Analysis of multiple *Brassica* transcriptomes reveals subgenome dominance in the response of *Brassica napus* to *Sclerotinia sclerotiorum*

by

Grant de Jong

BSc (Honours), University of Toronto, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Genome Science and Technology)

The University of British Columbia

(Vancouver)

January 2020

© Grant de Jong, 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Analysis of multiple *Brassica* transcriptomes reveals subgenome dominance in the response of *Brassica* napus to *Sclerotinia* sclerotiorum

submitted by **Grant de Jong** in partial fulfillment of the requirements for the degree of **Master of Science** in **Genome Science and Technology**.

Examining Committee:

Dr. Keith Adams, Botany Supervisor

Dr. Loren Rieseberg, Botany Supervisory Committee Member

Dr. Xin Li, Botany Supervisory Committee Member

Dr. Leonard Foster, Biochemistry and Molecular Biology Additional Examiner

Abstract

Whole-genome duplication (WGD; polyploidy) events have played an extensive role in the evolution of flowering plants. The sudden doubling of genetic material can expedite rapid novel changes to polyploid transcriptomes. For example, polyploids formed via an interspecific hybridization of closely related species, known as allopolyploids, can exhibit inconsistent expression patterns between their parental genome. These incipient disparities in parental gene dosage can have profound effects on the transcriptome of newly formed polyploids, which in turn can influence their response to environmental stressors. In particular, the extent to which the transcriptomic shock of polyploidization modulates the biotic stress response of plant species remains a nascent topic in polyploidy research. To elicit such a response. I subjected both natural and newly formed lines of *Brassica napus* to pathogen infection with the fungal necrotroph Sclerotinia sclerotiorum. To understand the origin of subgenome divergence in the newly formed polyploid, I also performed infections on the diploid parents of the resynthesized *B. napus*, *Brassica rapa* and Brassica oleracea. RNA-seq analyses of these pathosystems revealed wide-spread divergence between polyploid subgenomes in terms of both constitutive gene expression and alternative splicing patterns. This manifested in a global expression bias towards the *B. oleracea*-derived (C) subgenome among both polyploid hosts, enhanced by widespread non-parental down-regulation of the B. rapa-derived (A) homeolog. In the resynthesized *B. napus* specifically, this resulted a disproportionate C subgenome contribution to plant innate immunity and pathogen defense response, characterized by biases in both transcript expression level and the proportion of induced genes.

Lay Summary

Allopolyploids are organisms which form due to the hybridization between two related parental species, resulting in a whole-genome duplication. This phenomenon is particularly common among agronomically relevant plant species, partly due to the adaptive potential conferred through an influx of genetic material, and correspondingly, enhanced genetic diversity. However, these species must reconcile divergent parental transcriptomes, which can have profound consequences on global gene expression patterns — especially when confronted with environmental stressors. I used RNA-sequencing to investigate not only the response of *Brassica napus* to pathogen stress — elicited by the fungal necrotroph *Sclerotinia sclerotiorum* — but also how this response manifested in its parental species, *Brassica rapa* and *Brassica oleracea*, and a resynthesized allopolyploid *B. napus* formed directly from these species. This allowed me to identify a distinct bias to towards the B. oleracea-derived genome in both polyploids, which consequentially influenced the defence response of *B. napus*.

Preface

This research was performed with the assistance of my supervisor, Dr. Keith Adams, who proposed assessment of polyploid *Brassica napus* gene expression under biotic stress. Experimental procedures, including the bioinformatic pipeline, were largely designed and performed myself, apart from the following exceptions:

(1) The *Sclerotinia sclerotiorum* germplasm was provided by Rodney Werezuk from InnoTech Alberta and plant inoculation procedure was performed with the advice of the lab of Dr. Xin Li.

(2) I extracted the RNA from each sample; however, the bioanalysis and RNA sequencing was performed by sequencing services offered by Gnome Qubec.

(3) Scripts used to detect alternative splicing were modified from scripts designed and written by Dr. David Tack.

Table of Contents

Abstrac	:t	iii		
Lay Su	mmary	iv		
Preface		v		
Table of	f Contents	vi		
List of 7	Fables	ix		
List of l	Figures	x		
List of S	Supplemental Material	xii		
List of A	List of Acronyms			
Glossar	y	xiv		
Acknow	vledgments	xvi		
1 Intr	oduction	1		
1.1	Overview: Polyploidy in flowering plants	1		
1.2	Allopolyploid transcriptomes: Gene expression	2		
1.3	Allopolyploid transcriptomes: Alternative splicing	4		
1.4	Brassica as a model polyploid lineage	5		
1.5	Pathogen stress in allopolyploids: the Brassica-Sclerotinia pathosys-			
	tem	6		

	1.6	Resear	ch Objectives	7
2	Ana	lysis of 1	multiple Brassica transcriptomes reveals subgenome dom-	
	inan	ice in th	e response of Brassica napus to Sclerotinia sclerotiorum.	8
	2.1	Metho	ds	8
		2.1.1	Biological materials and preparation	8
		2.1.2	Inoculation and tissue extraction	9
		2.1.3	RNA extraction, library construction, and mRNA-sequencing	9
		2.1.4	Data preprocessing and read mapping	10
		2.1.5	Differential expression and differential alternative splicing	
			analyses	10
		2.1.6	Gene ontology enrichment analyses	11
		2.1.7	Analysis of synteny and the identification of orthologs and	
			homeologs	12
		2.1.8	Identification of defense-associated orthologs in Brassica	
			species	12
		2.1.9	Analyses related to subgenome expression bias and expres-	
			sion level dominance	13
	2.2	Results	S	13
		2.2.1	Brassica transcriptome sequencing and read alignment	13
		2.2.2	RNA-seq analysis suggests considerable reprogramming	
			of Brassica transcriptomes at 24 hours post S. sclerotiorum	
			inoculation	15
		2.2.3	Assessment of differential alternative splicing profiles dur-	
			ing early S. sclerotiorum infection	18
		2.2.4	Detection of subgenome-specific homeologs in B. napus	
			and orthologous genes in B. rapa and B. oleracea through	
			an analysis of synteny	20
		2.2.5	Quantitative analysis of Brassica host transcriptome simi-	
			larity using homeologous expression patterns	21
		2.2.6	Analysis of subgenome expression bias demonstrates a slight	
			C subgenome homeolog expression bias in both natural	
			and synthetic <i>Brassica napus</i>	22

	2.2.7	Analysis of subgenome-specific alternative splicing pat-	
		terns in both natural and synthetic Brassica napus	23
	2.2.8	Influence of pathogen infection on homeolog expression	
		bias in <i>B. napus</i> hosts	24
	2.2.9	Expression level dominance in the resynthesized B. napus	26
	2.2.10	splicing level dominance in the resynthesized <i>B. napus</i>	28
	2.2.11	Homeolog-specific expression categorization in the resyn-	
		thesized <i>B. napus</i>	29
	2.2.12	Impact of asymmetric subgenome expression on the pathogen	
		stress response of polyploid hosts	33
2.3	Discus	sion	35
	2.3.1	Transcriptome response of Brassica napus and its progen-	
		itor species to the pathogen Sclerotinia sclerotiorum	35
	2.3.2	Global patterns of subgenome expression bias in both nat-	
		ural and resynthesized Brassica napus	38
	2.3.3	Expression bias in the resynthesized Brassica napus is un-	
		derpinned by ancestral expression and a high degree of	
		transgressive down regulation	40
	2.3.4	Global bias in alternative splicing ratios in both natural and	
		resynthesized Brassica napus	41
	2.3.5	Non-parental AS event expression in the resynthesized Bras-	
		sica napus	43
	2.3.6	Biotic stress modulates patterns of subgenome expression	
		and splicing biases	44
	2.3.7	The influence of subgenome expression biases on the resyn-	
		thesized Brassica napus - Sclerotinia sclerotiorum pathosys-	
		tem	46
2.4	Conclu	usions	47
2.5	Tables	and Figures	49
Dir R.			
DIDIIOgi	ыюпоgrapny		

List of Tables

Table 2.1	Read mapping summary of Brassica species and S. sclerotiorum	61
Table 2.2	Summary of sample FPKM normalized expression data	62
Table 2.3	Overview of AS event distributions among Brassica species un-	
	der mock and pathogen inoculations	63
Table 2.4	Overview of AS event expression among homologous junctions	63

List of Figures

Figure 2.1	Hierarchical clustering results of RNA sequencing read counts	
	for each <i>Brassica</i> samples	49
Figure 2.2	Differential gene expression analyses of each <i>Brassica</i> host	50
Figure 2.3	Congruence in the gene ontology (GO) enrichment analyses of	
	differentially expressed genes for each host	50
Figure 2.4	Mean response (log ₂ FC) of hormone biosynthesis pathways,	
	hormone signalling pathways, and gene families commonly	
	implicated in Brassica napus stress response for each Brassica	
	host species.	51
Figure 2.5	GO enrichment analyses of differentially alternatively spliced	
	(DAS) genes for each <i>Brassica</i> host compared to enriched GO	
	terms of differentially expressed genes.	52
Figure 2.6	Alternative splicing and gene expression exhibit opposing re-	
	actions to pathogen stress.	53
Figure 2.7	Hierarchical clustering of host pathogen stress response	54
Figure 2.8	Subgenome expression biases under uninfected and S. sclero-	
	<i>tiorum</i> -infected conditions	55
Figure 2.9	Subgenome alternative splicing biases before and after pathogen	
	stress.	56
Figure 2.10	Transcriptome-wide change in subgenome bias between unin-	
	fected and infected expression states.	57
Figure 2.11	Parental expression and alternative splicing levels compared to	
	a homeolog-consolidated polyploid expression state	58

Figure 2.12	Homeolog expression and alternative splicing patterns relative	
	to parental expression states	59
Figure 2.13	Influence of subgenome expression bias on the defense re-	
	sponse of the resynthesized <i>Brassica napus</i>	60

List of Supplemental Material

Table S1	Enriched GO terms among up-regulated and down-regulated differentially
	expressed gene across all hosts.
Table S2	GO enrichment results for DAS genes across all hosts.
Table S3	Genes previously found to be associated with a biotic stress response (refer-
	enced in figure 4).
Figure S1	Relationship between constitutively mapped reads and alternatively mapped
	reads demonstrates that AS event expression scales with gene expression.

List of Acronyms

AS	Alternative splicing
BP	Base pair
DAMP	Damage-associated molecular pattern
DAS	Differential alternative splicing
ELD	Expression-level dominance
FDR	False discovery rate
FPKM	Fragments per kilobase of transcript per million mapped reads
HE	Homeologous exchange
NMD	Nonsense-mediated decay
PAMP	Pathogen-associated molecular pattern
PSI	Percent spliced-in
PTI	Pattern-triggered immunity
RBH	Reciprocal best-hit
TE	Transposable element
TF	Transcription factor
WGD	Whole-genome duplication

Glossary

Allopolyploidy: A form of whole-genome duplication (polyploid) in which the two constituent genomes are derived from an interspecific hybridization event between closely related species.

Expression-level dominance: A form of non-additive expression in which the total transcriptional dosage of a homeolog pair is more similar in expression level to a single "dominant" parental ortholog.

Homeolog: A gene duplicate (paralog) that has arisen due to a whole-genome duplication event.

Homeolog expression bias: Unequal expression between a set of homeologous genes, typically assessed at the genic level.

Reciprocal dominance: A given homeolog is more similar in expression to the alternative parental ortholog rather than the ortholog from which it is derived.

Regulatory dominance: Both homeologs are expressed to a level equivalent to only one of their parental orthologs.

Splicing-level dominance: Splicing-level dominance occurs when alternative transcript proportions of a subgenome-conserved alternative splicing event in the polyploid is equivalent to one of the parental splicing ratios for that event. Similar to expression-level dominance, polyploid alternative splicing is summarized by the proportion of total alternatively mapped reads at a given junction for an alternative splicing event that is shared between subgenomes.

Subgenome expression dominance: Subgenome expression dominance, used synonymously with subgenome expression bias, describes a subgenome which demonstrates a consistently higher levels of expression or homeolog expression bias.

Transgressive expression: A non-parental polyploid expression pattern character-

ized by higher or lower expression than either parents. This could apply to both homeologs or a single homeolog. For example, transgressive up-regulation refers to polyploid gene expression higher than either parent.

Acknowledgments

I would like to thank my supervisor, Dr. Keith Adams, for his mentorship and guidance. He provided me with many unique educational opportunities during the course of my Master's research. I would also like to thank my committee members, Dr. Xin Li and Dr. Loren Rieseberg, for their guidance and advice.

I also appreciate the valuable advice from members of the Li and Rieseberg labs, without whom my time in the wet lab would have been met with much greater difficulty. Of the Rieseberg lab, I would like to extend a special thanks to Winnie Cheung. Her support with administrative matters, among other trials and tribulations, was abundantly helpful.

Finally, I would like to thank past and present members of the Adams Lab, for their indispensable advice, support, and assistance — namely, Yichun Qiu, Ryan Bailey, Tonya Severson, and John Lee.

Chapter 1

Introduction

1.1 Overview: Polyploidy in flowering plants

Polyploidy has extensively shaped eukaryote evolution. The sudden multiplication of whole sets of homologous chromosomes is ancestral to many plant and animal lineages [1-5], suffusing these taxa with the genomic fodder for subsequent gene family diversification. However, as a recent evolutionary phenomenon, polyploidy is of demonstrably greater importance to plant evolution [4, 6, 7]. Angiosperms, or flowering plants, are particularly influenced by this process. Not only are they the most taxonomically numerous land plants - with over 300,000 extant species - all flowering plants share at least one episode of polyploidy in their evolutionary history [8]. Furthermore, these ancient whole-genome duplications have been thought to underpin the diversification and subsequent innovations present among ancestral angiosperm lineages [7, 9]. Importantly, many of these polyploids originated from interspecific hybridization between two closely related species (i.e. allopolyploidy) rather than as a result of intraspecific duplication of a single chromosome set (i.e. autopolyploidy). Allopolyploids are uniquely poised for a selective advantage, given that they contain two sets of diverged genomes, and therefore, an immediately increased genetic diversity. This has been shown to increase adaptive potential for wild angiosperm species, such as Compositae [10], and it was inferred that a higher percentage of domesticated crop plants are polyploid than wild species [11]. Moreover, this interaction between two divergent genomes can expedite large scale changes in gene expression [12, 13], alternative splicing [14], and epigenetic modifications (reviewed in [15, 16]) — the culmination of which can influence how polyploids respond to their environmental stressors (reviewed in [17]).

1.2 Allopolyploid transcriptomes: Gene expression

The interspecific hybridization event that underpins allopolyploidy introduces diverged sets of chromosomes and regulatory modules. The resulting transcriptome shock can have profound effects on the global expression profile of the newly formed polyploid. Despite the evolutionary distance between the merged parental genomes that constitute allopolyploid genomes — which are referred to as subgenomes — genic regions are largely syntenically conserved. This results in the expression of whole-genome gene duplicates, or homeologs, which in turn have their own suite of possibly divergent regulatory programs. The consequence of this is a vast array of potentially maladaptive transcriptional dosage differences. Studies have shown that polyploids compensate through widespread genomic and epigenomic modifications [18]. Accordingly, asymmetries in subgenome gene expression profiles begin to emerge in which one homeolog develops a complete loss in expression or homeolog pairs develop considerable disparities in transcript dosage, also known as homeolog expression bias [19].

Understanding the nature and extent of homeolog expression biases has been of particular importance to the field of polyploid evolution. Early analyses of homeolog expression bias have shown widespread transcriptional differences between homeologous gene pairs [20–22] and the impact these expression disparities might have on organ development and stress has been well-documented in allopolyploid cotton [23, 24]. Non-parental gene expression patterns have also long been shown to manifest in recent polyploids [12]. For example, comparisons between allopolyploid cotton and its parental species using microarray data highlighted extensive non-parental homeolog expression patterns [25], culminating in subgenome expression dominance. Given the ultimate fate of polyploids is thought to be the reversion to diploid status through biased fractionation of their constituent subgenomes, it is unsurprising that homeolog expression dominance can manifest non-randomly as subgenome expression dominance.

The application of RNA-seq technologies has duly resulted in a better understanding of transcriptome-wide patterns of subgenome expression dominance. Comparative transcriptome analyses in a number of polyploid species have shown abundant evidence of homeolog expression bias manifesting at a subgenome level [26–30]. For example, in allopolyploid *Mimulus peregrinus* and tetraploid cotton, synthetic polyploid lines were used which highlighted the fact that subgenome expression dominance likely occurs in the first polyploid generation. Moreover, this bias was shown to increase in magnitude over time [26, 28]. Similar to Rapp et al. [25], RNA-seq based studies have also compared parental expression with recently formed polyploid subgenome expression [26, 30, 31]. These studies pooled homeolog transcripts into a single polyploid expression level which was then categorized based its expression level relative to the parental transcriptomes. Ultimately, the distribution of these categories has provided evidence for parental dominance in early polyploid formation that might not be otherwise apparent [26, 30].

The implications of homeolog expression biases have been seen in the way in which polyploids respond to their environment. Studies have shown growth, as well as abiotic and biotic stress responses, are influenced by disparities in homeolog transcript levels [23, 24, 32]. Earlier work in natural and synthetic *Gossypium hirsutum* demonstrated widespread changes in relative expression of homeologs under various abiotic stress conditions [24]. More recently, transcriptome sequencing in hexaploid wheat infected with a fungal pathogen described consistent bias towards the B and D subgenomes. They also identified biases in the responsiveness of homeolog induction bias [32]. While this provides an intriguing way to analyze stress-induced homeolog expression differences, it remains to be seen whether the time since polyploidization plays a role in the compartmentalization of defense response. Furthermore, it is not currently known if different layers of homeolog expression — such as homeolog alternative splicing patterns — are additionally modulated in response to stressors.

1.3 Allopolyploid transcriptomes: Alternative splicing

Alternative splicing (AS) increases transcriptional diversity through the regulated processing of primary mRNA transcripts at alternative splice sites. Mechanistically, it is one of the many facets of transcriptional regulation and is ubiquitous among eukaryotes. As reviewed in [33], approximately 70% of intron-containing plant genes are capable of AS. Furthermore, AS events derived from plant genes are categorically distinct from other eukaryotes such as animals in that intron-retention (IR) is the most common form of AS [34]. Retention of introns in protein coding transcripts typically results in the presence of pre-mature stop codons (PTC) which can trigger nonsense-mediated decay (NMD) [35, 36], remain sequestered in the nucleus for later processing [37], or possibly result in a non-functional protein isoform. While seemingly maladaptive, the transcriptional inclusion of introns is thought to have regulatory function in plant cells, as AS-mediated NMD has been shown to modulate the abundance of constitutively spliced transcripts in Arabidopsis thaliana [35, 36, 38]. AS-coupled NMD (AS-NMD) can also influence transcription under environmental signals such as abiotic stress [37, 39, 40] and biotic stress [41, 42]. More generally speaking, studies have shown the AS-landscape of plants seems to react context-dependently and can modulate the abundance of key transcripts under altered expression states [43, 44]. Recent examples of this can be seen in the transcriptome sequencing of *Brachypodium distachyon* under viral infection [45], or the abiotic stress response of Arabidopsis thaliana seedlings [46].

