. Genotype of *CCD4* loci

Ortholog-specific genetic markers were developed to genotype *B. napus* individuals for the following *CCD4* gene family members: *Bo.CCD4.a*, *Bo.CCD4.b*, *Br.CCD4.a*, *Br.CCD4.b*. *B. rapa*, *B. oleracea*, and synthetic hybrid (50:50 mixture of *B. rapa* and *B. oleracea* genomic DNA) were used as controls throughout genotyping procedures. An ortholog was determined present with at least one successful detection of expected bands by gel electrophoresis; an ortholog was determined missing if it went undetected during two or more separate replications of genotyping experimentation.

All four *CCD4* orthologs were detected in  $S_0$  (TW78.6) and  $S_1$  individuals (Table 1). Progeny of  $S_1$  individuals were genotyped for the same *CCD4* gene family members (Table 2). Only two progeny of the yellow-flowering 2.5 lineage lacked an ortholog, namely *Bo.CCD4.a*.

**Table 1: *CCD4* genotypes of *B. napus* parents.** Ortholog-specific genetic markers were used to genotype *B. napus* S<sub>1</sub> individuals by the absence (-) or presence (+) of expected band sizes. The four *CCD4* orthologs were detected in each S<sub>1</sub> individual, regardless of flower color.

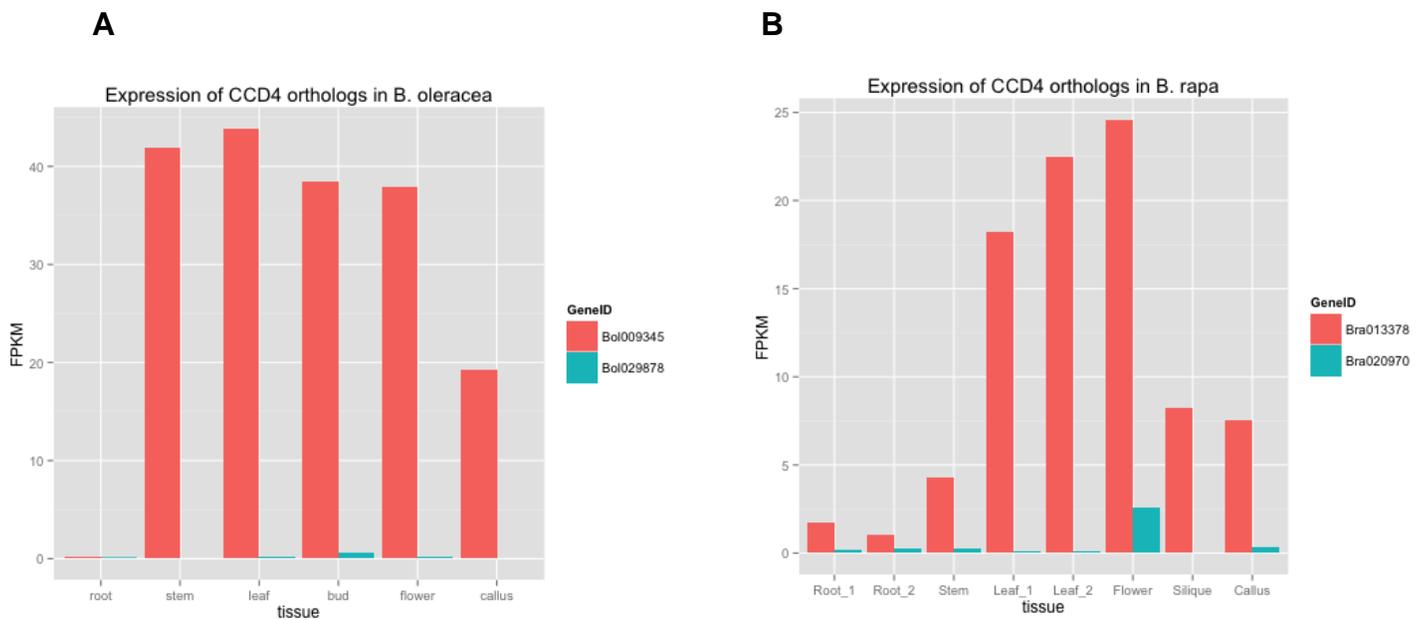
Parent	Lineage Color	<i>Bo.CCD4.a</i>	<i>Br.CCD4.a</i>	<i>Bo.CCD4.b</i>	<i>Br.CCD4.b</i>
1.1	white	+	+	+	+
14.1	white	+	+	+	+
2.5	yellow	+	+	+	+
9.1	yellow	+	+	+	+
28.3	yellow	+	+	+	+

