Justus Liebig University, Giessen
Institute of Agronomy and Plant Breeding I
Department of Plant Breeding

Regulatory and non-coding patterns associated to genome diversification and early development in an allopolyploid crop

Inaugural Dissertation for a Doctorate Degree in Natural Sciences
in the Faculty of Agricultural Sciences, Nutritional Sciences
and Environmental Management

Examiners
Prof. Dr. Rod J. Snowdon
Prof. Dr. Matthias Frisch

Submitted by
Mauricio Orantes-Bonilla
San Salvador, El Salvador
Giessen, 2023

In loving memory of

A.E. Bonilla & M.T. Haggard

Table of contents

1. General introduction 4

1.1 Hybrid breeding and genetic diversity in *B. napus* 5

1.2 Molecular principles of heterosis 7

1.3 Genomic structural variants in *B.napus* 10

1.4 Structural variants in the context of heterosis 13

1.5 Differential gene expression and methylation in heterosis 15

1.6 Objectives 18

2. Frequent spontaneous structural rearrangements promote genome diversification in a *Brassica napus* F1 generation 19

3. Transgressive and parental dominant gene expression and cytosine methylation during seed development in *Brassica napus* hybrids 34

4. Discussion 60

4.1 Spontaneous rearrangements and genomic diversity 61

4.2 Gene regulation in seed development 64

4.3 Transcriptomic and epigenomic patterns in early plant developmental stages 66

4.4 Effects of structural variants in gene regulation and methylation 70

4.5 Conclusion 90

5. Summary 75

6. Zusammenfassung 77

7. References 79

List of abbreviations 102

Declaration 104

Acknowledgements 105

1. General introduction

1.1 Hybrid breeding and genetic diversity in *Brassica napus*

Oilseed rape, *Brassica napus* (genome AACC; 2n =38), is a major crop grown for edible oil, feed, and biofuel worldwide. It is estimated that the global production of oilseed rape has risen from 59 to 72 million tonnes over the last decade, making it the most produced oil crop after oil palm and soybean (FAO 2022). Like many other relevant crops, oilseed rape is a polyploid, specifically an allopolyploid, derived from a spontaneous interspecific hybridization between *Brassica rapa* (AA, 2n = 20) and *Brassica oleracea* (CC, 2n =18) followed by a chromosome doubling event. Although the exact origin of the *B.napus* progenitors is still undetermined, recent genome-wide studies suggest European turnip (*B. rapa* ssp. *rapa*) as the maternal ancestor and at least four *B. oleracea* subspecies as putative paternal ancestors (Lu et al. 2019). Despite its recent origin, oilseed rape has diverged genetically into winter, spring, semi-winter and swede ecotypes adapted to diverse agroecological environments (Schiessl et al. 2017; Wu et al. 2019; Song et al. 2020). The adaptability to different environments is partially associated to whole genome duplications that *B. napus* has experienced and its allopolyploid nature that allows diversification through gene presence-absence and copy number variation driven by genomic rearrangements between homoeologous regions (Hurgobin et al. 2018; An et al. 2019; Walden et al. 2020).

Oilseed rape has been cultivated for oil and technical purposes since the 18th century in central Europe, however its usage became broader once low-glucosinolate and low-erucic acid traits were refined in *B. napus* in the 1970s and 1980s (Friedt et al. 2018; Canola Council of Canada 2021). Nowadays, oilseed rape is extensively grown worldwide, with most varieties in large oilseed rape producing countries like Canada, China and Germany being developed using hybrid breeding systems (Bonjean et al. 2016; Malla and Brewin 2019; BSA 2022). Hybrid breeding refers to the crossing of two genetically distant, highly homozygous lines to develop hybrids that display heterosis, i.e. an enhanced performance in comparison to the parents at a

given environmental setting. The use of hybrids has been facilitated by the development of cytoplasmic male sterility (CMS) or genic male sterility (GMS) systems and has resulted into increasing yields in many globally important crops including maize, sugarbeet, sunflower, tomato, rice and oilseed rape (Steeg et al. 2022; Bohra et al. 2016). Despite these achievements, the understanding of heterosis remains to be fully elucidated.