Considering the extent to which allopolyploidy can influence the constitutive expression of transcripts, the influence whole-genome duplication has on AS is also of interest. Widespread changes in expression level, activity, or methylation of genes encoding splice factors could have far-reaching consequences on subgenome AS profiles [47]. A previous study by Zhou et al. addressed this through the analysis of AS patterns from two resynthesized lines of allotetraploid *Brassica napus* and their progenitor species. They demonstrated widespread non-parental AS changes in the polyploid (26-30% of duplicated genes), a large number of which involved a loss of AS event expression in the polyploids [14]. Additionally, recent analyses of allopolyploid AS have shown considerable subgenome AS divergence [48, 49]. Zhang et al., in particular, found non-parental AS patterns between duplicated

chromosomes in segmental allotetraploid rice, particularly among subgenome regions homogenized by homeologous exchanges (HEs) [50]. Moreover, a study in hexaploid wheat demonstrated differential AS responses to abiotic stress [51]. This indicates that AS patterns are not only influenced by polyploidization, but AS of homeologous genes is further modulated by stress. However, a paucity of research has been published on how stress responses might influence pre-existing homeolog splicing patterns, or whether these divergent homeolog AS patterns are parental or non-parental in origin.

1.4 *Brassica* as a model polyploid lineage

The plant genus Brassica, a member of the mustard family Brassicaceae, is comprised of several economically important crop species. The relationship between six of these species, in particular, was famously explained in U's triangle [52] as a model for polyploid evolution. This diagrammatic representation of wholegenome duplication contains three diploid species — B. rapa, B. oleracea, and B. nigra — all of which have hybridized to form three tetraploid species — B. napus, B. juncea, and B. carinata. Of these polyploid species, B. napus (A_nA_nC_nC_n, 2n=4x=38) has garnered considerable attention in the study of polyploidy, having arisen only 7,500 years ago from a hybridization event between the ancestors of B. rapa (A_rA_r , 2n=2x=20) and B. oleracea (C_oC_o , 2n=2x=18) [53]. Due to its recent formation, it has been a useful model in studying genomic and transcriptomic consequences of allopolyploidization [14, 18, 30, 54–56]. One particular advantage is that allopolyploid lines can be resynthesized from modern *B. oleracea* and *B.* rapa cultivars, allowing researchers to use synthetic *B. napus* as a model for recent polyploidization. These resynthesized lines have been particularly useful for comparative analyses between newly formed polyploids and their parental genotypes to identify differences in homeolog expression bias [30], AS divergence [14], and small RNA populations [57].

1.5 Pathogen stress in allopolyploids: the *Brassica-Sclerotinia* pathosystem

Mathematical models have suggested that the hybrid origins of recent allopolyploids, and the subsequent increase in genetic diversity, confers enhanced resistance to pathogens [58]. This has been further supported by work in soybean showing increased resistance gene (R-gene) diversification following polyploidy [59]. However, few studies have directly investigated how allopolyploids and their constituent subgenomes respond to pathogen stress. Research in Arabidopsis tetraploids showed that differential interactions of homeologous WRKY transcription factors lead to alterations in PR1 induction [60]. Additionally, an analysis of homeolog expression patterns between hexaploid wheat subgenomes [32] revealed subgenome bias in the induction of defense response genes. In the context of Brassicas, recent efforts to construct B. napus and B. oleracea pangenomes have shown that presence/absence variants between different cultivars are enriched for disease resistance genes [56, 61], which suggests that R-genes are highly influenced by homeologous exchanges in established Brassica lines. However, comprehensive analyses of subgenome-specific response to biotic stress have not been performed in Brassicas.

Given the global importance of *Brassica* crops, a wealth of research has amassed around the effect of economically disruptive pathogens. In *B. napus*, one of the most devastating infections arises from the generalist fungal necrotroph *Sclerotinia sclerotiorum*. Interestingly, *Brassica* species lack high levels of resistance to *S. sclerotiorum*; consequently, no evidence of effector-triggered immunity (ETI) has been identified in commercially available *Brassica* species [62], indicating that each species must rely on the innate plant immune response for defense against the pathogen. This typically involves the recognition of pathogen-associated molecular patterns (PAMPs), via pattern-recognition receptors, which invoke patterntriggered immunity (PTI). The immune response is then triggered by signalling cascades facilitated by mitogen-activated protein kinases (MAPKs) and MAPKindependent pathways such as the calcium (Ca⁺²) pathway [63]. Among necrotrophic pathogens, the plant hormones jasmonic acid (JA) and ethylene (ET) also play key roles in controlling the defense response. This leads to the expression of transcription factors (TFs), such as WRKY33, which in turn activate specific defense responses, including pathogensis-related (PR) proteins, reactive oxygen species (ROS) accumulation, detoxification, plant-wall restructuring enzymes, and secondary metabolites — reviewed in [64].

The lack of qualitative resistance to *S. sclerotiorum* in *B. napus* is also interesting in the context of allopolyploid evolution, as infection might highlight nonparental or subgenome-biased manifestation of the innate immune response. For example, earlier work using *B. napus*, and its diploid progenitors *B. rapa* and *B. oleracea*, showed evidence of an A subgenome bias in peroxidase gene expression [65]. Additionally, quantitative trait loci (QTL) are distributed throughout both subgenomes [66, 67]; however, HEs [68, 69] and subgenome expression dominance might result in a biased defense response. Transcriptome analyses of the subgenome contribution to *S. sclerotiorum* defense in *B. napus* has not yet been studied but may provide useful information.

1.6 Research Objectives

While previous studies have assessed homeolog expression bias in polyploids [28, 32, 70, 71] and patterns of expression level dominance between polyploids and their progenitor species [25, 26, 30], few have assessed how these phenomena are modulated by pathogen stressors [32]. Furthermore, a dearth of data exist on the extent to which homeologous alternative splicing may exhibit biases [48]. To evaluate how different layers of allopolyploid transcriptomes — i.e. gene expression and AS — respond to pathogen stress, I surveyed the transcriptomes of *B. napus*, its diploid progenitors *B. rapa* and *B. oleracea*, and a resynthesized tetraploid *B. napus* formed directly from the genotypes of the parental diploid lines. Through this, I investigated the suite of subgenome expression changes that ensued from merger of divergent polyploid subgenomes and the way by which these disparities in expression influence responses to biotic stress in polyploids.

Chapter 2

Analysis of multiple *Brassica* transcriptomes reveals subgenome dominance in the response of *Brassica napus* to *Sclerotinia sclerotiorum*

2.1 Methods

2.1.1 Biological materials and preparation

Seeds of *Brassica napus* (cv. Sentry), *Brassica rapa* (IMB218), *Brassica oleracea* (TO1000), and a resynthesized *Brassica napus* formed from spontaneously doubled IMB218 and TO1000 lines (EL9450sp; polyploid Line 47) were held in a 0.5% solution of Plant Preservative Mixture for 15 minutes and rinsed for two hours. Seeds were then hydroprimed in Miracloth soaked in 20mL of dH₂O for 24 hours without light. Germinated plants were grown at 20°C and 60% humidity under a 16/8h photoperiod (200 E/m²/s) in identical growth chambers for three weeks pre-inoculation. Dehydrated ascospore disks of a single *S. sclerotiorum* isolate were provided by InnoTech Alberta and were stored at 4°C in a desiccator and without light.

2.1.2 Inoculation and tissue extraction

S. sclerotiorum inoculum was prepared by pipetting 50mL of a 5×10^5 ascospores per mL dH₂O solution onto potato dextrose agar at 20°C, which was stored under low light. Following 48 hours of fungal growth, 7mm diameter mycelial agar plugs were taken from the growing margin of pathogen and brushed with 2% agar solution. Mycelial plugs were then placed on healthy *Brassica* leaves and were fixed in place by sterilized needles to simulate wounding. Plants were immediately placed in plastic bags and stored in a growth chamber at 18°C and 60% humidity to foster *S. sclerotiorum* infection. Mock inoculations were treated with the same protocol using only sterile PDA. Both mock and pathogen inoculations occurred at approximately 12PM. At 24 hours post-inoculation (hpi), sterilized 6mm hole punches were used to excise two discs of leaf tissue roughly 2mm adjacent to developed lesions for each plant. Discs were immediately frozen in liquid nitrogen and stored at 70°C.

2.1.3 RNA extraction, library construction, and mRNA-sequencing

Leaf discs from six plants were pooled for each biological replicate (totalling 12 discs) and RNA was extracted using the Invitrogen RNAqueousTM Total RNA Isolation Kit with the supplementary Plant RNA Isolation Aid in accordance with the protocol provided by the manufacturer. RNA quality assessment was a composite of Nanodrop quantitation, Qubit Fluorometer (Invitrogen) RNA quantitation, and gel electrophoresis to ensure samples were of high RNA purity and integrity. Samples were then sent to Gnome Qubec Innovation Centre for Bioanalyzer (Agilent) analysis, library preparation, and mRNA-sequencing. Libraries were constructed using the TruSeq Stranded mRNA Library Prep Kit and sequenced on the Illumina NovaSeq 6000 platform, resulting in 2x100bp paired-end reads.

2.1.4 Data preprocessing and read mapping

Read quality was assessed using FastQC [72] and poor quality reads were removed with Trimmomatic v0.32 [73]. Only reads satisfying the following conditions were retained: (i) reads larger than 50 bp following a six base head crop, and (ii) reads possessing an average PHRED score of 20. High-quality reads were aligned to the reference genome of each species using STAR v2.7 [74] (see supplemental material for full set of commands). Given the uncertain lineage of the Sentry cultivar of *Brassica napus*, and therefore to accommodate possible presence/absence variants, reads were aligned to a composite of the pangenome [56] and the *Brassica napus* v5 reference genome [53]. Diploid sample reads were mapped to the *Brassica* rapa v3 [75] and the Brassica oleracea TO1000 reference genome [76], respectively. The resynthesized *Brassica napus* samples were mapped to a hybrid of the aforementioned Brassica rapa and Brassica oleracea genomes as few generations have elapsed since the original interspecific hybridisation event. All samples were also mapped to the Sclerotinia sclerotiorum genome [77]. To compensate for possible ambiguity in homeolog read assignment, STAR was performed such that only uniquely mapped reads were used for downstream analyses.

2.1.5 Differential expression and differential alternative splicing analyses

Constitutive gene expression was quantified using FeatureCounts from the Subread package such that only exon-spanning fragments were counted [78]. Raw counts were normalized using the weighted TMM library-specific normalization of edgeR 3.20.9 [79]. For the purposes of identifying prominent global transcriptome changes, genes were considered significantly differentially expressed with a false discovery rate (FDR) < 0.001 and an absolute fold change (FC) > 3.

Alternative splicing was detected using custom R and Python scripts [80]. Splicing was quantified based on a metric referred to as percent spliced-in index (PSI). This was represented by the proportion of alternatively mapped reads that support a specific splicing event (alt_{ij}) compared to the sum of constitutively (*cons*)

and alternatively mapped reads (alt) at a given junction (j) such that:

$$\psi_{ij}(PSI) = \frac{alt_{ij}}{\sum (alt + cons)_j}$$

Only splicing events that were (i) supported by at least one read per biological replicate and (ii) totaled at least eight reads were retained for further analysis [53]. Differentially alternatively spliced (DAS) events were calculated by performing a logistic regression on event-specific PSI values with the following model (as originally described by Tack et al.):

$$pr(alt) \sim bi(p = logit^{-1}(b_0 + b_{condition}))$$

If the logistic regression model provided support for a quantitative difference in condition (*FDR* <0.01) and the absolute $\log_2(\text{fold change}) > 1$, alternative splicing events were considered significantly differentially spliced.

2.1.6 Gene ontology enrichment analyses

To facilitate ontological comparison between all *Brassica* species, the genomes of *Brassica napus*, *Brassica rapa*, and *Brassica oleracea* were analyzed using the same functional annotation pipeline to produce comparable lists of gene ontology (GO) terms and putative gene descriptions for each genome. The first step in assembling these annotations involved searching the full list of confirmed and predicted protein sequences of each plant species against the InterPro database using InterProScan5 [81]. This was followed by the use of GO FEAT [82], which integrates several databases (such as InterPro, UniProt, KEGG, NCBI Protein, and Pfam). Pannzer2 [83] was then used to provide an estimate of GO terms for the remaining unannotated genes. Annotations were then merged and redundancies were removed. Finally, GO enrichment analysis was performed using the cluster-Profiler R package [84], with GO terms considered significantly enriched with a *FDR* <0.05.

2.1.7 Analysis of synteny and the identification of orthologs and homeologs

The quantitative comparison between polyploid subgenomes and their respective progenitor genomes requires a list of orthologous gene pairs. While such a list exists for earlier versions of the *B. rapa* genome [53], the use of a contemporary *B. rapa* genome assembly necessitated an updated list of genes. The synteny analysis pipeline described in Chalhoub et al. was employed to assemble a matrix of homologous genes by first identifying regions of conserved positional orthology, in the form of syntenic blocks, followed by a Reciprocal Best Hit (RBH) BLAST search to detect orthologs or homeologs that are not retained in syntenic blocks [53]. The use of RBH BLAST, MCScan and JCVI libraries [85] resulted in a list 42,712 possible gene pairs among the four genomes (i.e. A_{NAPUS}, C_{NAPUS}, A_{RAPA}, C_{OLERACEA}) and a total of 20,062 that shared among all aforementioned genomes.

The list of orthologs was also necessary for intergenomic comparisons of AS; however, junction-specific AS quantification requires specific processing of orthologous gene data. As previously mentioned, AS is detected based on a ratio of intronic reads and exonic reads at a given exon-exon junction; therefore, such AS comparisons must occur within junctions that have homologous exon-exon structure. To identify these homologous exon-exon junctions, exon sequences from the representative genes models of each species were extracted from their requisite genome assemblies. Pairwise alignments of orthologous exons were performed for each genome (A_N : C_N : A_R : C_O) using RBH BLAST (dc-megablast). A total of 94,180 consecutive exons were logged as possible homologous splice junctions across all four species and 76,770 putative splice junctions were shared among all species.

2.1.8 Identification of defense-associated orthologs in *Brassica* species

To query the response of specific biological pathways and gene families to early pathogen infection in each *Brassica* species, lists of *Arabidopsis thaliana* orthologs implicated in the defense against necrotrophic pathogens were assembled. These genes were then BLAST searched against *B. napus*, *B. rapa*, and *B. oleracea* gene

sequences [63, 86]. This list was comprised of three primary sets of genes: (i) genes related to plant hormone biosynthesis pathways, (ii) gene families of receptor proteins, and (iii) genes involved in either the propagation or accumulation of reactive oxygen species.

2.1.9 Analyses related to subgenome expression bias and expression level dominance

To study subgenome-specific differences in homeolog expression patterns, both subgenome expression bias (relative homeolog expression) and expression level dominance (the sum of homeolog transcripts compared to parental expression levels) [25] were investigated. Subgenome expression bias was analyzed for both natural and the resynthesized B. napus using edgeR. Subgenome differential expression incorporated sample-wide library sizes and gene length offsets to avoid underestimation of subgenome-specific mapping rate. Homeologs were defined as differentially expressed with a FDR < 0.01. Analyses of expression level dominance (ELD) involved a prerequisite analysis of differential gene expression between diploid transcriptomes and the concatenated resynthesized *B. napus* subgenome for each condition separately. Differential gene expression was performed using edgeR. Homeologous genes were considered differentially expressed if their absolute $\log_2 FC$ ratios were > 1 and they had a FDR < 0.01. Parental comparisons with the individual resynthesized *B. napus* subgenomes were similar to ELD; however, comparisons between homeologs incorporated the same library size considerations used for calculations of subgenome expression bias. Analyses of differential alternative splicing bias between polyploid subgenomes and quantitative comparisons between parental-polyploid event homologs were consistent with the aforementioned differential alternative splicing analysis.

2.2 **Results**

2.2.1 Brassica transcriptome sequencing and read alignment

RNA was extracted from a total of 24 samples: three biological replicates for each of the four *Brassica* genotypes at 24 hours following mock or *S. sclerotiorum* in-

oculation. This resulted in a sum of 2439.3 million clean reads — an average of 130.5 million reads for the polyploid samples and 72.8 million reads for the diploid samples (Table 2.1). The clean reads from each sample were mapped to the appropriate reference genome, of which 89.0% and 70.7% of reads in mock-inoculated and pathogen-inoculated samples, respectively, were uniquely mapped. Among the pathogen-inoculated samples only, an average of 16.4 million reads per sample also mapped to the *S. sclerotiorum* reference genome.

Transcript expression was only considered if all replicates in a given condition had a FPKM value ≥ 1 . FPKM normalized expression among replicates was significantly correlated with a Pearson correlation coefficient (R) ranging from 0.879 to 0.997 and a median of 0.975 ($P < 1 \ge 10^{-16}$). Hierarchical clustering on sample FPKM values resulted in condition-specific groups, suggesting there were large differences between mock and inoculated samples among each species (Figure 2.1A). Furthermore, these condition-specific groups were retained following hierarchical clustering based on interspecific ortholog FPKM values (Figure 2.1B). The proportion of genes expressed in the mock inoculated samples was consistent across all four species, with a mean of 44.4%, while the pathogen-inoculated samples expressed a lower proportion of genes on average (41.4%). The highest proportion of expressed genes was observed in *B. rapa* (47.2% to 49.3%) (Table 2.2).

To illustrate the distribution of gene expression for each sample, genes were categorized based on expression level: $1 < \text{FPKM} \le 5$ (low), $5 < \text{FPKM} \le 50$ (moderate), and 50 < FPKM (high) (Table 2.2). Overall, the distribution of these expression categories was highly consistent between mock and inoculated samples for all hosts. Generally speaking, all four hosts also boasted similar categorical proportions. For instance, across all hosts the majority of genes were moderately expressed, with a mean of 49.3% among the polyploid hosts and 55.5% among the diploid hosts. However, there was a notable difference in the incidence of lowly expressed genes between the polyploid (43.2%) and diploid hosts (28.8%).

2.2.2 RNA-seq analysis suggests considerable reprogramming of *Brassica* transcriptomes at 24 hours post *S. sclerotiorum* inoculation

The global transcriptome response of each *Brassica* host was first investigated through an analysis of differential gene expression, providing a general insight into of how the pathogen stress response of the resynthesized *B. napus* compares to that of its progenitor genotypes — the *B. rapa* and *B. oleracea* hosts — and the more distantly related natural *B. napus*.

Overall, *B. oleracea* and the polyploid species exhibited comparable expression changes under pathogen stress. Using a modest threshold of *FDR* <0.01 and an absolute fold change >1 revealed 11,100 DEGs in *B. oleracea*, 22,732 DEGs in the natural *B. napus*, and 20,441 DEGs in the resynthesized *B. napus* (41.3%, 43.1% and 44.9% of genes designated as expressed in each genotype, respectively). The global transcriptome response was notably less severe in *B. rapa*, with 7275 DEGs or 30.5% of the transcriptome differing between mock and pathogen treatments. To identify genes which are characteristic of a response to nectrotrophic pathogens, a more stringent DEG threshold was used for each species (*FDR* <0.001 and an absolute FC >3) (Figure 2.2). This revealed 4185 genes (8.4%) in natural *B. napus*, 5592 genes (12.3%) in the resynthesized *B. napus*, 2176 genes (8.1%) in *B. oleracea*, and 1395 genes (5.8%) in *B. rapa* that were highly responsive to the pathogen treatment. While *B. rapa* has the lowest proportion of highly differentially expressed genes, it also exhibited the lowest quantity of *S. sclerotiorum* RNA (Table 2.1) and the highest proportion of expressed genes.