**Table 2: Progeny *CCD4* genotypes.** Ortholog-specific genetic markers were used to genotype *B. napus* S<sub>2</sub> individuals by the absence (-) or presence (+) of expected band sizes. The presence of each *CCD4* ortholog is displayed as the proportion of progeny with the ortholog from a given parent. The four *CCD4* orthologs were detected in all S<sub>2</sub> individuals, with the exception of the *Bo.CCD4.a* ortholog in 2 progeny of the yellow-flowering parent 2.5.

Parent	Lineage Color	<i>Bo.CCD4.a</i>	<i>Br.CCD4.a</i>	<i>Bo.CCD4.b</i>	<i>Br.CCD4.b</i>
1.1	white	7/7	7/7	7/7	7/7
14.1	white	7/7	7/7	7/7	7/7
2.5	yellow	5/7	7/7	7/7	7/7
9.1	yellow	9/9	9/9	9/9	9/9
28.3	yellow	3/3	3/3	3/3	3/3

### 3.5. CCD4 Expression

Expression of *CCD4* orthologs was examined using published RNAseq datasets. In both *B. rapa* and *B. oleracea*, one ortholog of *CCD4* is expressed while the other is mostly silent (Figure 9). The highly expressed orthologs are *Bra013378* (*Br.CCD4.a*) and *Bol009345* (*Bo.CCD4.a*). Little *CCD4* RNA is detected in roots in either species. *CCD4* RNA is detected in leaf, flower, silique, and callus.



**Figure 9: CCD4 within species out-paralog expression *B. rapa* and *B. oleracea* assessed in various plant tissues. CAROTENOID CLEAVAGE DIOXYGENASE 4 (CCD4) out-paralog expression in *B. rapa* (A) and *B. oleracea* (B). *Bra013378* (*Br.CCD4.a*) and *Bol009345* (*Bo.CCD4.a*) are dominantly expressed in all plant tissues.**

## 4. DISCUSSION

### 4.1. Characterization of Petal Color in *B. napus*

Resynthesized *B. napus* exhibited distinct, non-additive petal color phenotype when compared to progenitor species. The midparent value from progenitor species phenotypes was significantly different than the allopolyploid phenotype. In the resynthesized *B. napus* analyzed here, the *B. oleracea* orthologs appears to be partially epistatic to the *B. rapa* orthologs. These findings suggest that, at the *CCD4* locus, the C subgenome is genetically dominant or has significantly higher expression than the A subgenome in *B. napus*.

Subgenome dominance appears to be a common phenomena among polyploid plants. In a typical polyploid, one subgenome (the “dominant” subgenome) will have higher expression and greater gene retention, while other subgenomes are silent and subject to pseudogenization and gene loss (Cheng et al, 2012). Here, the non-additive appearance of petal color phenotype in *B. napus* is consistent with the concept of subgenome dominance, where the loci of one subgenome, subgenome C, is dominant over the other, subgenome A.

Flower color in Brassica is largely determined by the interplay between carotenoid synthesis and carotenoid breakdown by *CCD4*. Each *Brassica* species exhibited distinct carotenoid accumulation and degradation patterns. In the TO1000 line of *B. oleracea*, carotenoids accumulate in the petals during the early stages of flower development and are broken down prior to flower opening (stage 2-3). The mature petals appear white. In contrast, the IMB218 line of *B. rapa* shows accumulation and maintenance of carotenoids throughout flower development, with a slight decline after flower opening (stage 4-5). In resynthesized *B. napus* (TO1000 and IMB218 progenitors), petal color phenotype is consistent with TO1000, being incompletely

epistatic to the IMB218 at the locus (or loci) that determine petal color. *Brassica napus* petals accumulate carotenoids until about stage 4, after which carotenoids are broken down. Unlike in TO1000, carotenoid breakdown in *B. napus* is not complete and a pale yellow color is visible.

#### 4.2. Petal Color in Segregating Populations of *B. napus*

In a screen for novel phenotypes among *B. napus* siblings, some plants with more yellow petal color were identified (Wang and Himelblau, unpublished). Here, we confirm these plants have greater carotenoid accumulation than populations of *B. napus* (from the same genetic screen) in which no variation for petal color was observed.