Comprehending heterosis is fundamental to break the genomic bottleneck that has largely arisen due to constant or limited germplasm use (Louwaars 2018). For instance, the high selection for low seed glucosinolate and erucic acid levels in oilseed rape enabled the release of suitable cultivars for animal feed and oil production. Nevertheless, it also led to a simultaneous reduction in genetic diversity (Körber et al. 2012; Friedt et al. 2018). One alternative approach for expanding genetic diversity in rapeseed breeding is the use of introgressions from close relatives that belong to the so-called “U triangle” (U 1935), a systematic description showing the relationships between the three diploid *Brassica* species *B. rapa*, *B. oleracea* (CC; 2n=18) and *B. nigra* (BB; 2n=16) and their respective allopolyploids *B. napus*, *B. juncea* (AABB, 2n=36) and *B. carinata* (BBCC; 2n=34). Despite genomic and transcriptomic shocks observed during subgenomic collisions and initial breeding barriers (Samans et al. 2017; Shin et al. 2022), many of the crosses within the U triangle can lead to relatively stable genotypes with interesting agricultural traits (Quezada-Martinez et al. 2021; Obermeier et al. 2022). In addition, multiple studies have shown that introgressions from genetically distant *B. napus* genotypes can lead to high performances varieties (Basunanda et al. 2010; Zou et al. 2018; Hu et al. 2021). The use of the U triangle in breeding revolves around the concept of polyploidy and its associated intrinsic effects like genome dosage variation and homoeologous exchanges, therefore dissecting and analyzing heterotic patterns in U triangles species, such as *B.napus*, is of utmost interest to advance polyploid hybrid breeding.

Oilseed rape and the other *Brassica* allopolyploids are the phylogenetically closest allopolyploid crop species to the model plant *Arabidopsis thaliana*, which is also a member of the Brassicaceae family. *A. thaliana* has been widely employed in research and described extensively due to its fast life cycle and the availability of divergent and diverse populations (Meinke et al. 1998; Provart et al. 2016). The close phylogenetical linkage to *A. thaliana* makes *B. napus* an ideal model polyploid for research of heterosis in agriculture, as many other major crops possess polyploid genomes including sugarcane, coffee, common oat, potato and wheat. Numerous studies analyzing heterosis at diverse stages and in different populations of *A. thaliana* are available (Alonso-Blanco et al. 2016; Moreno et al. 2022) which can be related to oilseed rape research with regard to flowering time, seed yield and cold tolerance traits (Leijten et al. 2018; Tang et al. 2021; Alahakoon et al. 2022). Lastly, the concept of heterosis in oilseed rape has been mostly explained by quantitative genetics and genomics (Radoev et al. 2008; Zhao et al. 2016; Werner et al. 2018), hence, detailed study of genomic mechanisms associated to heterosis in allopolyploids remains fundamental to elucidate heterotic clues in allopolyploids.

1.2 Molecular principles of heterosis

Heterosis was described as early as the late 19th century by biologist Charles Darwin who observed plant hybrids having taller height and earlier flowering time in comparison to their parents (Darwin 1876). However, the concept of heterosis still required further elucidation of inheritance hypotheses before being properly established. Earlier research on trait inheritance in peas by the pioneer of genetics, Gregor Mendel, suggested that information was being passed from parents to offspring, yet the underlying mechanisms remained unknown (Mendel 1866). It was not until almost half a century later that independent cytological studies in sea urchins (Sutton 1902) and western lubber grasshoppers (Boveri 1902) demonstrated that chromosomes were indeed the carriers of hereditary information. Nonetheless, the protein coding regions responsible for specific phenotypes remained undiscovered. In 1909, the term “gene” was

coined by botanist Wilhelm Johannsen as the basic inheritance unit determining distinct phenotypes (Johannsen 1909), which consequently triggered a revolution in biology and genetics. Just half a decade later, botanist George Shull not only applied the recent gene concept in intraspecies crosses but also confirmed enhanced hybrid performance in diverse plant species. Hence, proposing the term "heterosis" for the first time to unify previous terminology like "heterozygosis" or "heterozygotic stimulation" under a single term (Shull 1914, 1948).