For functional insight into this conservative assembly of genes involved in the *Brassica-Sclerotinia* pathosystem, enriched biological processes were assessed. This was first accomplished by the compilation of unique functional annotations for each plant genome. A custom annotation pipeline identified putative gene ontologies for 84.1% of the *B. napus* pan-genome, 81.7% of the combined *B. oleracea* - *B. rapa* genome to which the resynethsized *B. napus* was mapped, 70.7% of the *B. oleracea* genome, and 85.9% of the *B. rapa* genome. Using these custom lists of GO terms, enrichment analyses were run using clusterProfiler on sets of up-regulated (FPKM > 3) and down-regulated (FPKM < -3) genes, respectively [84].

The polyploid plant hosts exhibited similarly large sets of significantly enriched GO terms, with 45 and 67 enriched terms for up-regulated genes in the natural and resynthesized *B. napus* (Supplementary Table 1). There was an overlap of 22 terms between both species, most of which were related to a PAMP/DAMP response, consistent with expectations for generalist necrotrophic fungi infections such as: response to wounding, innate immune response, defense response, defense response signaling pathway, chitin catabolic process, and cell wall macromolecule catabolic process. There were also several enriched plant hormone ontologies; most notably, jasmonic acid and salicylic acid. Considerably more GO term enrichment occurred among down-regulated genes than up-regulated genes; however, there was similar concertedness, with an overlap of 25 gene ontologies. In both polyploids, the most significantly enriched categories were photosynthesis, photosystem II assembly, and photosynthesis, light harvesting ($FDR < 1 \ge 10^{-25}$). Multiple terms associated with energy metabolism were sporadically enriched between the two polyploids without a clear pattern: carbon fixation, starch biosynthetic process, fructose metabolic process.

While diploid and polyploid hosts had similar enriched ontologies, the diploid species showed less GO enrichment. Up-regulated genes in *B. oleracea* and *B. rapa* were significantly enriched for 30 GO terms (Supplementary Table 1). Overlapping terms between the two progenitor species were congruent with the polyploids, for which the most enriched categories were response to wounding, defense response, and response to jasmonic acid. Enrichment analysis of down-regulated gene sets revealed considerably fewer significantly enriched terms than in the polyploids, with 10 in *B. oleracea* and 9 in *B. rapa*. There was no direct agreement among diploid GO terms. However, these terms were also related to energy metabolism and growth — i.e. a reduction in the expression of auxin signalling and phototropism in *B. oleracea* and photosynthesis in *B. rapa*.

Taken together, each plant species demonstrated similar enrichment of biological processes, characterized by an up-regulation of genes related to response to wounding, innate immunity, jasmonic acid signalling, chitin catabolism, oxidant detoxification, and cell wall macromolecule catabolism (Figure 2.3). *B. oleracea* appeared to be more similar to the polyploid hosts than to *B. rapa*. This was evidenced by enrichment of terms such as nodulation and salicylic acid metabolic process. Additionally, these three hosts had a very significant enrichment for genes related to cellular response to hypoxia. An analysis of putative gene descriptions suggests that up-regulated genes in this ontology are broadly associated with a biotic stress response — i.e. VQ-containing motif [87], patatin-like genes [88], flotillin-like genes, PAMP-induced secretory genes, and several oxidoreductase genes. The enzymatic activities of these oxidoreductase genes (such as putative tetrahydroberberine oxidases and reticuline oxidase-like genes) likely produce hydrogen peroxide (Supplementary Table 1), suggesting the polyploid hosts and *B. oleracea* exhibit more pronounced expression of enzymes which produce reactive oxygen species.

Manifestation of immune response at 24 hpi in each of the Brassica hosts was investigated by compiling lists of genes associated with resistance to necrotrophic pathogens for each species (Figure 2.4). All four hosts exhibited a response consistent with previous studies in *B. napus* [63, 89, 90]. The model response to S. sclerotiorum infection first involves PAMP/DAMP recognition via pattern recognition receptors (PRR) localized on the cell surface. Examples of this can be seen in the significant up-regulation of chitin receptor kinases, lectin receptor kinases, cysteine-rich receptor-like kinases, and PEP receptors (Supplemental 2). These were highly diverse and ubiquitously up-regulated across all hosts. Upon activation, these enzymes trigger signal transduction cascades which principally consist of the mitogen-activated kinases (MAPK). Of all MAPKKK-MAPKK-MAPK module genes analyzed, only MAPKKK18 and MAPKKK19 were highly up-regulated across all hosts (FDR < 0.01 and $\log_2 FC > 2$). These signal transductions likely then activate *Brassica* defense responses. At 24hpi, WRKY transcription factors are consistently up-regulated (Supplemental 3). WRKY6, most notably, was highly up-regulated across all hosts ($\log_2 FC > 4$); however, most WRKY transcription factors were broadly up-regulated (mean $\log_2 FC = 4.4$). Additionally, jasmonic acid (JA) and ethylene (ET) signalling and biosynthesis pathways were up-regulated. The former was highly enriched among up-regulated GO terms, and both of which exhibited an extensive increase in expression. Of the JA signalling and biosynthesis genes, JAZ10, OPR1, AOC3, and particularly, PR3/4 were strongly induced across all species. ACS1/2 were also universally induced $(\log_2 FC > 5)$, however other ethylene biosynthesis genes were also induced but less concertedly. Such transcriptional reprogramming also facilitates expression of specific defense genes, of which chitinases were strongly represented (P < 0.01).

2.2.3 Assessment of differential alternative splicing profiles during early *S. sclerotiorum* infection

Gene expression inferred from exonic count data offers a general glimpse into the transcriptomic activity of the *Brassica* hosts, but the process of gene expression is not limited to constitutively spliced transcripts. To provide a more comprehensive analysis of gene expression patterns under pathogen stress, alternative splicing patterns of the *Brassica* hosts were also examined. Considering the fact that the dataset consisted of short-read sequencing data, I used a splice junction-centric approach to splicing event detection. An average of 84,865 splicing events were detected in the polyploids and 44,260 splicing events in the diploids. An enumeration of the most common angiosperm AS event classes revealed comparable event distributions for each species (Table 2.3). These were similar to event ratios outlined in Chalhoub et al., as intron retention was the most frequently observed event, corresponding to approximately 66.7% of all AS events, followed by ALTA, ALTD, ALTP, and skipped exons[53]. The predominance of IR events is also consistent with other studies in both monocots and dicots [45, 91, 92].

To provide insight into large-scale AS changes under pathogen stress, I first identified differences in splicing event identity between mock and infected treatments. In *B. napus*, 27.7% of (22,471) AS events were unique to infected samples. The proportion was somewhat comparable with the diploid hosts, in which 25.2% (10,076) and 25.3% (11,186) of events in *B. rapa* and *B. oleracea*, respectively, were unique to infected expression states. Slightly more agreement was found between infected and uninfected resynthesized *B. napus* samples, which had an overlap of 21.7% (11,186 events unique to infected samples). Of the splicing events that were shared between conditions, I also identified events that showed significantly different ratios of alternative to constitutively mapped reads for each *Brassica* host. Overall, the incidence of DAS was highest in *B. napus*, in which I identified 12,709 events (13.3% of all *B. napus* AS events) which originated from 8225 genes (or 8.1% of annotated genes). There were notably fewer AS events identified in the resynthesized *Brassica napus*, with 8863 DAS events (11.0%) in

5241 genes (5.2%). The frequency of DAS in *B. oleracea* was similar to the polyploids, which had 5853 (12.2%) DAS events among 3642 genes (6%). Analogous to the gene expression analyses, *B. rapa* also displayed the lowest levels of DAS, with 4087 significantly different events (9.6%) among 2387 genes (5.2%).

Of the differentially spliced events, *B. napus* and *B. oleracea* had similar regulatory patterns, exemplified by a higher number of up-regulated events and fewer down-regulated events. This trend was more pronounced in *B. napus*, which had 5667 (5.9%) up-regulated and 2347 (2.5%) down-regulated AS events, while *B. oleracea* exhibited a less notable difference of 1971 (4.1%) up-regulated and 1501 (3.1%) down-regulated events. In contrast, the majority of DAS events in the resynthesized *B. napus* and *B. rapa* were down-regulated: 1468 (1.8%) compared to 4096 (5.1%) AS events in the polyploid and 920 (2.2%) compared to 1625 (3.8%) AS events in the diploid.

Similar to DEGs, GO enrichment analyses were performed on genes presenting DAS events. Unlike the enrichment analyses of DEGs, little agreement was found between Brassica hosts when specific GO terms were compared (Figure 1.5). While B. napus, the resynthesized B. napus, and B. oleracea exhibited some GO term inconsistency, up-regulated terms were functionally associated with photosynthesis, growth, and housekeeping genes. For example, photosynthesis, light harvesting ($FDR < 2.4 \times 10^{-7}$) in *Brassica napus* and thylakoid membrane organization (FDR < 6.8 x 10^{-3}) in the resynthesized Brassica napus are related to photosynthetic processes. B. oleracea and B. napus were also enriched for lateral root development ($FDR < 1.1 \times 10^{-2}$) and regulation of cell growth (FDR $< 3.8 \text{ x} 10^{-2}$), suggesting an up-regulation of AS in genes functionally correlated with plant growth (Supplementary Table 2). These trends did not extend to B. rapa, which lacked any significantly enriched GO terms among genes containing up-regulated AS events. Enrichment of the genes containing down-regulated DAS events displayed a lack of interspecific overlap that was comparable to up-regulated AS events. While all four *Brassica* hosts were enriched for the tricarboxylic acid cycle (2.1 x $10^{-13} < FDR < 7.2 \times 10^{-4}$), it was the only GO term shared among all species. However, the polyploids and B. oleracea showed enrichment for GO terms associated with biotic stress: response to wounding, regulation of defense response, and regulation of jasmonic acid mediated signalling pathway in the resynthesized

B. napus; methylglyoxal catabolic process to lactate, glyoxal removal [93], cellular oxidant detoxification, and MAPK cascade (FDR < 0.05) in *B. napus*; and response to wounding (FDR < 0.05) in *B. oleracea* (Supplementary Table 1).

The biological theme exhibited by the up-regulated and down-regulated differentially spliced genes was inversely related to GO enrichment of DEGs (Figure 1.5). This suggests a possible negative correlation between the gene expression and AS of a given gene in response to pathogen stress. To investigate the possibility of such a relationship, the change in expression of a given gene was compared to changes in the AS ratio for genes presenting DAS events. Comparisons across the transcriptome of each host using Pearson correlation revealed a moderate negative correlation between AS and GE, ranging from R = -0.43 ($P < 3.4 \times 10^{-309}$) in *B. na*pus to R = $-0.29 (P < 2.2 \times 10^{-43})$ in *B. rapa* (Figure 1.6A). Such evidence suggests that, across each host transcriptome, AS and GE tend scale negatively. To ensure this relationship was not due to a static background expression of AS reads, the log₂FC of constitutively mapped reads was compared against the log₂FC of intronmapped reads, resulting in a significant correlation across all species (Pearson correlation cofficient (R) ranging from 0.90 to 0.84; $P < 2.2 \text{ x } 10^{-16}$; Supplementary Figure 1). Considering how this relationship manifests at a single-gene resolution, genes implicated in hormone biosynthesis and signalling pathways were also individually analyzed. In *B napus*, the majority of significantly responsive genes i.e. DAS events and DEGs - are directionally opposed in their log-fold changes (65%; Figure 1.6B).

2.2.4 Detection of subgenome-specific homeologs in *B. napus* and orthologous genes in *B. rapa* and *B. oleracea* through an analysis of synteny

An interspecific comparative analysis of gene expression patterns requires the identification of homeologs between the polyploid subgenomes and their respective parental orthologs. An analysis of synteny between the *B. napus* subgenomes and the parental genomes of *B. oleracea* and *B. rapa* (i.e. A_N : C_N : $A_R:C_O$) resulted in a total of 42,712 gene pairs across all genomes and subgenomes, 24,573 pairs between *B. napus* subgenomes, and 28,196 between the *B. rapa* and *B. oleracea* genomes. Overall, 20,062 genes had an existing homeolog in both *B. napus*
subgenomes and an ortholog in both diploid parents.

Comparison of the AS patterns of *B. napus* subgenomes to *B. oleracea* and *B. rapa* not only requires orthologous and homeologous genes, but also the identification of homologous exon-exon junctions shared between these genes. An analysis of exon homology revealed a total of 94,180 possible exon junction pairs. Of all detected junctions, 85,923 of these junctions were shared between the *B. napus* subgenomes, 80,893 junctions were shared between *B. oleracea* and *B. rapa*, and 76,770 junctions were shared across all hosts. However, not all exonexon junctions contain active splice junctions. Using expression data to inform splicing potential, a total of 32,996 junctions exhibited AS event expression in at least one species. Additionally, similar rates of event expression were detected for each genome or subgenome (Table 2.4), of which mean number splice junctions equalled 14,542.

2.2.5 Quantitative analysis of *Brassica* host transcriptome similarity using homeologous expression patterns

To examine global transcriptome similarities between polyploid subgenomes and diploid progenitor genomes, species were clustered based on homeolog response to pathogen stress (Figure 2.7A). Overall, each subgenome of the recently synthesized *B. napus* appears to cluster with its respective parental genome, while the natural B. napus subgenomes form a distinct cluster. Given the recent genesis of the resynthesized *B. napus*, and the relatively small number of generations that have elapsed since it diverged from its progenitor genomes, this result would be expected. Considering the high sequence identity between *B. oleracea* and *B. rapa* orthologs, homeologs from the *B. napus* genome were also combined to reflect overall changes in transcript dosage (Figure 2.7B). Reminiscent of the GO enrichment analyses, the resulting dendrogram suggests that the overall pathogen stress response of the resynthesized *B. napus* more strongly resembles the response of B. oleracea than the response of B. rapa. Given the strong similarities between the individual A and C subgenomes and their respective progenitor genomes, this disparity between subgenomes is possibly indicative of an expression bias towards the C subgenome of the resynthesized *B. napus*.

Changes in host alternative splicing due to pathogen infection were also clus-

tered for each species (Figure 2.7C). When compared to the gene expression data, the resulting dendrogram displayed markedly different clustering patterns. While the alternative splicing response of the natural *B. napus* subgenomes also formed a cluster with *B. oleracea*, the resynthesized *B. napus* subgenomes exhibited similarity to the response of *B. rapa*. This is at odds with the comparative analysis of gene expression, and perhaps suggests that subgenome biases in alternative splicing ratios are independent of biases of constitutive transcript expression.

2.2.6 Analysis of subgenome expression bias demonstrates a slight C subgenome homeolog expression bias in both natural and synthetic *Brassica napus*

Given the expression biases suggested by interspecific transcriptome comparisons, it is possible that the C subgenome is preferentially expressed in the resynthesized *B. napus*. To investigate this potential subgenome expression bias, the relative expression levels of homeologs from the A and C subgenomes of the polyploid hosts were compared (Figure 2.8). This involved the assessment of differential homeolog expression among gene pairs that were expressed in both subgenomes.

Of the 26,663 the homeolog pairs analyzed in the mock samples of the resynthesized *B. napus*, there were 8562 pairs that were biased toward the C subgenome and 3471 pairs which exhibited an A subgenome bias (Figure 2.8A). This trend was even more prominent among the 26,284 homeologs expressed between the infected replicates, of which 10,306 homeologs showed an expression bias favouring the C subgenome and 4332 pairs with expression levels biased towards the A subgenome. While the majority of homeologs expressed in both conditions were equivalent in their expression levels (15,119 and 11,646 gene pairs, respectively), the resynthesized B. napus displays a significant preference for the C subgenome across both treatments ($FDR < 2.2 \times 10^{-16}$, Fisher's exact test). Although somewhat attenuated when compared to the subgenome expression bias shown in the resynthesized polyploid, the natural B. napus displayed a similarly C-biased pattern of homeolog expression (Figure 2.8A). This is illustrated by an imbalance in the frequency of C-biased gene pairs (4617 and 3961 in the mock and inoculated treatments, respectively) compared to A-biased gene pairs (3388 and 3004, respectively). While there is a notable difference in the number of C-biased gene pairs

between polyploid hosts, the natural *B. napus* also showed a slight, yet significant, preference toward the C subgenome ($FDR < 6.1 \times 10^{-9}$, Fisher's exact test).

Subgenome expression bias can be exemplified by an analysis of relative homeolog expression; however, previous studies have shown large scale epigenomic restructuring immediately after polyploidization (recently reviewed in [15, 16]); therefore, polyploids may also exhibit biases in the subgenome-specific silencing of genes. To assess this possible bias, the number of expressed homeologs was calculated for both polyploid hosts (Figure 2.8B). A reduction in the number of A subgenome genes expressed occurs across both hosts. In the natural *B. napus*, this was only evident in mock samples (P < 0.05, Wilcoxon rank sum test). However, the resynthesized *B. napus* exhibited a significant difference in both mock and inoculated samples (P < 0.05, Wilcoxon rank sum test). Additionally, transcriptomewide rates of gene expression were analyzed between the requisite resynthesized *B. napus* subgenomes due to the their known origins. The proportion of genes exhibiting FPKM > 1 were calculated, resulting in roughly equivalent rates of expression for each subgenomes (P = 0.51).

2.2.7 Analysis of subgenome-specific alternative splicing patterns in both natural and synthetic *Brassica napus*

While biased gene expression among allopolyploid subgenomes is well established, the extent to which alternative splicing levels vary between homeologous genes is less clear. Moreover, alternative splicing is a complex phenomenon and biases in alternative splicing patterns might manifest in different forms. For instance, consistent biases in the rate of AS event inclusion among events conserved between both subgenomes might suggest higher levels of AS in a single subgenome. However, due to the speed with which alternative splicing evolves relative to gene expression, it may be more appropriate to represent alternative splicing bias as an increase in the likelihood of AS event expression at a given intron. To provide a comprehensive assessment of potential subgenome biases in alternative splicing patterns, comparisons were made at three different resolutions: event-specific comparisons, exon-exon junction comparisons, and overall subgenome rates of AS.

To assess whether one subgenome had an overall higher number of alternatively spliced transcripts, AS ratios for each pair of homeologous splicing events were compared (Figure 2.9A). In contrast to the analyses of subgenome expression bias, the resulting AS bias was less pronounced in both natural and resynthesized polyploid hosts. In the resynthesized *B. napus* AS event expression seemed to favour the A subgenome, with mock-inoculated samples displaying 1644 A-biased events compared to 1457 C-biased events (P < 0.03). This trend was similar among infected samples, in which 1120 A-biased events were detected compared to 953 C-biased events (P = 0.07). However, the majority of events (62.6% Mock; 64.1% Infected) were not significantly different between A and C subgenomes. This frequency of unbiased events was similarly high in the natural *B. napus* (64.2% Mock; 67.9% Infected), yet unlike the resynthesized *B. napus*, the natural polyploid cultivar exhibited a slight C subgenome bias in AS ratios.