The progeny of yellow-flowering parents had significantly more yellow petals than individuals from white-flowering parents. Increased carotenoid pigmentation seen in the yellow-flowered population strongly suggests this phenotype has genetic origin.

Previous studies have demonstrated the significant role of the *CCD4* gene in determining plant tissue coloration. Here, genetic changes at the *CCD4* locus in *B. napus* did not correspond to changes in flower color.

#### 4.3. Phylogeny of *CCD4* Gene Family

Local phylogeny of in-frame translated *CCD4* orthologs confirms out-paralog relationships between *B. rapa* and *B. oleracea*. The two pairs of out-paralogs (termed “a” and “b”) are most similar by ancestral descent from a *Brassica* ancestor. Originated from a whole genome triplication event in the *Brassica* ancestor, these paralogs are referred to as “ohnologs.” Ohnologs may retain similar function, but differential selection pressures allows for variable evolution of each gene copy over time. Between species

out-paralog relationship may explain the dominant expression patterns of the “a” ortholog exhibited in both *B. rapa* and *B. oleracea*.

#### 4.4. Genotype of *CCD4* Loci

Despite flower color variation, individuals did not significantly vary genetically at *CCD4* loci. All four *CCD4* orthologs were detected in almost all individuals ( $S_0$ ,  $S_1$ , and  $S_2$ ). These results suggest large genetic changes at the *CCD4* loci are not responsible for variation in flower color. Genetic markers were not dosage sensitive; partial loss/gain or mutations in *CCD4* orthologs could explain some petal color variation.

Pyrosequencing or dosage-sensitive genetic markers for *CCD4* orthologs may reveal a genetic basis for petal color variation in *B. napus*.

#### 4.5. Expression of *CCD4* orthologs

Although various genetic changes due to mutations and chromosomal translocations can explain gene gain/loss, differences in gene expression contribute to an additional layer of how a phenotype is manifested in an allopolyploid. New allopolyploids are believed to undergo massive epigenetic reconfigurations, including altered methylation, acetylation, and chromatin patterns/configuration (Chen, 2007; Ha et al., 2009). Epigenetic changes are thought to occur rapidly and/or stochastically in new *Arabidopsis* allopolyploids (Wang et al., 2004).

Within each subgenome, one *CCD4* out-paralog (*Bo.CCD4.a* and *Br.CCD4.a*) is dominantly expressed, consistent with examples of subgenome dominance from other polyploids (Woodhouse et al., 2014). Interestingly, expression of the *Bo.CCD4.a* is approximately 2x higher than the expression of *Br.CCD4.a*. Phenotypic analysis of flower color in resynthesized *B. napus* suggests, at least for *CCD4*, a new expression

hierarchy is established in the allopolyploid in which a single *B. oleracea* *CCD4* locus has the principal effect on petal color. *Brassica oleracea* (TO1000) was the egg donor in the resynthesis of the *B. napus* lines, possibly influencing gene epistasis in the allopolyploid. Future investigation of *CCD4* epistasis should include reciprocal resynthesized *B. napus* (i.e. resynthesized plants in which TO1000 was the pollen donor and IMB218 the egg donor) (Wertheim et al., 2013).

Expression of *CCD4* could explain the incomplete epistasis observed for petal color. For example, even though all *CCD4* loci are present in the genome, expression of *CCD4* could be predominantly from *B. oleracea* *CCD4* orthologs, leading to a white phenotype. RNA from developing flower buds and inflorescence tissue was extracted to investigate the role of differential expression of *CCD4* in determining phenotype; however, this analysis was not performed as part of the current work.

## 5. CONCLUSION

Collectively, we have provided insight as to the genetic and epigenetic roles of *CCD4* orthologs in *Brassica napus*. Although large genetic changes at the *CCD4* loci did not strongly explain changes in flower color phenotype, partial dosage shifts (CCAA to CCCA) could contribute to phenotypic variation seen in selfed lineages. Future research should include dosage-sensitive genetic markers to investigate if partial dosage shifts influence flower color phenotype.

Evidence of subgenome dominance was seen at *CCD4* loci in *B. rapa* and *B. oleracea* species. The expression of *CCD4* gene family members should be assessed in *B. napus* to investigate relative orthologous gene expression in an allopolyploid state. To investigate the role of genomic imprinting of pollen and egg donor genomes, genotypic and expressional data should be compared to reciprocally resynthesized *B. napus* lines.