Heterosis has been largely studied under three different, non-mutually exclusive hypotheses explaining gene allele interactions, known as the dominance, over-dominance and epistasis hypotheses, respectively (Fujimoto et al. 2018; Wu et al. 2021b). Dominance refers to one allele being responsible for a positive phenotype, as proposed first by Bruce (1910), while over-dominance stands for the interaction of two distinct alleles from the same gene that result into a desirable phenotype (Shull 1908). The discovery of non-allelic gene interactions, known as epistasis, was also argued by Powers (1944) to be a source of heterosis. Further research in the early and mid-20th century also proved that, contrary to a one gene-one enzyme concept, single genes could sometimes account for multiple phenotypes, a phenomenon later described as pleiotropy (Stearns 2010). Moreover, the complexity of certain traits also led to the discovery of polygenes, i.e. a set of multiple genes related to a single phenotype (Mather 1949). The accumulation of such works paved the way for quantitative genetics by pioneers like Ronald Fisher and Sewall Wright which consequently provided the basis of the modern genetic-based heterosis hypotheses (Hallauer 2007).

Some of the first reports of heterosis in oilseed rape remount to silique number, oil content and root development studies in the late 50's and 60's, and to biomass increase in swede and forage types in the 70's (Johnston 1971; McNaughton and Munro 1972; Wagner 1954). Advancements in molecular biology techniques like the polymerase chain reaction (PCR), microarrays and sequencing enabled detailed validation studies that linked specific genes and quantitative trait

loci (QTL) to heterotic traits in rice, maize and Arabidopsis (Hochholdinger and Hoecker 2007). Early investigations involving restriction fragment length polymorphism (RFLP), amplified fragment length polymorphisms (AFLP) and simple sequence repeats (SSR) revealed vernalization and flowering loci relevant for performance as well as heterotic patterns in oilseed rape traits influenced by dominance, over-dominance and epistatic effects (Ferreira et al. 1995; Diers et al. 1996; Basunanda et al. 2010; Radoev et al. 2008; Shi et al. 2011). Breakthroughs in next-generation sequencing (NGS) have led to innovative heterotic analyses where low coverage sequencing and microarrays allowed massive genotyping of diverse population and identified heterotic haplotypes and yield related genes (Snowdon et al. 2015; Lu et al. 2017; Luo et al. 2019). Long-read sequencing has also demonstrated a great potential in finding heterotic loci masked by repetitive sequences as exemplified in maize and rice (Li et al. 2020a; Wang et al. 2020; Sun et al. 2022).

One remarkable aspect of heterosis is that it manifests at specific developmental stages and tissues. For instance, advantages in hybrids are sometimes notably observed in mid-late plant development stages like flowering onset (Dickert and Tracy 2002; Fang et al. 2021) but also during early stages like seed and seedling development (Hochholdinger and Hoecker 2007; Jahnke et al. 2010; Liu et al. 2020). Early stage heterosis in *B. napus* has been recorded in seed and seedling traits through genome-wide and genetic analyses (Basunanda et al. 2010; Hatzig et al. 2015; Hatzig et al. 2018). In this regard, seed formation constitutes an appealing stage for heterotic research as two parental genomes with diverse allele compositions collide, in some cases, leading to F_1 plants with increased biomass and seed yield (Jahnke et al. 2010; Wang et al. 2017; Zhu et al. 2020). Moreover, epigenomic features like gamete-specific methylation and genomic imprinting are known to be more prominent during seed formation in most angiosperms (Bird et al. 2018; Batista and Köhler 2020). Although quantitative genetics have brought a more complete understanding on heterosis in crops, it does not fully exploit all

molecular mechanisms that might be associated to it, like structural variants, small RNAs (sRNAs) and epigenomics features.

1.3 Genomic structural variants in *B. napus*

Genomic structural rearrangements have been recorded since the early and mid-20th century as part of gene duplications and chromosome aneuploidy in diverse species (Bridges 1936; Nowell and Hungerford 1960; Escaramís et al. 2015). Cytological and karyotypic research allowed the distinction of some of the earliest variants in plants such as transposable elements (TEs) derived-insertions and chromosomal rearrangements (Stebbins 1945; McClintock 1950; Jackson 1982). The development of NGS accelerated the identification of different kinds of sequence variants (Saxena et al. 2014) which are now typically classified by their length as single nucleotide polymorphisms (SNPs), small insertion-deletion polymorphisms (indels), structural variants (SVs) and larger-scale rearrangements such as presence-absence variation (PAVs), copy number variation (CNVs) and homoeologous exchanges (HEs).