In addition to the rate of event inclusion, subgenomes may be biased in the diversity and frequency of AS event expression among homologous exon-exon junctions. To investigate such biases, the rate of event expression was calculated for the most common alternative splicing event classes — i.e. intron retention, alternative acceptor, alternative donor, and alternative position (Figure 2.9B). In the natural *B. napus*, the proportion of expressed events was consistently higher in the C subgenome; however, this difference was only significant among infected samples (P < 0.05, Wilcoxon rank sum test). This bias was more notable in the resynthesized polyploid host, in which there was a significant bias towards the C subgenome among both infected and uninfected samples for the most common event class — specifically intron retention (IR) and alternative acceptor (ALTA) (P < 0.05, Wilcoxon rank sum test). Interestingly, alternative position (ALTP) events exhibited a bias towards the A subgenome; however, ALTP events were relatively uncommon in comparison (they were detected in <1% of homeolog introns).

2.2.8 Influence of pathogen infection on homeolog expression bias in *B. napus* hosts

To appreciate the extent to which biases in gene expression or alternative splicing contribute to biotic stress responses, it is pertinent to assess how pathogen infection modulates patterns of subgenome expression bias. This was accomplished by assessing the retention of subgenome expression bias between mock-inoculated and *S. sclerotiorum*-inoculated conditions (Figure 2.10).

Of the 24,737 homeologous gene pairs in the resynthesized *B. napus* that were expressed in both conditions, 17,594 (71.1%) of the homeologs retained the same patterns of expression bias (Figure 2.10A). While the majority of these homeologs remained unbiased (8666, or 35.0% of all homeologs), homeologs exhibiting C subgenome bias were also highly consistent (6701; 26.3%). This was in contrast with the A subgenome, which was conserved among only 2537 homeologs (9.7%). Interestingly, the resynthesized *B. napus* displayed significantly more A and C subgenome biased homeologs under pathogen infection than during uninfected growth ($P < 2.2 \text{ x } 10^{-16}, \chi^2$ test). Similar to the resynthesized *B. napus*, genes in the naturally occurring *B. napus* exhibited consistent patterns of bias between both mock and infected samples (16,521 homeologs out of a total of 22,537 homeologs; 72.1%). However, unlike the synthetic polyploid, these homeologs consisted of mostly unbiased homeolog pairs (11,952, or 53.0% of all expressed homeologs). Of the homeologs which retained a subgenome bias, there were more C subgenome biased homeologs in both conditions (2511 C-biased pairs compared to 1788 A-biased pairs). However, these only consisted of 11.1% of total homeolog pairs. Additionally, across both subgenomes there were fewer homeologs showing unequal expression levels when the plants were subjected to pathogen stress (P <2.8 x 10⁻⁷).

To provide a more comprehensive assessment of pathogen-mediated changes in homeolog expression bias, shifts in homeolog expression ratios between infection conditions were quantitatively assessed using a permutation test based on the HomeoRoq pipeline [94]. In agreement with the assessment of pathogen-mediated changes in expression bias, the vast majority of homeolog expression ratios did not change significantly between conditions (87.7% in the resynthesized *B. napus* and 90.3% in the natural *B. napus*). Additionally, the majority of all homeologs were biased towards the C subgenome — 64.9% in the resynthesized *B. napus* and 54.6% in the natural *B. napus*. Of the homeologous pairs that were significantly different between the synthetic polyploid treatments (2832 pairs) (P < 0.05), 64.5% of homeolog expression ratios gained an enhanced bias in the pathogen infected treatment. An homeolog expression ratio analysis of the natural *B. napus* showed 1960 significant pathogen-altered homeolog ratios, which also largely shifted towards a greater expression bias (70.8%).

The influence of pathogen stress on subgenome splicing biases lies predominantly with splicing event identity, as comparison between uninfected and infected expression states reveals many condition-specific splicing events. This is most evident in the natural B. napus, for which there were 4966 events unique to uninfected samples and 3422 unique to infected samples. Of the 5771 events conserved between conditions, 75.8% retained the same splicing bias under both treatments (Figure 2.10B). Of the events that exhibited a notable bias, slightly more were C subgenome-biased (748 > 681); however, both subgenomes exhibited a notable decrease in the frequency of bias during the pathogen treatment (P < 0.002). The resynthesized B. napus demonstrated markedly different shifts in biases. While the mock-inoculated samples expressed a large number of condition-specific events (4311), the pathogen treatment was less divergent, expressing 1815 unique events. Of the 3970 events shared between conditions, 78.7% of events retained the same bias between conditions. Unlike the natural B. napus, these events suggested a slight bias towards the A subgenome (633 > 530). However, consistent with the natural B. napus, the resynthesized polyploid also exhibited a significant decrease in bias during pathogen infection (P < 0.003).

2.2.9 Expression level dominance in the resynthesized *B. napus*

Analyses of expression level dominance (ELD) have been used previously to summarize changes in expression pattern between parental transcriptomes and descendant polyploid sub-transcriptomes [25, 26, 30]. This involves the comparison of parental expression levels to the concatenated expression levels of a given homeologous gene pair, thereby facilitating the detection of additive, transgressive, and parental-dominant expression patterns. Considering the fact that the transcriptomes of contemporary *B. rapa* and *B. oleracea* cultivars do not reflect the true parentpolyploid relationship of natural *B. napus*, comparisons were only made if the parental genotypes were direct constituents of the polyploid genome. Using the resynthesized *B. napus* transcriptome, this resulted in the analysis of 26,014 homeolog pairs among the mock-inoculated samples and 26,001 homeolog pairs among the pathogen-inoculated samples, which were further binned into 13 expression pattern categories (Figure 2.11). In general, most homeologous gene pairs across both the mock and pathogeninoculated treatments exhibited no significant differences in expression (57.0% and 60.2% of homeolog pairs, respectively). Of the remaining categories, the most commonly represented were C-expression level dominant, A-expression level dominant, and transgressively down-regulated genes. Categories related to expression level dominance were among the most common of the aforementioned categories, consisting of 29.8% of genes among the mock-inoculated samples and 30.0% of all genes in the pathogen-inoculated samples. Of these genes, significantly more were C subgenome dominant (5318 and 5603 genes in the mock and pathogen-inoculated samples, respectively) than A subgenome dominant (2423 and 2231 genes) (P < 0.001, Fischer's exact test).

When comparing expression-level categories between mock and pathogen-inoculated treatments, category distribution differed notably. While additively and transgressively up-regulated genes exhibited consistently low incidence and varied little between treatments, most other categories showed significant change while under pathogen stress. For instance, a higher number of genes displayed equivalent and C-dominant expression during infection (P < 0.01, Fischer's exact test). There was also a reduction in number of transgressively down-regulated genes and A-dominant genes in the infected treatment (P < 0.01, Fischer's exact test).

To better understand how pathogen infection might influence these comparative expression levels, shifts in expression level category between treatments were also analyzed. Overall, a total of 25,299 homeologous gene pairs were actively expressed in both mock and pathogen-inoculated samples, which were sorted into 78 categorical comparisons (Figure 2.11). Of these inter-conditional comparisons, the majority (52.5%) of genes were categorically consistent between treatments. This included categories such as equivalent (10,530 genes) and C dominant (1248 and 391 genes) expression levels. Categories that displayed the most dramatic interconditional change involved a shift towards or from equivalency. For instance, C-dominant to equivalent, equivalent to C-dominant, A-dominant to equivalent, and transgressive down-regulation to equivalent.

2.2.10 splicing level dominance in the resynthesized *B. napus*

As previously described, alternative splicing is often depicted as the proportion of event-supporting reads in a given junction (PSI). Given the large number of conserved alternative splicing events that remain between homeologous gene pairs, their splicing ratios were consolidated into a transcriptome-wide rate of event inclusion. Similar to ELD, comparisons of these ratios to their respective ancestral orthologs could elucidate general changes in AS patterns as a result of polyploidization. To identify such deviations from the ancestral splicing patterns of homeologous events, differential splicing ratios were calculated for the homeologous events shared between the resynthesized *B. napus* and its parents for both mock (8074 events) and inoculated (7187) conditions. Analogous to analyses of ELD, these splicing events were then sorted into 13 categories (Figure 2.11).

Across both conditions, most homeologous splicing events did not display significant change between the polyploid transcriptome and those of its parental genotypes. This category constituted 66.7% (5431) of conserved mock-inoculated splicing events and 68.7% (4959) of total pathogen-inoculated splicing events. A large number of events also exhibited bias towards one of the parental subgenomes. C subgenome dominant splicing ratios was the most frequent dominance category, representing 17.7% of mock-inoculated events and 15.6% of pathogen infected samples. However, A subgenome dominant ratios were substantially higher than additive and non-parental (transgressive) categories, which comprised 10.2% and 9.6% of mock and pathogen-inoculated events, respectively. Interestingly, a large number of transgressively up-regulated alternative splicing events were detected, in contrast to ELD categorization.

Similar to ELD, distributions of parent-polyploid splicing levels varied significantly between control and pathogen infected conditions. An overall reduction in frequency was observed for most categories, notably transgressively up-regulated, highly spliced A- and C-dominant events, and additive events favouring the expression of the *B. rapa* homeolog (P < 0.01, Fisher's exact test). Interestingly, the frequency of polyploid splicing events that favoured the lowly expressed parental subgenome increased under pathogen treatment. Considering the broad changes in parental splicing level categorization, splicing level patterns were analyzed for

each event individually. Of all homeologous events expressed across both mock and pathogen conditions, only 31.0% (4757) of events were shared between the two expression states. Additionally, 65.7% of these conserved events showed no change in splicing level categorization. For example, equivalent splicing levels, Cdominant, and A-dominant events alone comprised 63.6% of retained splicing level patterns. Of the categories that demonstrated shifts in splicing level patterns, the majority (23.0% of all cross-conditionally conserved events) alternated between equivalent expression and A- or C subgenome dominance.

2.2.11 Homeolog-specific expression categorization in the resynthesized *B. napus*

While analyses of ELD can be useful for inferring patterns of homeolog transcript dosage in polyploid-parent expression levels, homeolog expression does not always scale additively and combined homeolog expression will obscure possibly informative non-parental expression patterns involving a specific homeolog. To investigate the contribution of each homeolog individually, comparisons were made between each polyploid subgenome and their respective parental genomes (Figure 2.12).

Comparisons of general gene expression patterns in the resynthesized *B. napus* and its constituent parental genotypes revealed a high similarity in the identity of genes expressed (FPKM > 1; Figure 2.12A-B). Among the uninfected samples, 96.2% of all gene expressed from the A subgenome were also expressed in *B. rapa*. Similarly, the C subgenome exhibited 95.4% overlap with *B. oleracea*. This overlap was consistently high in the infected samples, of which the A subgenome showed a 98.4% overlap with its parental transcriptome and the C subgenome, again, showed a 95.4% overlap with *B. oleracea*. Additionally, the polyploid subgenomes appear to express fewer genes than their diploid counterparts across both conditions. For instance, 10.0% of *B. oleracea* genes (across both conditions) are either lowly expressed or silenced in the resynthesized *B. napus*. This disparity between subgenome and diploid gene expression patterns is significantly more pronounced in *B. rapa*, of which 16.5% and 21.0% of genes are diploid-specific under uninfected and infected conditions, respectively ($P < 2.2 \ge 10^{-16}$, Fisher's exact test). To better understand possible divergence and conservation of subgenome-specific expression patterns in the resynthesized *B. napus*, when compared to *B. rapa* and *B. oleracea*, homeolog-ortholog expression patterns were assigned to 27 expression-level categories. Of these classifications, the predominant categories are equivalence, ancestral expression, subgenome dominance, and transgressive downregulated (Figure 2.12). The most common category across both conditions is equivalence, containing 46.3% and 43.2% of homeologs in uninfected and infected conditions, respectively. This is followed by ancestral expression patterns, which are represented by 21.0% and 25.1% of homeologs. Of the ancestral expression patterns, the majority (16.0% and 19.0% of all homeologs) favour C subgenome expression. This is similar to the subgenome-dominant expression categories, of which C dominance is the most frequently observed pattern. Interestingly, of the transgressive down-regulated homeolog categories, the vast majority of homeologs exhibited a significant reduction of expression in the A subgenome homeolog (5.5% and 8.2% of all homeologs).

To help elucidate the origins of subgenome expression dominance in the newly formed B. napus, relationships between homeologs and their corresponding parental genes were compared to polyploid genes exhibiting significant bias (absolute $\log_2 FC$ (A/C) > 1 and a FDR < 0.01). The resulting expression categorization revealed differing origins of expression bias between A and C subgenome (Figure 2.12E). The majority C subgenome-dominant homeologs conformed to ancestral expression patterns (58.6% and 57.9% of C-biased homeologs in uninfected and infected samples, resectively). This was followed by transgressive down-regulated A-homeologs (20.1% and 25.0%), which were also common among C-biased homeolog pairs. While the A subgenome also showed a high frequency of ancestral dominance (43.8% and 48.3% of A biased homeologs) and transgressive downregulation of the C homeolog (17.7% and 11.8%), there were a high proportion of A homeologs that had expression levels comparable to *B. oleracea* rather than *B.* rapa (15.0% and 14.0%). These were referred to as reciprocally dominant homeologs due to the similarity in expression level to the non-parental homeolog. Furthermore, it is also possible for homeologs to exhibit no bias within the polyploid yet have a notable bias towards one of the parents, suggesting a possible regulatory dominance (Figure 2.12E). While these parental-polyploid expression patterns

were somewhat uncommon in the context of an A subgenome bias (<5%), there were a considerable number of homeologs that exhibited a putative C subgenome regulatory dominance (8.1% and 7.2% of total homeolog pairs).

The changes in distribution of homeolog-specific expression level categories were also analyzed between uninfected and infected conditions. Most categories were reduced as a result of pathogen stress, most notably equivalent and regulatory dominant expression categories (P < 0.01, Fisher's exact test). However, the retention of ancestral gene expression patterns increased from a total of 21.0% and 25.1% of all homeologs ($P < 3.1 \times 10^{-7}$, Fisher's exact test). The frequency of transgressive down-regulated A subgenome homeologs also rose significantly from 5.5% to 8.2% ($P = 5.5 \times 10^{-34}$, Fisher's exact test). Examining the change in subgenome-specific expression levels was consistent with the ELD analysis, with 51.6% of homeologs retaining the same category assignment across conditions. Of homeologs that retained category assignment across treatments, the most frequent expression patterns were equivalent expression, ancestral C expression bias, and transgressive down-regulation of the A homeolog (58.0%, 18.9%, 6.9%) of homeologs with retained expression patterns, respectively). Furthermore, the majority of pathogen-mediated fluctuations in category involved a shifts between equivalent expression and a gain in ancestral C-subgenome bias or equivalent expression and regulatory C dominance (16.1% and 14.9% of categorically changed homeologs).

General comparisons of parent-polyploid splicing event identity revealed widespread changes in AS event expression of the resynthesized *B. napus*. An assessment of event overlap between the *B. napus* A subgenome and *B. rapa* showed an overlap of 28,623 events (76.4% of A subgenome events and 70.1% of *B. rapa* events). Among infected samples, 23,783 events overlapped — representing 85.5% of the A subgenome events and only 59.6% of *B. rapa* events. Overall, there were 16,577 events exclusive to *B. rapa*, and of these events, 33.4% were shared between conditions. Interestingly, 9481 alternative splicing events were novel in the resynthesized polyploid. These events were also generally specific to either uninfected and infected expression states, with 10.5% of events retained between conditions. Comparisons between the *B. napus* C subgenome and *B. oleracea* resulted in similar trends of polyploid-parent alternative splicing conservation. In uninfected sam-

ples, 41,032 events were conserved (77.2% of C subgenome events and 78.8% of *B. oleracea* events). Infected samples showed similar rates of conservation, with 32,497 overlapping events (79.3% of C subgenome events and 73.5% of *B. oleracea* events). Homeologous alternative splicing events originating from *B. oleracea* also exhibited low levels of conservation. Of *B. oleracea* genes, 14,822 were not expressed in the polyploid and 13,273 events were novel to the resynthesized *B. napus*.

Analogous to parental-polyploid expression level patterns, parental-polyploid alternative splicing ratios were also divided into 27 categories. Of these categories, the most frequently observed were equivalent (37.4% in the mock inoculated samples and 39.7% of the pathogen-inoculated samples), non-equivalent but ancestral in origin (24.5% and 23.0%), transgressively up-regulated (12.5% and 8.7), and subgenome dominant (10.5% and 12.1%) splicing levels (Figure 2.12E). As with gene expression, these analyses of splicing ratios also provide a better understanding of homeolog splicing bias. Of the homeologous AS events with a C subgenome bias, the majority of events exhibited ancestral splicing ratios (57.1% and 58.5% of biased splicing events in uninfected and infected conditions, respectively). In contrast to parent-polyploid gene expression patterns, the second most common AS category among C-biased homeologous events was transgressive upregulation (16.3% and 10.5%). Events biased towards the A subgenome were similarly categorized. Ancestral splicing patterns were the most common source of A-biased homeologous event splicing (44.8% and 43.5%). Interestingly, a significantly higher number (P < 0.01, Fisher's exact test) of transgressive up-regulated events contributed to splicing bias in the A subgenome (33.7% and 26.7% of Abiased AS event in uninfected and infected conditions). Intriguingly, a number of homeologous events also exhibited a possible regulatory dominance — in other words, both homeologs were biased to one of the ancestral subgenomes. Within this particular splicing categorization, 4.4% and 5.3% of all homeologous events were biased towards B. rapa and 6.1% and 6.8% were biased towards B. oleracea.

There were also notable shifts in the frequency of homeolog splicing level categories between uninfected and infected conditions. While the frequency of the most common categories remained consistent (equivalent and ancestral splicing levels), a significant decrease in the frequency of subgenome dominant events was observed with high splicing ratios ($P < 3.5 \times 10^{-5}$, Fisher's exact test). Furthermore, the frequency of dominant subgenomes with low splicing ratios increased following pathogen treatment. Tracking the change in categorization of specific events, similar to the analysis of global splicing ratios, highlighted the conditionspecific nature of alernative splicing event expression. Between homeologs and parental orthologs, 3012 events were shared across conditions, totalling in 39.9% of all events. Of the events that were preserved between conditions, 59.7% retained the same the splicing level category and an additional 10.6% exhibited ancestral splicing patterns. The remaining changes in splicing pattern involved a shift from parental dominance towards ancestral expression when subjected to pathogen stress (8.2%).

2.2.12 Impact of asymmetric subgenome expression on the pathogen stress response of polyploid hosts

To assess whether subgenome expression bias might play a role in the pathogen stress response of the resynthesized *B. napus*, an analysis of induction bias was performed. As described in Powell et al. [32], subgenome induction bias assesses the relative contribution of homeologs to the combined homeolog stress response. To accomplish this, read counts were summed across A and C homeologs and a differential gene expression analysis was performed using edgeR (FDR < 0.01). The generalized subgenome response to S. sclerotiorum was then compared to the relative induction of each homeolog. Out of 3093 up-regulated (FC > 1) gene pairs, the majority (2725) of homeologs exhibited symmetric induction. However, C homeologs were induced more frequently than A homeologs, suggesting a slight C subgenome induction bias (244 > 122) (P < 0.001, χ^2 test). Furthermore, to resolve the relative subgenome contribution to the homeolog-consolidated transcriptome response, differentially expressed genes were assessed for underlying subgenome biases. Among the up-regulated gene cohort, the majority (1695 of 3162) of genes exhibit a C subgenome bias ($P < 2 \ge 10^{-16}$, χ^2 test). This is consistent with the down-regulated gene set, of which 49.4% are significantly C subgenome biased (compared to the 19.3% which display an A subgenome bias). Taken together, this suggests that the pathogen stress response of the resynthesized *B. napus* is disproportionately influenced by the transcriptional contribution of *B.*

oleracea.