## REFERENCES

- Adami M, Franceschi P, Brandi F, Liverani A, Giovannini D, Rosati C, Tartarini S. 2013. Identifying a Carotenoid Cleavage Dioxygenase (ccd4) Gene Controlling Yellow/White Fruit Flesh Color of Peach. *Plant Molecular Biology Reporter*, 31: 1166–1175.
- Adams KL. 2007. Evolution of duplicate gene expression in polyploid and hybrid plants. *The Journal of Heredity* 98: 136–141.
- Auldridge ME, Block A, Vogel JT, Dabney-Smith C, Mila I, Bouzayen M, Klee HJ. 2006. Characterization of three members of the Arabidopsis carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family. *The Plant Journal : For Cell and Molecular Biology* 45:982–993.
- Becker HC, Engqvist GM, Karlsson B. 1995. Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers. *Theoretical and Applied Genetics* 91: 62–67.
- Buggs, Richard JA, Chamala S, Wu, W Jennifer AT, Schnable PS, Soltis DE, Soltis PS, Barbazuk WB. 2012. Rapid, Repeated, and Clustered Loss of Duplicate Genes in Allopolyploid Plant Populations of Independent Origin. *Current Biology* 22: 248–252.
- Chen ZJ. 2007. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annual Review of Plant Biology* 58: 377–406.
- Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, Pingxia L, Wei H, Wang X. 2011. BRAD, the genetics and genomics database for Brassica plants. *BMC Plant Biology* 11:136.

- Cheng F, Wu J, Fang L, Sun S, Liu B, Lin K, Bonnema G, Wang X. 2012. Biased Gene Fractionation and Dominant Gene Expression among the Subgenomes of *Brassica rapa*. *PLoS ONE* 7(e36442): 1-9.
- Comai L. 2005. The advantages and disadvantages of being polyploid. *Nature Reviews. Genetics* 6: 836–846.
- Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC. 2007. Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. *The Plant Cell* 19: 3403-3417.
- Gaeta RT, and Pires JC. 2010. Homoeologous recombination in allopolyploids: the polyploidy ratchet. *New Phytologist* 186: 18-28.
- Gonzalez-Jorge S, Ha SH, Magallanes-Lundback M, Gilliland LU, Zhou A, Lipka AE, Nguyen YN, Angelovici R, Lin H, Cepela J, Little H, Buell CR, Gore MA, Dellapenna D. 2013. Carotenoid cleavage dioxygenase4 is a negative regulator of  $\beta$ -carotene content in *Arabidopsis* seeds. *The Plant Cell* 25: 4812–26.
- Ha M, Kim E, Chen JZ. 2009. Duplicate genes increase expression diversity in closely related species and allopolyploids. *Proceedings of the National Academy of Sciences U.S.A.* 106: 2295-2300.
- Kim ST, Sultan SE, Donoghue MJ. 2008. Allopolyploid speciation in *Persicaria* (*Polygonaceae*): Insights from a low-copy nuclear region. *PNAS* 105: 12370-12375.
- Lagercrantz U. 1998. Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements *Genetics* 150: 1217-1228.

- Leitch IJ and Bennett MD. 1997. Polyploidy in angiosperms. *Trends in Plant Science*. 12: 470–476.
- Leitch AR and Leitch IJ. 2014. Genomic Plasticity and the Diversity of Polyploid Plants. *Science* 320: 481–483.
- Lim K, Yoong, Soltis DE, Soltis PS, Tate J, Matyasik R, Srubarova H, Kovarik A, Pires JC, Xiong Z, Leitch AR. 2008. Rapid Chromosome Evolution in Recently Formed Polyploids in *Tragopogon* (Asteraceae). *PLoS ONE* 3:1- 13.
- Lysak MA, Koch MA, Pecinka A, Schubert I. 2005. Chromosome triplication found across the tribe *Brassicaceae*. *Genome Research* 15: 516-525.
- McClintock B. 1983. The Significant Responses of the Genome to Challenge. *Science* 226: 792–801.
- Ohmiya A. 2013. Involvement of CCD4 in Determining Petal Color, Carotenoid Cleavage Products. *ACS Symposium Series* 1134: 21-30.
- Ozkan H, Levy AA, Feldman M. 2001. Allopolyploid-Induced Rapid Genome Evolution in the Wheat (*Aegilops-Triticum*) Group. *The Plant Cell* 13: 1735-1747.
- Parkin, Isobel AP, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ. 2005. Segmental Structure of the *Brassica napus* Genome Based on Comparative Analysis With *Arabidopsis thaliana*. *Genetics Society of America* 171: 765–781.
- Pontes O, Neves N, Silva M, Lewis MS, Madlung A, Comai L, Viegas W, Pikaard CS. 2004. Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. *Proceedings of the National Academy of Sciences U.S.A.* 101: 18240-18245.
- Rana D, Boogarrt T, O'Neill CM, Hynes L, Bent E, Macpherson L, Park JY, Lin YP, Bancroft I. 2004. Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. *The Plant Journal* 40: 725-733.

- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya M, Sandhu D, Valliyodan B, Lindquis E, Pet M, Gran D, Shu S, Goodstein D, Barry K, Futrell-Grigg M, Abernath B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang C, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA. 2010. Genome sequence of the palaeopolyploid soybean. *Nature* 463: 178–183.
- Song K, Lu P, Tang K, Osborn TC. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proceedings of the National Academy of Sciences of the United States of America* 92: 7719–7723.
- Tong C, Wang X, Yu J, Wu J, Li W, Huang J, Dong C, Hua W, Liu S. 2013. Comprehensive analysis of RNA-seq data reveals the complexity of the transcriptome in *Brassica rapa*. *BMC Genomics* 14(689): 1-10.
- U N. 1935. Genomic analysis of *Brassica* with special reference to the experimental formation of *B. napus* and its peculiar mode of fertilization. *Japanese Journal of Botany* 7: 389–452.
- Wang J, Tian L, Madlung A, Lee HS, Chen M, Lee JJ, Watson B, Kagochi T, Comai L, Chen ZJ. 2004. Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics*, 167: 1961–73.
- Wendel JF. 2000. Genome evolution in polyploids. *Plant Molecular Biology* 42: 225–49.
- Wertheim B, Beukeboom LW, Zande L. 2013. Polyploidy in animals: effects of gene expression on sex determination, evolution and ecology. *Cytogenetic and Genome Research* 140:256–69.

- Widmer A, Lexer C, Cozzolino S. 2008. Evolution of reproductive isolation in plants. *Heredity* 102: 31-38.
- Woodhouse MR, Cheng F, Pires JC, Lisch D, Freeling M, Wang X. 2014. Origin, inheritance, and gene regulatory consequences of genome dominance in polyploids. *Proceedings of the National Academy of Sciences U.S.A.* 111: 5283-5288.
- Wolfe KH. 2001. Yesterday's Polyploids and the Mystery of Diploidization. *Nature Reviews. Genetics* 2: 333–341.
- Yamamizo C, Kishimoto S, Ohmiya A. 2010. Carotenoid composition and carotenogenic gene expression during *Ipomoea* petal development. *Journal of Experimental Botany* 61: 709–19.
- Yang, Yau-Wen, Lai K, Tai P, Li W. 1999. Rates of Nucleotide Substitution in Angiosperm Mitochondrial DNA Sequences and Dates of Divergence Between *Brassica* and Other Angiosperm Lineages. *Journal of Molecular Evolution* 48: 597–604.
- Yu J, Tehrim S, Zhang F, Tong C. Huang J, Gong C, Zhou Y, Hua W, Liu S. 2014. Genome-wide comparative analysis of NBS-encoding genes between *Brassica* species and *Arabidopsis thaliana*. *BMC Genomics* 15(3): 1-18.
- Zhu C, Bai C, Sanahuja G, Yuan D, Farre G, Naqvi S, Shi L, Capell T, Christou P. 2010. The regulation of carotenoid pigmentation in flowers. *Archives of Biochemistry and Biophysics* 504: 132-141.

## APPENDIX

### PRIMER SEQUENCES

*Bo.CCD4.a* & *Br.CCD4.a* Sall dCAPS markers:

Forward Primer: AAC TTC GCT CCT GTC GTC GA

Reverse Primer: TCG TGT TCG CTA AAC CAA TG

(*Bo.CCD4.a* undigested band size: 379bp; *Br.CCD4.a* digested band size: 363bp)

*Bo.CCD4.b* & *Br.CCD4.b* markers:

Forward Primer: CCT CAC TCC TCA CAG GCT TC

Reverse Primer: GTG GAT CGA TGA ACG TGT TG

(*Bo.CCD4.b* amplified product: 374bp; *Br.CCD4.b* amplified product: 354bp).