Although there is no consensus on which length ranges specifically defines each variant type, indels are regarded as presence-absence variants not longer than 50 bp, whereas SVs are normally defined as variants surpassing such a length, which can be classified into insertions, deletions, duplications, inversions and translocations (Gabur et al. 2019; Mahmoud et al. 2019). Structural variants that are longer than 1 Mbp tend to be further labelled based on their origin or population-wide effects. CNVs for instance refer to large segmental deletions or duplications that lead to changes in the number of copies from a specific genomic region. Similarly, PAVs are extremely large forms of CNV present in some genotypes within a population, and consequently, dividing accessions by core and disposable gene sets (Springer et al. 2009). HEs refer to large genomic rearrangements intrinsic to allopolyploids between homoeologous chromosomes of distinct subgenomes (Song et al. 1995).

Although the presence of small variants had been postulated already as a source of genetic variation in the first decade of this century. Nonetheless, further novel genomic studies at that time revealed that mid and large scale variants could generate genetic and phenotypic diversity (Escaramís et al. 2015; Yuan et al. 2021). The origin of such structural variants has been associated to distinct events like DNA replication errors, non-allelic recombination and double-strand break repair (Escaramís et al. 2015; Belyeu et al. 2021). Although determining the exact origin of structural variants in each study is not always feasible, the number of publications in agronomy referring to "structural variation" in the last decade alone has increased from 7075 in 2011 to 22654 in 2021 (Digital Science 2018). This highlights the growing relevance of structural rearrangements in agricultural research, and the putative role and potential of decreasing sequencing costs and advancements in NGS for discovery of structural variants (Muir et al. 2016).

Oilseed rape is prone to extensive rearrangements partially due to non-homologous chromosome pairing and its flexibility to cross with its ancestors and related species (U 1935; Attia and Röbbelen 1986; Ihien Katche et al. 2022). RFLP markers have aided in validating non-reciprocal homoeologous exchanges (NRHE) in oilseed rape (Parkin et al. 1995; Sharpe et al. 1995; Song et al. 1995). Contrasting to NRHE, where one subgenomic locus is duplicated and its homoeologous copy deleted, reciprocal homoeologous exchanges (RHE) involve homoeologous loci that are proportionally shifted between subgenomes. This kind of event can also be validated through RFLP analyses in *B. napus* segregating populations derived from diverse winter and spring *B. napus* accessions (Osborn et al. 2003; Udall et al. 2005).

In addition, the use of microsatellite markers also contributed to analyzing genetic variation among *B.napus* accessions and other Brassicaceae species at small basepair resolution (Plieske and Struss 2001; Hasan et al. 2006). The development of high-throughput genotyping through SNP microarrays accelerated research on genomic variation and the identification of QTL

associated with flowering time, disease response and seed quality in *B. napus* (Hayward et al. 2012; Liu et al. 2013; Dalton-Morgan et al. 2014; Raman et al. 2014; Clarke et al. 2016). SNP arrays have also helped in inferring large scale variants like CNV and HEs in *B.napus* (Stein et al. 2017; Samans et al. 2017; Higgins et al. 2018).

The assembly of the first oilseed rape reference genome by Chalhoub et al. (2014) facilitated the identification of structural variation calling through reference alignment of short read libraries (Mahmood et al. 2016; Wang et al. 2018; An et al. 2019). Short reads not only enabled a more precise variant breakpoint identification but also served in analyzing loci and genotypes not covered in standard SNP arrays. Long read sequencing technologies like Pacific Biosciences (PacBio) and Oxford Nanopore Technology (ONT) platforms have facilitated genome-wide comparisons through more complete reference assembling (Lee et al. 2020; Rousseau-Gueutin et al. 2020) and have contributed to detecting mid and large structural variants in complex genomic loci with high repetitive sequences. In oilseed rape, long read sequencing has enabled the detection of variants responsible for expanding genomic diversification and adaptation (Song et al. 2020; Chawla et al. 2021) and for modifying flowering time and disease resistance phenotypes (Vollrath et al. 2021b; Vollrath et al. 2021a).

Due to ancestral genome-wide duplications and polyploidization events, allopolyploid crops often lead to more complex breeding scenario where gene copy number variation, homoeologous exchanges and subgenomic biases shape the phenotype in a different manner than in diploids. *B. napus* is an adequate model crop for polyploidy studies where breakthroughs in analyses of structural variants can be used to enhance current understanding of rearrangements, and their implications in hybrid breeding and heterosis in polyploids crops.