In addition to an examination of induction bias, genes previously implicated in a defense response against S. sclerotiorum were analyzed for evidence of biased subgenome expression. Overall, among pathogen infected samples, there were roughly equivalent proportions of genes expressed from each subgenome (37.0% of genes originated from either A or C homeologs, respectively). However, a clear C subgenome bias in the expression level of homeologs was found to be associated with a defense response (Figure 2.13B). Of the pathogen-responsive homeologs related to hormone biosynthesis, transcriptional regulation, or pattern recognition, 328/639 displayed a C subgenome bias, whereas 140/639 displayed an A subgenome bias ($P < 5.2 \times 10^{-5}$) (Figure 2.13). Furthermore, the mean ratio of expression between A and C homeologs, represented by log(A/C), was -1.07 in the infected samples. In other words, the average expression level of homeologs associated with a pathogen response was 2.1 times greater in the C subgenome. This expression bias was consistent across most families that were highly up-regulated during pathogen stress and notably greatest among PEP receptors, wall-associated receptor kinases, ethylene biosynthesis genes, and chitin receptor kinases (Figure 2.13A).

To illustrate the impact subgenome expression asymmetry might have on the pathogen stress response of resynthesized *B. napus*, expression patterns of homeologs from the ethylene biosynthesis pathway were analyzed (Figure 2.13C). The plant hormone ethylene (ET), along with jasmonic acid (JA) are integral aspects of a plant response to a necrotrophic pathogen, and consequently, important signalling molecules in the *Brassica-Sclerotinia* pathosystem [64]. Furthermore, I found ET biosynthesis and signalling genes were consistently induced across all four *Brassica* host transcriptomes (Figure 2.4), allowing clear interspecific comparisons. Using gene lists assembled from *A. thaliana* and *B. napus* orthology, I compared 10 sets of orthologs encoding ET biosynthesis enzymes, and 4 sets of orthologs implicated in downstream ET and ET/JA signalling. The process of ethylene biosynthesis begins with the conversion of the amino acid methionine to S-adenosylmethionine via methionine adenosyltransferase (MAT). Examination of host MTO3/MAT4 expression revealed strong C homeolog bias across both sets of genes (*FDR* < 1.6 x 10⁻⁴). Interestingly, this seemed to involve apparent trans-

gressive down-regulation of the A homeolog and ancestrally high C subgenome expression. This was followed by the conversion of S-adenosylmethionine to 1aminocyclopropane-1-carboxylic acid (ACC) by 1-aminocyclopropane-1-carboxylate synthase (ACS), a major rate-limiting step of the pathway. In contrast to MAT homeologs, consistent induction of the A homeolog occurred among (ACS) genes. Two sets of genes (ACS1 and ACS2) demonstrated intermediate expression levels with a bias towards the C homeolog (FDR < 0.005); however, ACS6 seemed markedly biased towards the ancestral *B*. rapa expression level (FDR < 0.001). Finally, ACC is then converted to ethylene via 1-aminocyclopropane-1-carboxylate oxidase (ACO). Of these four ortholog gene sets, the majority of highly up-regulated genes are significantly biased towards the C homeolog ($FDR < 7.1 \times 10^{-8}$). ET signalling proceeds to elicit the expression of several TFs and pathogen-defense genes. Interestingly, examining a subset of the most highly up-regulated genes involved in ET/JA signalling reveals similar expression level patterns. Homeologous PDF1.2c and PR4 expression follow the ancestral B. oleracea expression pattern $(FDR < 1.8 \times 10^{-9})$. A and C PR3 homeologs show a slight *B. oleracea*-derived attenuation when compared to the ancestral PR3 response. Contrastingly, ERF1 exhibits a slight A subgenome bias but patterns of induction are relatively consistent between homeologs.

2.3 Discussion

2.3.1 Transcriptome response of *Brassica napus* and its progenitor species to the pathogen *Sclerotinia sclerotiorum*

To investigate the interplay between polyploid subgenomes under biotic stress, a natural cultivar of *Brassica napus*, a resynthesized allopolyploid *Brassica napus*, and its corresponding parental genotypes of *Brassica rapa* and *Brassica oleracea* were infected with the necrotrophic pathogen *Sclerotinia sclerotiorum*. Host defense in the *Brassica-Sclerotinia* pathosystem is largely characterized by quantitative resistance and therefore considerable emphasis was placed on comparisons between the response of the resynthesized polyploid and that of its direct parental genotypes — i.e. the diploids *B. rapa* and *B. oleracea*. Furthermore, the specific

patterns of host response are significantly associated with specific homeologous regions on both sets of diploid chromosomes [67], and consequently, the resulting genomic and transcriptomic shock of allopolyploidization may have dramatic effects on the expression of genes involved in the defense response of the polyploid.

Global transcriptome analyses of the four host species at 24 hours post-infection (hpi) revealed large scale changes in host expression-state. While considerably more genes were down-regulated in the polyploids, which could be attributed to the transcriptome shock following polyploidization, the resulting expression changes were mostly consistent between *Brassica* hosts. This is exemplified by the interspecific congruence of the GO enrichment among up-regulated genes (Figure 2.3A). These co-enriched ontologies — for example, defense response, response to wounding, and innate immune response — are highly consistent with patterntriggered immunity (PTI). PTI is characterized by the recognition of pathogenassociated molecular patterns (PAMPs) or damage-associated molecular patterns (DAMPs) and is a form of innate plant immunity elicited in response to infection from a wide variety of pathogen hosts [95, 96]. Considering the fact that S. sclerotiorum is a broad host-range necrotroph, and there is little evidence to support a pathogen-specific resistance for *Brassica-Sclerotinia* pathosystems [89, 95], all assayed hosts likely share similar PTI-mediated resistance to the pathogen. Using genes outlined in previous studies [63, 89], I identified orthologs for all Brassica hosts species to interrogate their defense response in more detail. The resulting analyses showed similarity to previous studies; however, direct interspecific comparisons are rendered inadequate by possible differences in host genotype and differences in expression could be the result of disparities in resistance level between hosts. To account for this, I only considered genes that were similarly up-regulated between all hosts, with particular emphasis on the parent-progeny relationship between the diploids and the resynthesized *B. napus*. Overall, across all species concerted up-regulation was observed for MAPK-module genes, WRKY TFs (WRKY6), jasminic acid and ethylene biosynthesis and signalling genes, and chitinases.

Although a notable degree of overlap exists between the four *Brassica* hosts, *B. rapa* appears to differ considerably (Figure 2.4). For example, the expression of receptor kinases appears somewhat attenuated compared to the polyploids and *B.*

oleracea, with exception of PEP receptor genes. Additionally, ET-associated genes are more strongly induced. Interestingly, among the polyploids and *B. oleracea*, genes related to a cellular response to hypoxia and cellular oxidant detoxification are significantly enriched. This is interesting given the relationship between *S. sclerotiorum* and reactive oxygen species metabolism [63, 95]: oxalic acid produced by the pathogen will first suppress ROS signalling, but will later promote ROS production and cell death. Overall, this suggests that perhaps the response of *B. rapa* to pathogen stress differs in comparison to the allopolyploids and *B. oleracea*. Importantly, such a discordance between the resynthesized *B. napus* and the response of its parental genotypes might imply that this recent allopolyploid is biased towards a C subgenome immune response contribution. Quantitative comparison of transcriptome-wide responses to *S. sclerotiorum* further validates this claim. While the A subgenome exhibits a stress response consistent with *B. rapa*, the combined response of both homeologs is more similar to *B. oleracea*, suggesting the expression of C subgenome homeologs is considerably higher.

Unlike gene expression, few studies have directly investigated the role of alternative splicing in the regulation of a pathogen stress response [97]. While some studies have used tools — such as Cufflinks [98] — to infer alternative transcript isoforms [89], transcript resolution might be best investigated through a composite of both long and short read sequencing [99–101]. Furthermore, the expression of alternatively spliced isoforms does not always result in a coordinated shift from one major transcript isoform to another. In plants, alternative splicing often acts to modulate the ratio of transcripts that actively contribute to the proteome; therefore, AS might represent an indirect form of gene regulation [102]. As our data were generated with short read sequencing, and to capture transcriptome-wide AS changes that may not be correlated with major transcript isoforms, I opted for a junction-centric approach to alternative splicing analysis. The increased sensitivity offered by this method of AS quantitation allows us to assess widespread fluctuations in splicing and the extent to which these patterns of AS may be modulated by pathogen stress.

I identified widespread changes in alternative splicing between uninfected and infected expression states. Across all hosts, there were a considerable number of infected-specific events (ranging from 27.7% to 21.7% of all infected AS events).

Additionally, only approximately 52.6% of all AS events detected within each host were shared between infection conditions. Among the splicing events that are shared between uninfected and infected samples, approximately 10-13% of events were differential alternative splicing (DAS) in all hosts. Similar to recent abiotic stress analyses [103], these were not functionally consistent with differentially expressed genes. In fact, up-regulated GO categories were inconsistent between species. However, these did share similar biological themes, most of which are consistent with down-regulated differentially expressed genes. Similarly, downregulated DAS genes were highly enriched for genes involved in redox reactions and occasionally defense response. Accordingly, I also identified a moderate negative correlation between gene expression and alternative splicing (Figure 2.6A). Significantly enriched gene ontologies support this relationship (Figure 2.5), as GO terms enriched among up-regulated alternative splicing events are thematically related to down-regulated constitutively expressed transcripts (e.g. energy metabolism and photoresipiration). This can also be seen by directly analyzing AS patterns of genes involved in pathogen defense. Changes in splicing ratio and expression tended to be directionally opposed (Figure 2.6B, represented by $\log_2 FC$). No evidence supports a causal link between alternative splicing and constitutive transcript isoform expression; however, a recent study has shown drought stress to enhance splicing efficiency of introns belonging to stress-related genes [104]. Furthermore, there is evidence of AS involvement in transcript regulation [36, 105]. Studies show that genic contribution to the proteome is substantially reduced by enhanced expression of transcript isoforms subject to nonsense-mediated decay (NMD) [38]. In this case, the constitutive transcript is indirectly negatively regulated as its dosage is diluted by the increased proportion of alternatively spliced transcript isoforms. However, molecular evidence is needed to confirm such speculations.

2.3.2 Global patterns of subgenome expression bias in both natural and resynthesized *Brassica napus*

Quantitative analyses of transcriptome-wide reprogramming between polyploid *Brassica* hosts were suggestive of a notable C subgenome expression bias. To more thoroughly investigate the possibility of subgenome expression bias, I evaluated

global differences in subgenome expression for both the natural and resynthesized *B. napus*. Prior to analyses of homeolog expression, I recognized biases between subgenome mapping rates could lead to erroneous differences in homeolog expression. To assess whether this might influence our inference of expression bias, I analyzed our dataset using several read-sorting methods [94, 106]. These results were in strong agreement with the use of STAR-aligned samples with stringent filters for read-mapping ambiguity [107] — this method also resulted in a notably higher retention of data.

Our analyses of homeologous expression in the natural and resynthesized *B. napus* confirmed both polyploid *Brassicas* exhibited a significant C subgenome bias. In corroboration with previous studies [30, 53], this bias towards the *B. oleracea*-derived subgenome was slightly attenuated in the natural *B. napus*. In contrast, the resynthesized *B. napus* exhibited extensive C subgenome expression bias (Figure 2.8). This result disagrees with the previous findings in other polyploids [26, 28, 108], which suggest expression bias increases over time. In the case of Yoo et al., the homeolog expression bias represents 1-2 million years of polyploid evolution, compared to the relatively recent origin of the natural *B. napus* (< 10,000 years). Considering this difference in age, it is possible extensive expression bias has yet to develop. Additionally, the natural *B. napus* has undergone extensive cultivation, which has been found to influence homeolog expression [26], and therefore could play a role in expression bias formation.

Furthermore, patterns of subgenome expression bias in *B. napus* are inconsistent among resynthesized varieties of *B. napus*, as different studies have shown evidence of both A and C subgenome dominance [18, 30, 53, 108]. For example, our analyses of subgenome expression bias and expression level dominance (ELD) in resynthesized *B. napus* showed extensive C subgenome expression dominance. However, previous studies demonstrated a slight bias towards the A subgenome despite sharing similar frequencies of ELD categories [30, 108]. Interestingly, this inconsistency in the establishment of a dominant subgenome could also be related to the fact that transposable element (TE) density is not strongly correlated with homeolog expression in *B. napus*. In opposition to the prevailing hypothesis of subgenome dominance, which cites TE-associated gene silencing as a prevailing factor in the establishment of a dominant subgenome, *B. napus* may be

more strongly influenced by the rate and extent of homeologous exchanges (HE) [53, 109]. For example, recent work in *B. napus* demonstrated large-scale changes to homeolog dosage were the result of HEs [110]. This is important to recognize when performing analyses of differential homeolog expression, as underestimating the extent and influence of HEs, or other large-scale chromosomal changes, could lead to erroneous conclusions regarding the causes or extent of homeolog expression biases.

2.3.3 Expression bias in the resynthesized *Brassica napus* is underpinned by ancestral expression and a high degree of transgressive down regulation

Many past analyses of expression dominance between polyploid subgenomes were not informed by parental transcriptome data [32, 94, 111, 112]. While unavoidable in many organisms due to the age of the polyploid, this is unlikely to provide evolutionary context for the origin of such biases as these analyses are predicated on the assumption of an equal parental expression ratio. To better understand the extent to which subgenome expression bias reflect the ancestral expression states of the parents, previous studies have assessed the individual homeolog contribution but only in the context of ELD-assigned categories [26, 30]. In this study I performed quantitative comparisons of expression between genes in the diploid of the resynthesized *B. napus* and their parental homologs to evaluate the parental origin of homeolog expression dominance. Expectedly, I found that the majority of subgenome expression bias originated from ancestral expression patterns among both A-biased and C-biased homeolog pairs (Figure 2.12B). Additionally, many biased homeologs also exhibited transgressive down-regulation of the opposite homeolog, supporting the idea that large-scale epigenomic or structural changes — such as HEs — may be responsible for the establishment of subgenome bias in B. napus. Interestingly, reciprocal dominance was common among homeologs biased towards the A subgenome - in which the A subgenome homeolog was more similar to B. oleracea in expression level than it was to either genes from the C subgenome or B. rapa. This curious pattern of transgressive up-regulation was suggestive of a possible regulatory mismatch between the ancestral B. oleracea and B. rapa genomes, which is hypothesized to play a role in the evolution of subgenome

dominance — discussed in [113]. Similarly, a number of homeologs exhibited comparatively invariant expression yet displayed a preference towards one of the two parental genomes (Figure 2.12B). While this suggests there is no homeolog expression dominance, it is indicative of a prospective regulatory dominance. For instance, the trans effectors might have affinity towards both subgenomes, or alternatively, one of the parental regulators is negligibly expressed. Further study is needed to both validate and investigate definitively; nevertheless, if the intention of the study is to investigate the evolutionary circumstances that led to homeolog expression dominance, these cases highlight the importance of assessing parental expression patterns in concert with subgenome expression bias.

2.3.4 Global bias in alternative splicing ratios in both natural and resynthesized *Brassica napus*

As previously discussed, subgenome expression bias and the emergence of subgenome dominance has been comprehensively studied. However, the extent to which alternative splicing (AS) patterns have diverged between allopolyploid subgenomes is less understood. Using the Pacific Biosciences long read sequencing platform, an analysis of isoform expression in tetraploid cotton has shown homeologous genes to be divergent in the number of transcript isoforms; however, they did not identify an overall bias in the frequnecy of alternative splicing [48]. While long read sequencing is superior to short read data in the resolution of unique subgenome-specific isoforms, in its current state, it is difficult to determine specific inclusion ratios using long read data. In this study, I address the difficulty of using short read data to investigate alternative splicing by adopting a junction-specific approach to alternative splicing detection. In doing so, I have determined the extent to which subgenome-specific AS events in *B. napus* have diverged in terms of AS event identity and splicing ratio.

Due to the age of the natural *B. napus*, and consequently the lack of true parental transcriptomes, I first studied subgenome AS patterns within each polyploid individually. Given the functional and regulatory complexities of AS, multiple comparisons of subgenome-specific AS were performed. To provide an understanding of how subgenomes differed in their overall expression of non-canonical transcript isoforms, I first explored disparities in splicing ratio bias between home-

ologous events. Both polyploids exhibited bias in splicing levels — which constituted approximately 42% and 44% of all AS events in the natural and resynthesized B. napus, respectively. However, natural B. napus yielded no evidence of significant subgenome splicing level dominance. Interestingly, a slight A subgenome bias was evident in the resynthesized *B. napus*. Similar to Wang et al. [48], another method of representing a possible AS bias is to assess the differences in AS occurence in each subgenome. To examine this I quantified the frequencies of each class of AS events to identify whether one particular subgenome exhibited higher levels of AS expression. While both polyploids tended towards a slight C subgenome bias in AS frequency, results showed that subgenome biases in AS event type were stronger in the nascent polyploid than in the natural polyploid. When considering both approaches to assessing splicing bias, splicing patterns are more evident in the younger polyploid line. While our study provides evidence of the attenuation of AS bias over time, additional work is needed to confirm whether polyploid age has a consistently strong influence on subgenome AS bias. Furthermore, it is not known if polyploid AS expression stabilizes over time, or whether subgenome-specific AS events gradually diverge in expression, retaining events conserved in expression pattern.

It is also important to note that the interaction of subgenomes, and notably subgenome-specific patterns of alternative splicing regulation, could result in the development of non-additive transcript splicing ratios. For example, trans-acting splicing factors might influence both constituent subgenomes, producing subgenome-dominant splicing patterns. Additionally, transcriptome shock could result in combined splicing ratios that are transgressively higher or lower than the parental splicing patterns. I investigated this by quantitatively comparing splicing levels between the resynthesized *B. napus* and its parental transcriptomes. Looking first at the previously described splicing level dominance categories — which involved the summation of all junction-mapped reads between subgenomes — I identified a number of AS events that deviated from expected splicing levels (Figure 2.11). These non-additive splicing levels involved a high number of events showing subgenome dominance and transgressive up-regulation, which were further investigated using a homeolog-specific approach to polyploid-parent splicing level categorization (Figure 2.12). This allowed us to establish putative origins for events exhibiting splic-

ing level bias. Similar to constitutive homeolog expression, the majority of biased events were the result of polyploid subgenomes expressing ancestral proportions of splicing events (40.2% of A dominant events and 51.4% of C dominant events). Interestingly, the second most common cause of subgenome splicing level dominance was transgressive up-regulation (43.9% of A dominant events and 27.1% of C dominant events), suggesting emergent changes in AS following polyploidization. There were also notable proportions of events in which both subgenomes exhibited dominance towards one of the ancestral splicing patterns (10.5% and 12.1% under mock and infected conditions, respectively). This pattern would not immediately appear evident in analyses of subgenome splicing bias or ELD, but would still represent non-additive splicing patterns. Such a case could be an example of the influence of subgenome-dominant trans-acting splicing regulators influencing splicing of both homeologs equivalently, although further research is needed to validate this possibility.

Altogether, the interplay between subgenome dominance and the emergence of polyploid AS patterns is complex. Patterns of subgenome-specific AS do not directly parallel gene expression. It appears that patterns of AS bias manifest in two distinct phenomena: (1) the splicing level, or proportion of event-supporting reads, for each event, and (2) the identity and number of AS events that are expressed. While the A subgenome displayed higher rates of AS, the number of events that were conserved between subgenomes represented only a fraction of the total AS detected in a given subgenome. Therefore, special attention has to be given to the extent to which AS event identity is shaped by polyploidization.

2.3.5 Non-parental AS event expression in the resynthesized *Brassica* napus

Our comparisons between polyploid subgenomes and their parental AS events show that polyploidization has profound effects on alloployploid AS event identity. While there is notable event overlap between the resynthesized *B. napus* and its parents, a high number of parental events were not expressed in either polyploid subgenome (Figure 2.12A). This suggests a general reduction in AS event expression between the diploids and the resynthesized *B. napus* under normal growth conditions. Additionally, a notable portion of the events were non-parental in origin. These include events that were not found in either parent, in addition to events that likely originated from the alternate parental genome. For example, 19.0% of AS events only found in the polyploid A subgenome were homeologous with AS events expressed in the *B. oleracea* genome. This highlights the potential for polyploidization to result in the emergent expression of non-parental AS events possibly due to the sudden integration of differing AS regulatory networks. Additionally, these cross-parental comparisons of AS identity are non-exhaustive and identified in genes found to be highly similar in terms of both sequence identity and exon organization. Therefore, it is possible that the extent to which AS regulatory networks in the polyploid are modulated by the alternate parent may be underestimated in this study. Furthermore, it is important not to divorce the role of homeolog expression bias and alternative splicing biases. Biases in splicing patterns or the emergence of non-parental splicing events could be result from unequal homeologous transcript expression. Future subgenome comparisons of splicing regulators, both in terms of their splicing patterns and overall gene expression patterns, might help elucidate the origins of splicing patterns unique to the polyploid.

2.3.6 Biotic stress modulates patterns of subgenome expression and splicing biases

Earlier work in hexaploid wheat has demonstrated that pathogen stress can result in a significant increase in homeolog expression bias [32]. This was somewhat consistent with our study in *Brassica napus*. Analyses of cross-conditional homeolog expression bias in the resynthesized *B. napus* revealed a significant increase in bias under pathogen stress. However, the natural B. napus exhibited roughly consistent levels of homeolog expression bias between treatments, indicating a pathogen-mediated increase in bias is not universal and may depend on the nature of the polyploid as well as other factors. Importantly, in either case the majority of genes retained the same patterns of bias between expression states. This suggests that, while pathogen stress can act to modify homeolog expression bias, the regulation of homeolog bias is stable.

Using parental transcriptome data from the resynthesized *B. napus*, I also demonstrated that the parental patterns of homeolog expression biases remain roughly consistent between conditions, (61.8% which retained the same pattern of bias).

However, there were a number of homeologs that did exhibit a pathogen-mediated modulation of expression categories. Unsurprisingly, these were largely related to parental expression patterns; for example, a shift from equivalent expression levels to ancestral expression patterns favoring the C homeolog and B. oleracea. I also found that non-parental expression level changes were particularly common among homeologs that retained patterns of bias. While this was evident from C biased homeologs changing from ancestral expression to a downregulaton of the A homeolog in the infected samples, non-parental expression patterns were particularly interesting among A subgenome biased homeologs. These were typified by a shift from transgressive down-regulation of the C homeolog to reciprocal dominance of B. oleracea. In other words, it is possible that A dominance manifests due to trans effects from the *B. oleracea* subgenome. A similar phenomenon was observed among non-biased homeologs: a sizeable proportion of non-biased homeologs fluctuated from equivalent expression to a regulatory dominance in favour of the C subgenome. Again, this suggests that pathogen treatment might result in a more pronounced C subgenome bias in the expression of transcription regulators.

Previous studies have demonstrated large-scale changes in AS patterns due to biotic stress [45, 50] and as a consequence of polyploidization [14, 48]. Accordingly, I also sought to investigate how polyploid AS patterns are modulated by biotic stress. Overall, disparity in subgenome AS identity was the most prominent consequence of biotic stress on the polyploid AS landscape in both the resynthesized and natural *B. napus*. There were also few homeologous events that were symmetrically expressed across conditions (39.8-40.6% of homeologous AS events in both species). Additionally, among the events shared between conditions, the vast majority of homeologous AS events retained the same patterns of splicing level bias between conditions (75.0% and 78.1% in the natural and resynthesized *B. napus* respectively), which suggested that homeologous AS patterns are more influenced by a change in AS identity than they are by splicing levels.

For the homeologous AS events conserved between mock and infected conditions, I nevertheless performed comparisons between homeologs and their parental transcriptomes. Like analyses of gene expression, patterns of AS level bias were largely consistent between conditions (74.0%). While the splicing levels of most of these events (58.3%) were similarly categorized, the remaining events exhibited shifts between equivalent splicing levels and non-parental splicing patterns — such as regulatory dominance or transgressive up-regulation. This increase in frequency of non-additive splicing levels further supports the evidence that pathogen stress elicits widespread non-parental AS patterns in the newly formed polyploid.

2.3.7 The influence of subgenome expression biases on the resynthesized *Brassica napus* - *Sclerotinia sclerotiorum* pathosystem

Duplicated regions of homeologous chromosomes have been shown to contribute to the overall response of *B. napus* to *S. sclerotiorum* [67, 114]. Consequently, the response of a polyploid to pathogen stress can be modulated by large-scale genomic changes such as HEs, or through biases in the global levels of subgenome expression. Previous work in hexaploid wheat infected with Fusarium pseudograminearum identified strong bias towards the B and D subgenomes in both expression and the degree of responsiveness to the pathogen [32]. They referred to this biased responsiveness as homeolog induction bias. I performed a similar analysis in the resynthesized tetraploid *B. napus* which showed a slight bias towards the C subgenome (244 > 122). Additionally, despite consistent rates of subgenome induction between both the A and C subgenomes, a significant majority of up-regulated events exhibit C subgenome bias in expression. This asymmetric preference for C subgenome expression under pathogen infection was corroborated by the clustering of host response ($\log_2 FC$) to S. sclerotiorum (Figure 2.7). Altogether, these analyses suggest that the defense response of the newly formed polyploid is disproportionately influenced by the B. oleracea-derived subgenome.

Considering the widespread bias towards the C subgenome among highly induced genes, I also assessed genes previously implicated in *S. sclerotiorum* infection [63, 89]. Although a number of genes exhibited an A subgenome bias, there was a significantly higher rate of C subgenome bias among the assessed defense response genes (illustrated in Figure 2.13A). For example, looking at the ET biosynthesis pathway, in addition to several genes regulated by JA/ET activity, the majority of the highly induced genes exhibited significant bias towards the C homeolog (Figure 2.13D). Interestingly, many of the genes exhibiting extensive C subgenome bias originated from the same *B. rapa* chromosome. For example, both MTO3/MAT4 gene sets and PR4 (Figure 2.13D), which exhibit substantial down-regulation of the A homeolog, are from chromosome A03. Perhaps large-scale chromosomal events such as homeologous exchange are responsible for this coordinated down-regulation of the A homeolog. Recent work has shown a number of quantitative trait loci (QTL) associated with *S. sclerotiorum* resistance in the *B. rapa*-derived chromosome A3 of *B. napus* [115], which is slightly more gene dense than other *B. rapa* chromosomes [75], suggesting genomic events in this region could result in an apparent coordinated down-regulation of the A homeolog among defense genes. Genomic assessment of the resynthesized *B. napus* in conjunction with transcriptome analyses would provide better support of the role these changes might have on homeolog expression. Nevertheless, it is clear that the emergence of subgenome biases occurs shortly after polyploidization and has a duly strong influence on both the identity and extent of defense gene induction.

2.4 Conclusions

The results of this research demonstrated considerable homeolog expression bias among natural and resynthesized *Brassica napus*, which occurs broadly across both subgenomes, culminating in a clear C subgenome bias. Through parental-polyploid comparisons facilitated by the direct parental relationship between resynthesized *B. napus* and the assayed diploid *Brassica* species, *Brassica rapa* and *Brassica oleracea*, I also determined that biased expression in the newly formed polyploid largely reflected ancestral expression patterns. However, there is a pattern of transgressive down-regulation of the A homeolog, highlighting that considerable transcriptomic changes manifest only a short time after polyploidization.

This study also assessed subgenome AS in each *Brassica napus* genotype. Of potential AS level biases, I found the most notable evidence of differing patterns between the overall expression of AS events, which was more frequent in the C subgenome, and unequal splicing levels in the resynthesized *B. napus*, which favoured the A subgenome. Particularly interesting were the emergent effects that polyploidization had on AS event identity in the resynthesized polyploid. While there was considerable AS event overlap, polyploid subgenomes demonstrated both non-parental AS events and AS event expression likely originating from the

alternate parent — highlighting the possible influence of regulatory mismatches.

While I found that the pathogen stress elicited by S. sclerotiorum had little influence on the patterns of homeolog expression bias exhibited by natural *B. napus*, pathogen infection had a profound influence on biased homeolog expression in newly formed *B. napus*. To further investigate these shifting patterns of expression bias, I also assessed pathogen-mediated changes in expression levels of the resynthesized polyploid relative to its diploid parents, showing a distinctly higher downregulation of the A subgenome — i.e. the expression-recessive subgenome under pathogen stress. Interestingly, these parent-polyploid analyses also showed extensive non-parental AS event expression, in addition to an increase in transgressive AS event up-regulation in response to biotic stress, further highlighting the dramatic reaction stress has upon the transcriptional interplay between subgenomes. I also demonstrated that these expression biases have a strong effect on the defense response of the newly formed polyploid, as genes implicated in pathogen defense displayed consistent homeolog expression biased towards the C subgenome. Additionally, C homeologs also exhibited significantly higher responsiveness to the pathogen stress.

Altogether, this study emphasizes the influence subgenome expression dominance has on multiple layers of polyploid transcriptomes. This can be seen in the divergence of subgenome transcript dosages, emergent AS patterns, and the orchestration of a stress response. These findings provide insight into the complex aspects of inter-subgenomic expression states and will hopefully serve as a foundation for future molecular analyses.

2.5 Tables and Figures



Figure 2.1: Hierarchical clustering results of RNA sequencing read counts for each *Brassica* sample. Species are represented by the following sample IDs: BR (*B. rapa*), BO (*B. oleracea*), BN (*B. napus*), BS (resynthesized *B. napus*). Suffixes denote mock-inoculation (M) and pathogen-inoculation (I). (A) Gene-wise FPKM values for infected and control samples of each host clustered separately using Wards method. Both coloured branches and shaded dendrogram backgrounds correspond to k-means clustering results. (B) Hierarchical clustering of homeolog and ortholog FPKM values using the aforementioned technique for all samples inclusive. Subgenomes of the polyploid are indicated in brackets such that AN and CN represent the A and C subgenomes, respectively.



Figure 2.2: Differential gene expression analyses of each *Brassica* host. Volcano plots for each host depict the number of differentially expressed genes under an FDR < 0.001 and an absolute $\log_2 FC > 3$.



Figure 2.3: Congruence in the gene ontology (GO) enrichment analyses of differentially expressed genes for each host. (A) Heatmap of significantly enriched GO terms (hypergeometric P < 0.01) among up-regulated genes that are co-enriched with other *Brassica* hosts. (B) Significantly enriched GO terms of down-regulated genes, highlighting co-enrichment with other *Brassica* species.



PRRs, MAPK modules, and WRKY transcription factors

Figure 2.4: Mean response $(\log_2 FC)$ of hormone biosynthesis pathways, hormone signalling pathways, and gene families commonly implicated in *Brassica napus* stress response for each *Brassica* host species. Colours are represented by a heatmap depicting up-regulated (red) and down-regulated (down) average responses.

B. rapa



Figure 2.5: GO enrichment analyses of differentially alternatively spliced (DAS) genes for each *Brassica* host compared to enriched GO terms of differentially expressed genes. (A) Significantly enriched GO terms of DAS genes (hypergeometric P < 0.01). Cell color corresponds to gene set (red represents up-regulated AS events; blue represents down-regulated AS events). (B) GO terms exhibiting significant enrichment in at least one host. Color corresponds to gene set of origin: red represents up-regulated genes, blue depicts down-regulated genes.



Figure 2.6: Alternative splicing and gene expression exhibit opposing reactions to pathogen stress. (A) Pearson correlation between change in gene expression (\log_2 FC FPKM) and change in alternative splicing (\log_2 FC PSI) across the entire transcriptome of each species. (B) Heatmap of hormone biosynthesis and signalling pathways that are implicated in stress response. Colors correspond to the average response (\log_2 FC) of (right) gene expression and (left) alternative splicing.



Figure 2.7: Hierarchical clustering of host pathogen stress response. (A) Hosts were clustered based on the response to pathogen stress (\log_2 FC of FPKM normalized reads) using orthologous and homeologous gene relationships. These clusters were further supported by k-means clustered prinicipal components analyses (PCA) of the same data. Color is indicative of k-means identified clusters. (B) Hosts were similarly clustered using a homeolog-consolidated polyploid transcriptome to illustrate changes in homeolog dosage. (C) Hosts clustered (using hierarchical clustering and PCA) based on the change in alternative splicing ratios (\log_2 FC of PSI values).



Figure 2.8: Subgenome expression biases under uninfected and *S. sclerotiorum*-infected conditions. (A) Subgenome expression level biases in the natural and resynthesized for *B. napus* under uninfected and infected expression states, respectively. The number of homeologous gene pairs that exhibited significant differences in expression level were assessed with a χ^2 goodness of fit test to identify subgenome-level expression biases. (B) Biases in the proportion of homeologous gene pairs expressed (FPKM > 1). Significant differences in proportion were evaluated with the Wilcoxon rank sum test. **P* < 0.05.



Figure 2.9: Subgenome alternative splicing biases before and after pathogen stress. (A) Subgenome bias in alternative splicing ratios within the natural and resynthesized *B. napus*, respectively. Subgenome-wide biases in alternative splicing levels were determined based on the number of biased homeologous AS events. The significance of these putative biases were reported based χ^2 goodness of fit test (B) Bias in the frequency of alternative splicing event expression across homeologous splicing junctions for the following event classes: intron retention (IR), alternative acceptor (ALTA), alternative donor (ALTD), and alternative position (ALTP). Significant differences in the proportion of spliced introns for each given event were assessed with the Wilcoxon rank sum test (**P* < 0.05).


Figure 2.10: Transcriptome-wide change in subgenome bias between uninfected and infected expression states. (A) Alluvial plot highlighting the change in subgenome expression bias between uninfected and *S. sclerotiorum*-inoculated conditions for both the natural and the resynthesized *B. napus*. Blue alluvial flow are genes that retained the same bias between conditions, whereas red depicts genes in which there has been a change in subgenome bias (B) Similarly, alluvial plots for the two *B. napus* hosts describing the influence of *S. sclerotiorum* on biases in alternative splicing.

	Addi	tivity	A–exp level do	ression minance	C-exp level do	ression minance	T do	ransgressive Transgressive wn regulation up regulation			No Change			
Inoculation	ASC	ASC	ASC	ASC	ASC	ASC	ASC	ASC	ASC	ASC	ASC	ASC	ASC	Totals
method	\searrow	1	-	1	\int	7	\mathbf{V}	V	$\overline{\mathbf{V}}$	$\langle \langle \rangle$	\wedge	[]	•-•-•	
	Gene expression													
Mock	539	631	525	1898	1173	4145	673	400	1137	15	38	23	14817	26014
Pathogen	412	600	577	1631	1502	4101	455	301	715	3	43	12	15649	26001
	Alternative splicing													
Mock	109	119	614	212	1183	246	1	13	4	12	173	24	5431	8141
Pathogen	99	247	312	380	802	321	5	18	2	5	63	5	4959	7218

Figure 2.11: Parental expression and alternative splicing levels compared to a homeolog-consolidated polyploid expression state. Patterns in gene expression and alternative splicing were arranged into 13 categories based on comparisons between (A) *B. rapa*, (S) *B. napus* (syn.), and (C) *B. oleracea*.

)	В					Genome A At C Ct	•	c	Ċ)	D	C	
		AI	bias			CI	oias				No bias			
Inoculation	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	AR AN CN CO	Totals
method	-	Λ.	\mathbb{V}	\mathcal{N}	\int	\mathbf{N}	\bigvee	М				1	••••	
						Ge	ne express	ion						
Mock	1266 (5%)	318 (1.3%)	512 (2%)	433 (1.7%)	4022 (16%)	391 (1.6%)	1379 (5.5%)	379 (1.5%)	632 (2.5%)	432 (1.7%)	1333 (5.3%)	698 (2.8%)	11663 (46.3%)	25180
Pathogen	1512 (6.1%)	336 (1.4%)	368 (1.5%)	439 (1.8%)	4731 (19%)	356 (1.4%)	2045 (8.2%)	412 (1.7%)	365 (1.5%)	388 (1.6%)	1006 (4%)	801 (3.2%)	10746 (43.2%)	24878
Alternative splicing														
Mock	674 (11.5%)	508 (8.7%)	48 (0.8%)	90 (1.5%)	778 (13.3%)	222 (3.8%)	48 (0.8%)	28 (0.5%)	161 (2.7%)	97 (1.7%)	322 (5.5%)	38 (0.6%)	2191 (37.4%)	5865
Pathogen	480 (10.4%)	295 (6.4%)	85 (1.8%)	142 (3.1%)	583 (12.6%)	104 (2.3%)	63 (1.4%)	36 (0.8%)	60 (1.3%)	185 (4%)	211 (4.6%)	101 (2.2%)	1829 (39.7%)	4612

Figure 2.12: Homeolog expression and alternative splicing patterns relative to parental expression states. (A-D) Overlap between polyploid subgenomes and their respective parental transcriptome with regards to the identity of expressed genes (A-B) and detected alternative splicing events (C-D) for both mock and pathogen conditions. Colours correspond to subgenome or parental genome: At (resynthesized B. napus A subgenome), Ct (resynthesized B. napus C subgenome), A (B. rapa), C (B. oleracea). (A-B) Gene expression profile of polyploid subgenomes (interior ring) compared to parental expression profiles (exterior ring) under uninfected (A) and infected (B) conditions. (C-D) Alternative splicing profiles of polyploid subgenomes compared to diploid progenitor AS profiles under (C) uninfected and (D) infected conditions. (E) The data presented reflects the most common expression-level and splicing level categories for AR - B. rapa, AN - the A subgenome of the resynthesized B. napus, CN - the C subgenome of the resynthesized B. napus, and CO - B. oleracea. Expression categories are organized based on contribution to homeolog expression or homeolog splicing biases, with each column summarizing the frequency of each category.



Figure 2.13: Influence of subgenome expression bias on the defense response of the resynthesized *Brassica napus*. (A) Average homeolog expression bias of gene sets commonly implicated in pathogen defense to *Sclerotinia sclerotiorum*, including: salyclic acid (SA), jasmonic acid (JA), ethylene (ET), WRKY transcription factors, wall-associated protein kinases (Wall), PEP-recptors (PEP), cysteine-rich receptor kinases, and chitin receptor kinases. (B) An assessment of homeolog expression bias among genes previous implicated in response to pathogen stress. (C) Enzymes involved in the ET biosynthesis pathway. Arrows are coloured by subgenome bias of genes involved in the pathway (D) The impact of homeolog expression on the ET biosynthesis pathway represented as FPKG.