1.4 Structural variants in the context of heterosis

Heterosis has been analysed in plants mostly from a quantitative genetic perspective where heterotic effects are explained as part of allelic combinations and interactions. Although alleles are in way a combination of gene variants that have occurred through diversification. Nonetheless, the extent at which intra- and intergenetic variations impact heterosis remains to be assessed. Structural variants have contributed to generating intragenetic allelic differences between populations in plant germplasms (Fuentes et al. 2019; Alonge et al. 2020; Guan et al. 2021) and have also helped in developing novel intergenetic variation with potential applications for genomic diversity and adaptation (Gullotta et al. 2022; Pokrovac and Pezer 2022).

An early model of structural variants leading to heterosis was proposed by Wasserman (1968) based on cytogenetic observations where inversions in hybrids were putatively linked to overall fitness. Springer et al. (2009) carried some of the first genomic and transcriptomic work analyzing the potential of structural variants in parental lines in maize. Their work revealed that parental CNVs and PAVs were likely involved in heterosis when crossing B73 and Mo17, two maize accessions frequently employed in hybrid breeding. Genome assembling through long read sequencing has allowed more complete maize references that have highlighted potential implications on heterosis through structural variants (Sun et al. 2018; Li et al. 2020a). Similar observations of parental specific structural rearrangements associated with hybrid performance have been reported in pigeon pea, rice and tomato (Sun et al. 2022; Yu et al. 2022; Saxena et al. 2021).

Structural rearrangements leading to gene duplication might be a mechanism used by polyploid species to overcome deleterious mutations and achieve enhanced performance, as described by Comai (2005). Moreover, polyploids are generally prone to increased genomic rearrangements triggered by genomic and transcriptomic shock during gamete and zygote formation (Chen and

Ni 2006; Otto 2007). Such rearrangements are not only relevant as coupling mechanisms to escape polyploidization dead-ends (Soltis et al. 2014; Mayrose et al. 2015), but also for expanding genomic diversity and modifying genome dosage (Lu et al. 2019; Walkowiak et al. 2020; Birchler and Veitia 2021; Bird et al. 2021b).

Genomic rearrangements with heterotic implications have been reported in polyploid plants as well. For instance, increase of paternal genome dosage induced heterotic height gain, and likely seed size, in *A. thaliana* F_1 triploids (Fort et al. 2016). Moreover, homoeologous exchanges in lines derived from allotetraploid F_1 hybrids led to quick genomic and phenotypic diversification in a rice population (Wu et al. 2021c). Unlike these examples, *B. napus* is a natural allopolyploid serving as a model for rearrangements and heterosis in other polyploid crops. Genomic introgressions from related *Brassica* species and *B. napus* genetic pools with distant genetic distance have shown to promote heterosis in *B. napus* (Zou et al. 2010; Qian et al. 2005; Girke et al. 2012). Introgressions employing U triangle species have also been shown to promote genomic diversity, as comprehensively reviewed by Quezada-Martinez et al. (2021) and demonstrated in previous studies (Gaebelein et al. 2019; Katche et al. 2021).

Genome-wide association studies (GWAS) and linkage mapping in oilseed rape have also revealed genomic variants significantly associated with performance and quality traits such as response to sclerotinia stem rot, branching control, seed oil content and seeds per silique (He et al. 2017a; Roy et al. 2021; Zhao et al. 2022; Liu et al. 2022b). Copy number variation on flowering time and disease resistance genes have also been reported in oilseed rape as a source of phenotypic diversification with implications in performance (Schiessl et al. 2017; Gabur et al. 2020; Dolatabadian et al. 2022). Genomic rearrangements in polyploids typically lead to genetic imbalances that modifies transcriptomic networks (Lloyd et al. 2018; Birchler and Veitia 2021). Thus, understanding the role of coding and regulatory features is pivotal to clarifying heterosis.