Species	Condition	Million	Million	Percent	Percent
		cleaned	mapped	Mapped	Mapped
		reads	reads	(Bn)	(Ss)
		(Bn)	(Bn)		
B. napus	Infected	132.4	92.20	69.60	15.90
(synthetic)		116.2	81.30	70.00	15.30
		116.6	87.50	75.10	10.60
B. napus		125.3	88.60	70.70	12.90
		115	64.70	56.30	24.60
		132.3	78.50	59.30	23.00
B. oleracea		67.8	43.60	64.30	21.40
		92.7	66.60	71.90	16.30
		67.2	51.40	76.50	24.30
B. rapa		89.5	73.90	82.60	6.60
		68.1	55.10	81.00	8.00
		84.3	59.50	70.50	16.20
B. napus	Mock	136.2	121.40	89.10	0.00
(synthetic)		126.5	111.30	88.00	0.00
		127.6	112.30	88.00	0.00
B. napus		114.1	97.70	85.60	0.00
		112.2	95.70	85.30	0.00
		211.1	178.20	84.40	0.00
B. oleracea		56.3	51.70	91.80	0.00
		56	52.20	93.10	0.00
		100*	92.20	92.20	0.00
B. rapa		67.4	60.90	90.30	0.00
		62.7	56.20	89.60	0.00
		62	56.00	90.40	0.00

Table 2.1: Read mapping summary of Brassica species and S. sclerotiorum

* Reads from this B. oleracea sample were randomly downsampled to 100000000 due to file size

Species	Condition	Number	Fraction	FPKM	FPKM	FPKM
		of genes	of	1-5	5-50	>50
		expressed	genome			
B. napus	Infected	41188	40.80	18867	19444	2877
		41083	40.70	18370	19746	2967
		42775	42.30	19402	20536	2837
B. napus (syn.)		39455	39.10	16468	19209	3778
		38990	38.60	16079	19061	3850
		38250	37.90	15770	18668	3812
B. oleracea		23190	38.30	7427	11877	3886
		24004	39.60	7332	12615	4057
		22465	37.10	7444	11247	3774
B. rapa		22526	48.70	6261	12491	3774
		21829	47.20	6547	11661	3621
		21816	47.20	6330	11722	3764
B. napus	Mock	45103	44.60	20784	22085	2234
		44206	43.80	20384	21657	2165
		44316	43.90	20526	21621	2169
B. napus (syn.)		42429	42.00	16964	22181	3284
		43511	43.10	17532	22635	3344
		42260	41.80	16736	22106	3418
B. oleracea		25734	42.50	7091	14904	3739
		25460	42.00	6947	14697	3816
		25529	42.10	7138	14765	3626
B. rapa		22409	48.50	5931	12824	3654
		22779	49.30	6126	12808	3845
		22633	48.90	6024	12885	3724

 Table 2.2: Summary of sample FPKM normalized expression data

Species	Events	Mock	Mock	Infected	Infected
			(%)		(%)
B. napus	IR	65088	65.80	56574	69.70
	ALTA	11971	12.10	8319	10.20
	ALTD	8497	8.60	6405	7.90
	ALTP	4274	4.30	2957	3.60
	SKIP	1181	1.20	1220	1.50
B. napus (syn.)	IR	62309	68.80	45633	66.40
	ALTA	9270	10.20	7068	10.30
	ALTD	6713	7.40	5629	8.20
	ALTP	5608	6.20	4661	6.80
	SKIP	682	0.80	610	0.90
B. oleracea	IR	34918	67.10	30119	68.10
	ALTA	5513	10.60	4434	10.00
	ALTD	4352	8.40	3758	8.50
	ALTP	3163	6.10	2501	5.70
	SKIP	451	0.90	552	1.20
B. rapa	IR	26489	64.90	24978	62.60
	ALTA	4623	11.30	4640	11.60
	ALTD	3390	8.30	3612	9.00
	ALTP	2935	7.20	2962	7.40
	SKIP	249	0.60	236	0.60

Table 2.3: Overview of AS event distributions among *Brassica* species under mock and pathogen inoculations

Table 2.4: Overview of AS event expression among homologous junctions

Genome	Condition	Number of	Number of in-
		AS events	trons
B. napus A	Mock	17880	15166
B. napus C		20313	17089
B. napus syn. A		16084	14017
B. napus syn. C		19857	17541
B. rapa		17077	14411
B. oleracea		19573	17201
B. napus A	Infected	15242	13228
B. napus C		17120	14710
B. napus syn. A		11551	9893
B. napus syn. C		14820	12972
B. rapa		16272	13439
B. oleracea		17057	14837

Bibliography

- [1] Sarah P Otto and Jeannette Whitton. Polyploid incidence and evolution. Annual review of genetics, 34(1):401–437, 2000. \rightarrow page 1
- [2] Barbara K Mable. Breaking down taxonomic barriers in polyploidy research. Trends in plant Science, 8(12):582–590, 2003.
- [3] BK Mable, MA Alexandrou, and MI Taylor. Genome duplication in amphibians and fish: an extended synthesis. <u>Journal of Zoology</u>, 284(3): 151–182, 2011.
- [4] Kevin Vanneste, Guy Baele, Steven Maere, and Yves Van de Peer. Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the cretaceous–paleogene boundary. <u>Genome research</u>, 24 (8):1334–1347, 2014. → page 1
- [5] Zheng Li, Anthony E Baniaga, Emily B Sessa, Moira Scascitelli, Sean W Graham, Loren H Rieseberg, and Michael S Barker. Early genome duplications in conifers and other seed plants. <u>Science advances</u>, 1(10): e1501084, 2015. → page 1
- [6] Keith L Adams and Jonathan F Wendel. Polyploidy and genome evolution in plants. Current opinion in plant biology, 8(2):135-141, 2005. \rightarrow page 1
- [7] Pamela S Soltis and Douglas E Soltis. Ancient wgd events as drivers of key innovations in angiosperms. <u>Current opinion in plant biology</u>, 30:159–165, 2016. → page 1
- [8] One Thousand Plant Transcriptomes Initiative et al. One thousand plant transcriptomes and the phylogenomics of green plants. <u>Nature</u>, 574(7780): 679, 2019. → page 1
- [9] Yuannian Jiao, Norman J Wickett, Saravanaraj Ayyampalayam, André S Chanderbali, Lena Landherr, Paula E Ralph, Lynn P Tomsho, Yi Hu,

Haiying Liang, Pamela S Soltis, et al. Ancestral polyploidy in seed plants and angiosperms. Nature, 473(7345):97, 2011. \rightarrow page 1

- [10] Justin Ramsey. Polyploidy and ecological adaptation in wild yarrow. <u>Proceedings of the National Academy of Sciences</u>, 108(17):7096–7101, 2011. \rightarrow page 1
- [11] Ayelet Salman-Minkov, Niv Sabath, and Itay Mayrose. Whole-genome duplication as a key factor in crop domestication. <u>Nature plants</u>, 2(8): 16115, 2016. → page 1
- [12] Mi-Jeong Yoo, Xiaoxian Liu, J Chris Pires, Pamela S Soltis, and Douglas E Soltis. Nonadditive gene expression in polyploids. <u>Annual review of</u> genetics, 48:485–517, 2014. → page 2
- [13] Guanjing Hu, Jin Koh, Mi-Jeong Yoo, Sixue Chen, and Jonathan F Wendel. Gene-expression novelty in allopolyploid cotton: a proteomic perspective. Genetics, 200(1):91–104, 2015. → page 2
- [14] Renchao Zhou, Noushin Moshgabadi, and Keith L Adams. Extensive changes to alternative splicing patterns following allopolyploidy in natural and resynthesized polyploids. <u>Proceedings of the National Academy of</u> Sciences, 108(38):16122–16127, 2011. → pages 2, 4, 5, 45
- [15] Mingquan Ding and Z Jeffrey Chen. Epigenetic perspectives on the evolution and domestication of polyploid plant and crops. <u>Current opinion</u> in plant biology, 42:37–48, 2018. → pages 2, 23
- [16] Jonathan F Wendel, Damon Lisch, Guanjing Hu, and Annaliese S Mason. The long and short of doubling down: polyploidy, epigenetics, and the temporal dynamics of genome fractionation. <u>Current opinion in genetics &</u> development, 49:1–7, 2018. → pages 2, 23
- [17] Mariska Te Beest, Johannes J Le Roux, David M Richardson, Anne K Brysting, Jan Suda, Magdalena Kubešová, and Petr Pyšek. The more the better? the role of polyploidy in facilitating plant invasions. <u>Annals of</u> botany, 109(1):19–45, 2011. → page 2
- [18] Robert T Gaeta, J Chris Pires, Federico Iniguez-Luy, Enrique Leon, and Thomas C Osborn. Genomic changes in resynthesized brassica napus and their effect on gene expression and phenotype. <u>The Plant Cell</u>, 19(11): 3403-3417, 2007. \rightarrow pages 2, 5, 39

- [19] CE Grover, JP Gallagher, EP Szadkowski, MJ Yoo, LE Flagel, and JF Wendel. Homoeolog expression bias and expression level dominance in allopolyploids. New Phytologist, 196(4):966–971, 2012. → page 2
- [20] Lex Flagel, Joshua Udall, Dan Nettleton, and Jonathan Wendel. Duplicate gene expression in allopolyploid gossypium reveals two temporally distinct phases of expression evolution. BMC biology, 6(1):16, 2008. → page 2
- [21] Richard JA Buggs, Linjing Zhang, Nicholas Miles, Jennifer A Tate, Lu Gao, Wu Wei, Patrick S Schnable, W Brad Barbazuk, Pamela S Soltis, and Douglas E Soltis. Transcriptomic shock generates evolutionary novelty in a newly formed, natural allopolyploid plant. <u>Current Biology</u>, 21(7): 551–556, 2011.
- [22] Z Jeffrey Chen. Genomic and epigenetic insights into the molecular bases of heterosis. Nature Reviews Genetics, 14(7):471–482, 2013. \rightarrow page 2
- [23] Zhenlan Liu and Keith L Adams. Expression partitioning between genes duplicated by polyploidy under abiotic stress and during organ development. Current Biology, 17(19):1669–1674, 2007. → pages 2, 3
- [24] Shaowei Dong and Keith L Adams. Differential contributions to the transcriptome of duplicated genes in response to abiotic stresses in natural and synthetic polyploids. <u>New Phytologist</u>, 190(4):1045–1057, 2011. → pages 2, 3
- [25] Ryan A Rapp, Joshua A Udall, and Jonathan F Wendel. Genomic expression dominance in allopolyploids. <u>BMC biology</u>, 7(1):18, 2009. → pages 2, 3, 7, 13, 26
- [26] MJ Yoo, Emmanuel Szadkowski, and JF Wendel. Homoeolog expression bias and expression level dominance in allopolyploid cotton. <u>Heredity</u>, 110 (2):171, 2013. → pages 3, 7, 26, 39, 40
- [27] Huakun Zhang, Xiaowan Gou, Ai Zhang, Xutong Wang, Na Zhao, Yuzhu Dong, Linfeng Li, and Bao Liu. Transcriptome shock invokes disruption of parental expression-conserved genes in tetraploid wheat. <u>Scientific reports</u>, 6:26363, 2016.
- [28] Patrick P Edger, Ronald Smith, Michael R McKain, Arielle M Cooley, Mario Vallejo-Marin, Yaowu Yuan, Adam J Bewick, Lexiang Ji, Adrian E Platts, Megan J Bowman, et al. Subgenome dominance in an interspecific hybrid, synthetic allopolyploid, and a 140-year-old naturally established

neo-allopolyploid monkeyflower. The Plant Cell, 29(9):2150–2167, 2017. \rightarrow pages 3, 7, 39

- [29] Yves Van de Peer, Eshchar Mizrachi, and Kathleen Marchal. The evolutionary significance of polyploidy. <u>Nature Reviews Genetics</u>, 18(7): 411, 2017.
- [30] Jian Wu, Li Lin, Meiling Xu, Peipei Chen, Dongxiao Liu, Qinfu Sun, Liping Ran, and Youping Wang. Homoeolog expression bias and expression level dominance in resynthesized allopolyploid brassica napus. BMC genomics, 19(1):586, 2018. → pages 3, 5, 7, 26, 39, 40
- [31] Aili Li, Dengcai Liu, Jun Wu, Xubo Zhao, Ming Hao, Shuaifeng Geng, Jun Yan, Xiaoxue Jiang, Lianquan Zhang, Junyan Wu, et al. mrna and small rna transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. <u>The Plant Cell</u>, 26(5): 1878–1900, 2014. → page 3
- [32] Jonathan J Powell, Timothy L Fitzgerald, Jiri Stiller, Paul J Berkman, Donald M Gardiner, John M Manners, Robert J Henry, and Kemal Kazan. The defence-associated transcriptome of hexaploid wheat displays homoeolog expression and induction bias. Plant biotechnology journal, 15 (4):533–543, 2017. → pages 3, 6, 7, 33, 40, 44, 46
- [33] Saurabh Chaudhary, Waqas Khokhar, Ibtissam Jabre, Anireddy SN Reddy, Lee J Byrne, Cornelia M Wilson, and Naeem H Syed. Alternative splicing and protein diversity: plants versus animals. <u>Frontiers in Plant Science</u>, 10, 2019. → page 4
- [34] Sergei A Filichkin, Henry D Priest, Scott A Givan, Rongkun Shen, Douglas W Bryant, Samuel E Fox, Weng-Keen Wong, and Todd C Mockler. Genome-wide mapping of alternative splicing in arabidopsis thaliana. Genome research, 20(1):45–58, 2010. → page 4
- [35] Maria Kalyna, Craig G Simpson, Naeem H Syed, Dominika Lewandowska, Yamile Marquez, Branislav Kusenda, Jacqueline Marshall, John Fuller, Linda Cardle, Jim McNicol, et al. Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in arabidopsis. <u>Nucleic acids research</u>, 40(6):2454–2469, 2011. → page 4
- [36] Gabriele Drechsel, André Kahles, Anil K Kesarwani, Eva Stauffer, Jonas Behr, Philipp Drewe, Gunnar Rätsch, and Andreas Wachter.

Nonsense-mediated decay of alternative precursor mrna splicing variants is a major determinant of the arabidopsis steady state transcriptome. The Plant Cell, 25(10):3726–3742, 2013. \rightarrow pages 4, 38

- [37] Sergei A Filichkin, Jason S Cumbie, Palitha Dharmawardhana, Pankaj Jaiswal, Jeff H Chang, Saiprasad G Palusa, ASN Reddy, Molly Megraw, and Todd C Mockler. Environmental stresses modulate abundance and timing of alternatively spliced circadian transcripts in arabidopsis. Molecular plant, 8(2):207–227, 2015. → page 4
- [38] Sridevi Sureshkumar, Craig Dent, Andrei Seleznev, Celine Tasset, and Sureshkumar Balasubramanian. Nonsense-mediated mrna decay modulates flm-dependent thermosensory flowering response in arabidopsis. <u>Nature</u> Plants, 2(5):16055, 2016. → pages 4, 38
- [39] Hee-Jeong Jeong, Young Jin Kim, Sang Hyon Kim, Yoon-Ha Kim, In-Jung Lee, Yoon Ki Kim, and Jeong Sheop Shin. Nonsense-mediated mrna decay factors, upf1 and upf3, contribute to plant defense. <u>Plant and cell</u> physiology, 52(12):2147–2156, 2011. → page 4
- [40] Sergei A Filichkin, Michael Hamilton, Palitha D Dharmawardhana, Sunil K Singh, Christopher Sullivan, Asa Ben-Hur, Anireddy SN Reddy, and Pankaj Jaiswal. Abiotic stresses modulate landscape of poplar transcriptome via alternative splicing, differential intron retention, and isoform ratio switching. Frontiers in plant science, 9:5, 2018. → page 4
- [41] Dorothee Staiger, Christin Korneli, Martina Lummer, and Lionel Navarro. Emerging role for rna-based regulation in plant immunity. <u>New</u> <u>Phytologist</u>, 197(2):394–404, 2013. → page 4
- [42] Jiradet Gloggnitzer, Svetlana Akimcheva, Arunkumar Srinivasan, Branislav Kusenda, Nina Riehs, Hansjörg Stampfl, Jaqueline Bautor, Bettina Dekrout, Claudia Jonak, José M Jiménez-Gómez, et al. Nonsense-mediated mrna decay modulates immune receptor levels to regulate plant antibacterial defense. Cell host & microbe, 16(3):376–390, 2014. → page 4
- [43] Dorothee Staiger and John WS Brown. Alternative splicing at the intersection of biological timing, development, and stress responses. The Plant Cell, 25(10):3640–3656, 2013. \rightarrow page 4
- [44] Tom Laloum, Guiomar Martín, and Paula Duque. Alternative splicing control of abiotic stress responses. <u>Trends in plant science</u>, 23(2):140–150, 2018. → page 4

- [45] Kranthi K Mandadi and Karen-Beth G Scholthof. Genome-wide analysis of alternative splicing landscapes modulated during plant-virus interactions in brachypodium distachyon. <u>The Plant Cell</u>, 27(1):71–85, 2015. → pages 4, 18, 45
- [46] Fu-Yuan Zhu, Mo-Xian Chen, Neng-Hui Ye, Lu Shi, Kai-Long Ma, Jing-Fang Yang, Yun-Ying Cao, Youjun Zhang, Takuya Yoshida, Alisdair R Fernie, et al. Proteogenomic analysis reveals alternative splicing and translation as part of the abscisic acid response in arabidopsis seedlings. The Plant Journal, 91(3):518–533, 2017. → page 4
- [47] Naeem H Syed, Maria Kalyna, Yamile Marquez, Andrea Barta, and John WS Brown. Alternative splicing in plants-coming of age. <u>Trends in</u> plant science, 17(10):616–623, 2012. → page 4
- [48] Maojun Wang, Pengcheng Wang, Fan Liang, Zhengxiu Ye, Jianying Li, Chao Shen, Liuling Pei, Feng Wang, Jiang Hu, Lili Tu, et al. A global survey of alternative splicing in allopolyploid cotton: landscape, complexity and regulation. <u>New Phytologist</u>, 217(1):163–178, 2018. → pages 4, 7, 41, 42, 45
- [49] Zhibin Zhang, Tiansi Fu, Zhijian Liu, Xutong Wang, Hongwei Xun, Guo Li, Baoxu Ding, Yuzhu Dong, Xiuyun Lin, Karen A Sanguinet, et al. Extensive changes in gene expression and alternative splicing due to homoeologous exchange in rice segmental allopolyploids. <u>Theoretical and</u> Applied Genetics, pages 1–14, 2019. → page 4
- [50] Hong Zhang, Rui Mao, Yanzhen Wang, Lu Zhang, Changyou Wang, Shikai Lv, Xinlun Liu, Yajuan Wang, and Wanquan Ji. Transcriptome-wide alternative splicing modulation during plant-pathogen interactions in wheat. Plant Science, 2019. → pages 5, 45
- [51] Zhenshan Liu, Jinxia Qin, Xuejun Tian, Shengbao Xu, Yu Wang, Hongxia Li, Xiaoming Wang, Huiru Peng, Yingyin Yao, Zhaorong Hu, et al. Global profiling of alternative splicing landscape responsive to drought, heat and their combination in wheat (triticum aestivum l.). <u>Plant biotechnology</u> journal, 16(3):714–726, 2018. → page 5
- [52] U Nagaharu. Genome analysis in brassica with special reference to the experimental formation of b. napus and peculiar mode of fertilization. Jpn J Bot, 7(7):389–452, 1935. → page 5