1.5 Differential gene expression and methylation in heterosis

Overcoming the genetic constrain caused through targeted germplasms usage has constituted one of the main concerns and challenges in modern plant breeding. Mining for trait diversity is evolving from marker assisted selection to a more pragmatical genomic selection (GS) approach which is leading key advancements in crop enhancement (Bhat et al. 2016). Nonetheless, the identification and classification of regulatory and non-coding features in the context of improved performance remains unelucidated (Zanini et al. 2022). A trend of transcriptomics and epigenomic analyses have been employed in the last decades to detect such features through gene expression and small RNA expression and methylation techniques.

Genes are often differentially expressed between contrasting genotypes at critical crop developmental stages with effects on flowering time, drought stress and diseases susceptibility (Kebede et al. 2018; Sharma et al. 2019; Jian et al. 2019). The delimitation of heterosis to a single gene has been out of research consideration for years since heterosis is associated to diverse traits that are regulated by multiple genes at various developmental stages. In this regard, RNA sequencing (RNA-Seq) has brought advancements in understanding how various gene patterns contribute to overall enhanced performance as reported in maize, rice and oilseed rape through expression QTL (eQTL) (Li et al. 2018; He et al. 2022; Liu et al. 2022a). RNA-Seq has also contributed in building gene expression databases for oilseed rape and other crops (One Thousand Plant Transcriptomes Initiative 2019; Chao et al. 2020; Liu et al. 2021a).

Furthermore, allopolyploids present a challenging scenario for gene expression assays due to altered genomic dosages through genomic rearrangements and increased gene homologs (He et al. 2017b; Lloyd et al. 2018). Nevertheless, gene expression analyses have already allowed detailed dissection of vernalization, disease response and seedling development patterns in important allopolyploid crops like oilseed rape and wheat (Kippes et al. 2015; Körber et al. 2015; Schiessl et al. 2019a; Hiebert et al. 2020; Chittem et al. 2020). Genes interreacting

through transcriptomic interactions are also responsible for increased seedling biomass and early seed development in hybrids. Thus, highlighting the significance of transcriptomic networks in heterosis (Jahnke et al. 2010; Wang et al. 2017; Zhu et al. 2020; Xiong et al. 2022).

Moreover, epigenomics factors such as small RNAs and genome-wide methylation can shape the transcriptomic landscape by targeting gene expression and masking genetic effects within transcriptomic networks. Small RNAs can be classified in a large group span based on their length and origin. However, small interfering (siRNAs) and micro RNAs (miRNAs) remain among the most studied kinds with a range of databases and resources for post-hoc analyses in plants (Griffiths-Jones et al. 2006; Lunardon et al. 2020). siRNAs are double stranded RNAs sequences, typically between 20-30 nucleotides with endogenous or exogenous origins that modify gene expression through translation inhibition and mRNA cleavage (Carthew and Sontheimer 2009). The diverse origins of siRNAs have limited their detailed characterization in plants when compared to miRNA resources.Nonetheless, studies have found that siRNAs are generally active during seed formation through RNA-directed DNA methylation (RdDM) and TE silencing (Mosher and Melnyk 2010; McCormick 2018; Grover et al. 2020) with implications on heterotic patterns as detected in maize, Arabidopsis and oilseed rape (Groszmann et al. 2011; Shen et al. 2017; Seifert et al. 2018a).

Contrastingly, miRNAs are single stranded RNA generally around 22 nucleotides long which are encoded by specific endogenous genes for post-transcriptional silencing and cleavage of target genes (O'Brien et al. 2018). At least more than 48,000 mature miRNAs have been identified and classified based on their sequence similarity among species, thus, enabling the formation of large databases for transcriptomic network analyses (Ambros et al. 2003; Kozomara et al. 2019; Huang et al. 2022). miRNAs have been demonstrated to regulate the response to pathogens and abiotic stresses and to orchestrate flowering time and seed development as reviewed in various crops by Kamthan et al. (2015) and Dong et al. (2022b).

miRNA families in the Brassica genus play critical roles in regulating flowering time, seed development, drought adaptation and pathogen response have also been identified (Huang et al. 2013; Huang et al. 2018; Cui et al. 2020; Schiessl et al. 2020). Heterotic patterns linked to miRNA families during seed development, flowering, grain filling and fruit formation have been further reported in maize, wheat, chili and oilseed rape (Ding et al. 2012; Li et al. 2014; Shen et al. 2017; Yang et al. 2021).

Lastly, genome-wide DNA methylation is one epigenomic feature induced by environmental stress and inherited patterns, whereby DNA methyltransferases add a methyl group to the 5th carbon of cytosines (Hotchkiss 1948; Heard and Martienssen 2014; Quadrana and Colot 2016). The methylation of cytosines (C) in plants occurs mainly in three contexts based on the identity of the surrounding nucleotides: CpG, CHG and CHH; where "H" stands for adenine, thymine or cytosine and "p" represents the phosphodiester bond between adjacent nucleotides on the same DNA strand. DNA methylation diverges between species and contributes to trait variability by altering gene expression, triggering cis-trans elements and developing imprinting genomic features during embryo formation (Gallego-Bartolomé 2020; Kumar and Mohapatra 2021; Muyle et al. 2022). Furthermore, genomic methylation has been linked to hybrid performance in Arabidopsis, rice, soybean and oilseed rape (Ran et al. 2016; Lauss et al. 2018; Zhou et al. 2021; Chen et al. 2022). Interestingly, seed formation constitutes a key stage where parental-specific methylation shapes the offspring genome and epigenome with relevant effects on hybrid performance and inheritance in early developmental stage heterotic studies (Meyer et al. 2012; Alonso-Peral et al. 2017). Integrating multiple omics approaches could prove beneficial to revealing potential heterotic patterns in polyploid crops and dissect features contributing to environmental adaptation and enhanced performance.

1.6 Objectives

Heterosis compromises one of the elemental approaches for crop improvement that has been exploited over decades. Nevertheless, the effect of non-genetic allele interactions in heterosis have not been fully explored, especially, in polyploid crops. Therefore, the aim of the present research is to dissect key genomic, transcriptomic and epigenomic patterns associated with hybrids and heterosis. Firstly, this introductory chapter presents hybrid breeding and genomic diversity in the context of heterosis. Genomic rearrangements are reviewed in oilseed rape with implications on performance, and examples of distinct heterotic hypotheses are provided. At the end of the chapter heterosis is also examined under the scope of gene expression and epigenomic factors.

Chapter 2 describes an evaluation of spontaneous rearrangements in F_1 hybrids from a cross between two genetically distant oilseed rape parental lines. Interestingly, large-scale genomic rearrangements were discovered to generate rapid genomic diversity by altering gene copy number variation and modifying methylation patterns. In Chapter 3, F_1 hybrids from the same cross were evaluated under field conditions during seed and seedling development stages. Parental and hybrid-specific patterns in terms of gene expression, small RNAs and methylation were detected and discussed in regard to their implications for heterosis. Finally, Chapter 4 discusses the cumulative findings in light of current research on spontaneous genomic variants and genomic diversity. Gene regulation and epigenomic clues are also discussed in the context of early heterosis and future implications and multiomics approaches are assessed and proposed.

2. Frequent spontaneous structural rearrangements promote genome diversification in a *Brassica napus* F1 generation

Mauricio Orantes-Bonilla, Manar Makhoul, HueyTyng Lee, Harmeet Singh Chawla, Paul Vollrath, Anna Langstroff, Fritz J. Sedlazeck, Jun Zou and Rod J. Snowdon

Frontiers in Plant Science 13:1057953 (2022)

doi: 10.3389/fpls.2022.1057953

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 18 November 2022
DOI 10.3389/fpls.2022.1057953

OPEN ACCESS

EDITED BY
Feng Cheng,
Insitute of Vegetables and Flowers (CAAS), China

REVIEWED BY
Margaret Woodhouse,
Agricultural Research Service (USDA), United States
Yue Hao,
Translational Genomics Research Institute (TGen), United States

*CORRESPONDENCE
Rod J. Snowdon
Rod.Snowdon@agrar.uni-giessen.de

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Bioinformatics, a section of the journal Frontiers in Plant Science

RECEIVED 30 September 2022
ACCEPTED 31 October 2022
PUBLISHED 18 November 2022

CITATION
Orantes-Bonilla M, Makhoul M, Lee H, Chawla HS, Vollrath P, Langstroff A, Sedlazeck FJ, Zou J and Snowdon RJ (2022) Frequent spontaneous structural rearrangements promote rapid genome diversification in a *Brassica napus* F1 generation.
Front. Plant Sci. 13:1057953.
doi: 10.3389/fpls.2022.1057953