- [53] Boulos Chalhoub, France Denoeud, Shengyi Liu, Isobel AP Parkin, Haibao Tang, Xiyin Wang, Julien Chiquet, Harry Belcram, Chaobo Tong, Birgit Samans, et al. Early allopolyploid evolution in the post-neolithic brassica napus oilseed genome. <u>science</u>, 345(6199):950–953, 2014. → pages 5, 10, 11, 12, 18, 39, 40
- [54] Keith L Adams, Richard Cronn, Ryan Percifield, and Jonathan F Wendel. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. <u>Proceedings of the</u> National Academy of sciences, 100(8):4649–4654, 2003. → page 5
- [55] Chaobo Tong, Rafaqat Ali Gill, Yang Xiang, Lixin Ma, Xiaohui Cheng, Junyan Huang, and Shengyi Liu. Fractionization of polyploid duplicated genes: Gene loss, expression divergence, and epigenetic regulation in brassica napus. In <u>The Brassica napus Genome</u>, pages 149–158. Springer, 2018.
- [56] Bhavna Hurgobin, Agnieszka A Golicz, Philipp E Bayer, Chon-Kit Kenneth Chan, Soodeh Tirnaz, Aria Dolatabadian, Sarah V Schiessl, Birgit Samans, Juan D Montenegro, Isobel AP Parkin, et al. Homoeologous exchange is a major cause of gene presence/absence variation in the amphidiploid brassica napus. <u>Plant biotechnology journal</u>, 16(7):1265–1274, 2018. → pages 5, 6, 10
- [57] Paulina Martinez Palacios, Marie-Pierre Jacquemot, Marion Tapie, Agnès Rousselet, Mamoudou Diop, Carine Remoué, Matthieu Falque, Andrew Lloyd, Eric Jenczewski, Gilles Lassalle, et al. Assessing the response of small rna populations to allopolyploidy using resynthesized brassica napus allotetraploids. <u>Molecular biology and evolution</u>, 36(4):709–726, 2019. → page 5
- [58] Benjamin P Oswald and Scott L Nuismer. Neopolyploidy and pathogen resistance. Proceedings of the Royal Society B: Biological Sciences, 274 (1624):2393–2397, 2007. → page 6
- [59] Roger W Innes, Carine Ameline-Torregrosa, Tom Ashfield, Ethalinda Cannon, Steven B Cannon, Ben Chacko, Nicolas WG Chen, Arnaud Couloux, Anita Dalwani, Roxanne Denny, et al. Differential accumulation of retroelements and diversification of nb-lrr disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. <u>Plant</u> <u>Physiology</u>, 148(4):1740–1759, 2008. → page 6

- [60] Jayami K Abeysinghe, Kai-Man Lam, and Danny W-K Ng. Differential regulation and interaction of homoeologous wrky 18 and wrky 40 in arabidopsis allotetraploids and biotic stress responses. <u>The Plant Journal</u>, 97(2):352–367, 2019. → page 6
- [61] Philipp E Bayer, Agnieszka A Golicz, Soodeh Tirnaz, Chon-Kit Kenneth Chan, David Edwards, and Jacqueline Batley. Variation in abundance of predicted resistance genes in the brassica oleracea pangenome. <u>Plant</u> biotechnology journal, 17(4):789–800, 2019. → page 6
- [62] A Taylor, K Rana, C Handy, and JP Clarkson. Resistance to sclerotinia sclerotiorum in wild brassica species and the importance of sclerotinia subarctica as a brassica pathogen. <u>Plant pathology</u>, 67(2):433–444, 2018. → page 6
- [63] Ian J Girard, Chaobo Tong, Michael G Becker, Xingyu Mao, Junyan Huang, Teresa de Kievit, WG Dilantha Fernando, Shengyi Liu, and Mark F Belmonte. Rna sequencing of brassica napus reveals cellular redox control of sclerotinia infection. Journal of Experimental Botany, 68(18): 5079–5091, 2017. → pages 6, 13, 17, 36, 37, 46
- [64] Zheng Wang, Lu-Yue Ma, Jun Cao, Yu-Long Li, Li-Na Ding, Ke-Ming Zhu, Yan-Hua Yang, and Xiao-Li Tan. Recent advances in mechanisms of plant defense to sclerotinia sclerotiorum. <u>Frontiers in Plant Science</u>, 10, 2019. → pages 7, 34
- [65] Jianwei Zhao, Lone Buchwaldt, S Roger Rimmer, Myrtle Brkic, Diana Bekkaoui, and Dwayne Hegedus. Differential expression of duplicated peroxidase genes in the allotetraploid brassica napus. <u>Plant Physiology and</u> Biochemistry, 47(7):653–656, 2009. → page 7
- [66] Jiqiang Li, Zunkang Zhao, Alice Hayward, Hongyu Cheng, and Donghui Fu. Integration analysis of quantitative trait loci for resistance to sclerotinia sclerotiorum in brassica napus. <u>Euphytica</u>, 205(2):483–489, 2015. → page 7
- [67] Sanjaya Gyawali, Myrtle Harrington, Jonathan Durkin, Kyla Horner, Isobel AP Parkin, Dwayne D Hegedus, Diana Bekkaoui, and Lone Buchwaldt. Microsatellite markers used for genome-wide association mapping of partial resistance to sclerotinia sclerotiorum in a world collection of brassica napus. <u>Molecular Breeding</u>, 36(6):72, 2016. → pages 7, 36, 46

- [68] Thomas C Osborn, David V Butrulle, Andrew G Sharpe, Kathryn J Pickering, Isobel AP Parkin, John S Parker, and Derek J Lydiate. Detection and effects of a homeologous reciprocal transposition in brassica napus. <u>Genetics</u>, 165(3):1569–1577, 2003. → page 7
- [69] Jianwei Zhao, Joshua A Udall, Pablo A Quijada, Craig R Grau, Jinling Meng, and Thomas C Osborn. Quantitative trait loci for resistance to sclerotinia sclerotiorum and its association with a homeologous non-reciprocal transposition in brassica napus l. <u>Theoretical and Applied</u> Genetics, 112(3):509–516, 2006. → page 7
- [70] James C Schnable, Nathan M Springer, and Michael Freeling. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. <u>Proceedings of the National Academy of</u> Sciences, 108(10):4069–4074, 2011. → page 7
- [71] Feng Cheng, Jian Wu, Lu Fang, Silong Sun, Bo Liu, Ke Lin, Guusje Bonnema, and Xiaowu Wang. Biased gene fractionation and dominant gene expression among the subgenomes of brassica rapa. <u>PloS one</u>, 7(5): e36442, 2012. → page 7
- [72] Simon Andrews et al. Fastqc: a quality control tool for high throughput sequence data. 2010, 2017. \rightarrow page 10
- [73] Anthony M Bolger, Marc Lohse, and Bjoern Usadel. Trimmomatic: a flexible trimmer for illumina sequence data. <u>Bioinformatics</u>, 30(15): 2114–2120, 2014. \rightarrow page 10
- [74] Alexander Dobin, Carrie A Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R Gingeras. Star: ultrafast universal rna-seq aligner. <u>Bioinformatics</u>, 29(1): 15–21, 2013. → page 10
- [75] Lei Zhang, Xu Cai, Jian Wu, Min Liu, Stefan Grob, Feng Cheng, Jianli Liang, Chengcheng Cai, Zhiyuan Liu, Bo Liu, et al. Improved brassica rapa reference genome by single-molecule sequencing and chromosome conformation capture technologies. <u>Horticulture research</u>, 5(1):50, 2018. → pages 10, 47
- [76] Isobel AP Parkin, Chushin Koh, Haibao Tang, Stephen J Robinson, Sateesh Kagale, Wayne E Clarke, Chris D Town, John Nixon, Vivek Krishnakumar, Shelby L Bidwell, et al. Transcriptome and methylome profiling reveals

relics of genome dominance in the mesopolyploid brassica oleracea. Genome biology, 15(6):R77, 2014. \rightarrow page 10

- [77] Mark Derbyshire, Matthew Denton-Giles, Dwayne Hegedus, Shirin Seifbarghy, Jeffrey Rollins, Jan Van Kan, Michael F Seidl, Luigi Faino, Malick Mbengue, Olivier Navaud, et al. The complete genome sequence of the phytopathogenic fungus sclerotinia sclerotiorum reveals insights into the genome architecture of broad host range pathogens. <u>Genome biology</u> and evolution, 9(3):593–618, 2017. → page 10
- [78] Yang Liao, Gordon K Smyth, and Wei Shi. featurecounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics, 30(7):923–930, 2013. → page 10
- [79] Mark D Robinson, Davis J McCarthy, and Gordon K Smyth. edger: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26(1):139–140, 2010. → page 10
- [80] David C Tack, William R Pitchers, and Keith L Adams. Transcriptome analysis indicates considerable divergence in alternative splicing between duplicated genes in arabidopsis thaliana. <u>Genetics</u>, 198(4):1473–1481, 2014. → pages 10, 11
- [81] Philip Jones, David Binns, Hsin-Yu Chang, Matthew Fraser, Weizhong Li, Craig McAnulla, Hamish McWilliam, John Maslen, Alex Mitchell, Gift Nuka, et al. Interproscan 5: genome-scale protein function classification. Bioinformatics, 30(9):1236–1240, 2014. → page 11
- [82] Fabricio Almeida Araujo, Debmalya Barh, Artur Silva, Luis Guimarães, and Rommel Thiago Juca Ramos. Go feat: a rapid web-based functional annotation tool for genomic and transcriptomic data. <u>Scientific reports</u>, 8 (1):1794, 2018. → page 11
- [83] Petri Törönen, Alan Medlar, and Liisa Holm. Pannzer2: a rapid functional annotation web server. <u>Nucleic acids research</u>, 46(W1):W84–W88, 2018. → page 11
- [84] Guangchuang Yu, Li-Gen Wang, Yanyan Han, and Qing-Yu He. clusterprofiler: an r package for comparing biological themes among gene clusters. <u>Omics: a journal of integrative biology</u>, 16(5):284–287, 2012. → pages 11, 15
- [85] H Tang, V Krishnakuar, and J Li. jcvi: Jcvi utility libraries. Zenodo.(doi: 10.5281/zenodo. 31631), 2015. \rightarrow page 12

- [86] Jian Wu, Qing Zhao, Qingyong Yang, Han Liu, Qingyuan Li, Xinqi Yi, Yan Cheng, Liang Guo, Chuchuan Fan, and Yongming Zhou. Comparative transcriptomic analysis uncovers the complex genetic network for resistance to sclerotinia sclerotiorum in brassica napus. <u>Scientific reports</u>, 6:19007, 2016. → page 13
- [87] Shu-Ye Jiang, Mayalagu Sevugan, and Srinivasan Ramachandran. Valine-glutamine (vq) motif coding genes are ancient and non-plant-specific with comprehensive expression regulation by various biotic and abiotic stresses. BMC genomics, 19(1):342, 2018. → page 17
- [88] Sylvain La Camera, Pierrette Geoffroy, Hala Samaha, Abdoulaye Ndiaye, Gwendoline Rahim, Michel Legrand, and Thierry Heitz. A pathogen-inducible patatin-like lipid acyl hydrolase facilitates fungal and bacterial host colonization in arabidopsis. <u>The Plant Journal</u>, 44(5): 810–825, 2005. → page 17
- [89] Jian Wu, Qing Zhao, Sheng Liu, Muhammad Shahid, Lei Lan, Guangqin Cai, Chunyu Zhang, Chuchuan Fan, Youping Wang, and Yongming Zhou. Genome-wide association study identifies new loci for resistance to sclerotinia stem rot in brassica napus. <u>Frontiers in plant science</u>, 7:1418, 2016. → pages 17, 36, 37, 46
- [90] Hongju Jian, Jinqi Ma, Lijuan Wei, Pu Liu, Aoxiang Zhang, Bo Yang, Jiana Li, Xinfu Xu, and Liezhao Liu. Integrated mrna, srna, and degradome sequencing reveal oilseed rape complex responses to sclerotinia sclerotiorum (lib.) infection. <u>Scientific reports</u>, 8(1):10987, 2018. → page 17
- [91] Tingting Lu, Guojun Lu, Danlin Fan, Chuanrang Zhu, Wei Li, Qiang Zhao, Qi Feng, Yan Zhao, Yunli Guo, Wenjun Li, et al. Function annotation of the rice transcriptome at single-nucleotide resolution by rna-seq. <u>Genome</u> research, 20(9):1238–1249, 2010. → page 18
- [92] Yamile Marquez, John WS Brown, Craig Simpson, Andrea Barta, and Maria Kalyna. Transcriptome survey reveals increased complexity of the alternative splicing landscape in arabidopsis. <u>Genome research</u>, 22(6): 1184–1195, 2012. → page 18
- [93] Mohammad Golam Mostofa, Ajit Ghosh, Zhong-Guang Li, Md Nurealam Siddiqui, Masayuki Fujita, and Lam-Son Phan Tran. Methylglyoxal–a signaling molecule in plant abiotic stress responses. <u>Free Radical Biology</u> and Medicine, 122:96–109, 2018. → page 20

- [94] Satoru Akama, Rie Shimizu-Inatsugi, Kentaro K Shimizu, and Jun Sese. Genome-wide quantification of homeolog expression ratio revealed nonstochastic gene regulation in synthetic allopolyploid arabidopsis. Nucleic acids research, 42(6):e46–e46, 2014. → pages 25, 39, 40
- [95] Tesfaye Mengiste. Plant immunity to necrotrophs. <u>Annual review of</u> phytopathology, 50:267–294, 2012. → pages 36, 37
- [96] Juan S Ramirez-Prado, Aala A Abulfaraj, Naganand Rayapuram, Moussa Benhamed, and Heribert Hirt. Plant immunity: from signaling to epigenetic control of defense. <u>Trends in plant science</u>, 23(9):833–844, 2018. → page 36
- [97] Raj Kumar Joshi, Swati Megha, Muhammad Hafizur Rahman, Urmila Basu, and Nat NV Kav. A global study of transcriptome dynamics in canola (brassica napus l.) responsive to sclerotinia sclerotiorum infection using rna-seq. Gene, 590(1):57–67, 2016. → page 37
- [98] Cole Trapnell, Brian A Williams, Geo Pertea, Ali Mortazavi, Gordon Kwan, Marijke J Van Baren, Steven L Salzberg, Barbara J Wold, and Lior Pachter. Transcript assembly and quantification by rna-seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature biotechnology, 28(5):511, 2010. → page 37
- [99] Kin Fai Au, Jason G Underwood, Lawrence Lee, and Wing Hung Wong. Improving pachio long read accuracy by short read alignment. <u>PloS one</u>, 7 (10):e46679, 2012. → page 37
- [100] Anthony Rhoads and Kin Fai Au. Pacbio sequencing and its applications. Genomics, proteomics & bioinformatics, 13(5):278–289, 2015.
- [101] Guoqiang Zhang, Min Sun, Jianfeng Wang, Meng Lei, Chenji Li, Duojun Zhao, Jun Huang, Wenjie Li, Shuangli Li, Jing Li, et al. Pacbio full-length cdna sequencing integrated with rna-seq reads drastically improves the discovery of splicing transcripts in rice. <u>The Plant Journal</u>, 97(2):296–305, 2019. → page 37
- [102] Misato Ohtani and Andreas Wachter. Nmd-based gene regulationa strategy for fitness enhancement in plants? <u>Plant and Cell Physiology</u>, 60(9): 1953–1960, 2019. → page 37
- [103] Feng Ding, Peng Cui, Zhenyu Wang, Shoudong Zhang, Shahjahan Ali, and Liming Xiong. Genome-wide analysis of alternative splicing of pre-mrna

under salt stress in arabidopsis. <u>BMC genomics</u>, 15(1):431, 2014. \rightarrow page 38

- [104] Geeng Loo Chong, Mung Hsia Foo, Wen-Dar Lin, Min May Wong, and Paul E Verslues. Highly aba-induced 1 (hai1)-interacting protein hin1 and drought acclimation-enhanced splicing efficiency at intron retention sites.
 <u>Proceedings of the National Academy of Sciences</u>, 116(44):22376–22385, 2019. → page 38
- [105] Jin Li, Yang Wang, Xi Rao, Yue Wang, Weixing Feng, Hong Liang, and Yunlong Liu. Roles of alternative splicing in modulating transcriptional regulation. BMC systems biology, 11(5):89, 2017. → page 38
- [106] Tony Kuo, Martin C Frith, Jun Sese, and Paul Horton. Eagle: Explicit alternative genome likelihood evaluator. <u>BMC medical genomics</u>, 11(2):28, 2018. \rightarrow page 39
- [107] Adam Voshall and Etsuko N Moriyama. Next-generation transcriptome assembly and analysis: impact of ploidy. Methods, 2019. → page 39
- [108] Dawei Zhang, Qi Pan, Chen Tan, Bin Zhu, Xianhong Ge, Yujiao Shao, and Zaiyun Li. Genome-wide gene expressions respond differently to a-subgenome origins in brassica napus synthetic hybrids and natural allotetraploid. Frontiers in plant science, 7:1508, 2016. → page 39
- [109] Kevin A Bird, Robert VanBuren, Joshua R Puzey, and Patrick P Edger. The causes and consequences of subgenome dominance in hybrids and recent polyploids. New Phytologist, 220(1):87–93, 2018. → page 40
- [110] Andrew Lloyd, Aurélien Blary, Delphine Charif, Catherine Charpentier, Joseph Tran, Sandrine Balzergue, Etienne Delannoy, Guillem Rigaill, and Eric Jenczewski. Homoeologous exchanges cause extensive dosage-dependent gene expression changes in an allopolyploid crop. <u>New</u> Phytologist, 217(1):367–377, 2018. → page 40
- [111] RH Ramírez-González, P Borrill, D Lang, SA Harrington, J Brinton, L Venturini, M Davey, J Jacobs, F Van Ex, A Pasha, et al. The transcriptional landscape of polyploid wheat. <u>Science</u>, 361(6403): eaar6089, 2018. → page 40
- [112] Tatsuhiro Nomaguchi, Yoshiaki Maeda, Tomoko Yoshino, Toru Asahi, Leila Tirichine, Chris Bowler, and Tsuyoshi Tanaka. Homoeolog expression bias in allopolyploid oleaginous marine diatom fistulifera solaris. <u>BMC genomics</u>, 19(1):330, 2018. → page 40

- [113] Samuel Bottani, Nicolae Radu Zabet, Jonathan F Wendel, and Reiner A Veitia. Gene expression dominance in allopolyploids: hypotheses and models. <u>Trends in plant science</u>, 23(5):393–402, 2018. → page 41
- [114] Berline Fopa Fomeju, Cyril Falentin, Gilles Lassalle, Maria J Manzanares-Dauleux, and Régine Delourme. Homoeologous duplicated regions are involved in quantitative resistance of brassica napus to stem canker. Bmc Genomics, 15(1):498, 2014. → page 46
- [115] Jian Wu, Peipei Chen, Qing Zhao, Guangqin Cai, Yue Hu, Yang Xiang, Qingyong Yang, Youping Wang, and Yongming Zhou. Co-location of qtl for sclerotinia stem rot resistance and flowering time in brassica napus. <u>The</u> Crop Journal, 7(2):227–237, 2019. → page 47