COPYRIGHT
© 2022 Orantes-Bonilla, Makhoul, Lee, Chawla, Vollrath, Langstroff, Sedlazeck, Zou and Snowdon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Frequent spontaneous structural rearrangements promote rapid genome diversification in a *Brassica napus* F1 generation

Mauricio Orantes-Bonilla[1†], Manar Makhoul[1†], HueyTyng Lee[1], Harmeet Singh Chawla[2], Paul Vollrath[1], Anna Langstroff[1], Fritz J. Sedlazeck[3], Jun Zou[4] and Rod J. Snowdon[1*]

[1]Department of Plant Breeding, IFZ Research Centre for Biosystems, Land Use and Nutrition, Justus Liebig University, Giessen, Germany, [2]Department of Plant Sciences, Crop Development Centre, University of Saskatchewan, Saskatoon, SK, Canada, [3]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, United States, [4]National Key Laboratory of Crop Genetic Improvement, College of Plant Science & Technology, Huazhong Agricultural University, Wuhan, China

In a cross between two homozygous *Brassica napus* plants of synthetic and natural origin, we demonstrate that novel structural genome variants from the synthetic parent cause immediate genome diversification among F1 offspring. Long read sequencing in twelve F1 sister plants revealed five large-scale structural rearrangements where both parents carried different homozygous alleles but the heterozygous F1 genomes were not identical heterozygotes as expected. Such spontaneous rearrangements were part of homoeologous exchanges or segmental deletions and were identified in different, individual F1 plants. The variants caused deletions, gene copy-number variations, diverging methylation patterns and other structural changes in large numbers of genes and may have been causal for unexpected phenotypic variation between individual F1 sister plants, for example strong divergence of plant height and leaf area. This example supports the hypothesis that spontaneous *de novo* structural rearrangements after *de novo* polyploidization can rapidly overcome intense allopolyploidization bottlenecks to re-expand crops genetic diversity for ecogeographical expansion and human selection. The findings imply that natural genome restructuring in allopolyploid plants from interspecific hybridization, a common approach in plant breeding, can have a considerably more drastic impact on genetic diversity in agricultural ecosystems than extremely precise, biotechnological genome modifications.

KEYWORDS
Brassica, polyploidy, genetic diversity, long-read sequencing, structural rearrangements

Introduction

Genetic and genomic diversity plays a key role in plant adaption to environmental changes. Plant breeding exploits such diversity to develop new varieties adapted to new growth environments or biotic and abiotic challenges (Louwaars, 2018). In many crops the exploitation of heterosis through hybrid breeding, achieved *via* expansion and differentiation of genetic and genomic diversity into distinct heterotic pools, has been crucial to breeding success (Labroo et al., 2021). However, both in hybrid and open-pollinated crops, genetic bottlenecks due to inbreeding or targeted trait selection are a constant obstacle in breeding programs (Hickey et al., 2019). Furthermore, numerous important crop species arose through allopolyploidization, a process that is sometimes described as an evolutionary dead-end because of the severe genetic bottleneck posed by small numbers of founders involved to allopolyploidization events (Soltis et al., 2014; Mayrose et al., 2015). Detailed analysis of genetic and genomic diversity in crop germplasm pools and breeding materials has become an important technique to identify and exploit diversity for breeding, and in recent decades genome-wide assays of sequence polymorphisms have become a relevant tool for genomic selection and gene discovery in crops. High-throughput sequencing technologies also provide a means to investigate the consequences and impact of polyploidization on genome structure and genome-scale diversity (Samans et al., 2017).

High-throughput genetic analysis of large plant populations normally implements single nucleotide polymorphism (SNP) arrays or genotyping-by-sequencing techniques which deliver low-cost, relatively simple and meanwhile reasonably standardized datasets for extrapolation or imputation of genome-wide DNA sequence patterns. New SNP variants can also arise spontaneously during mitosis and meiosis. Mutations of this kind occur in plants at different rates depending on the genome size and ploidy. For example, in *Arabidopsis thaliana*, between 1 and 5 *de novo* intra-varietal mutations have been shown to occur per generation, whereas higher rates are found in rice and other plants with larger genomes (Singer et al., 2021). However, such estimates do not account for internal and external variables that can induce mutation, like cell age, epigenetics or temperature (Schoen and Schultz, 2019).

Meiosis also generates other forms of genomic diversity, not only by recombination through crossovers (CO), but also *via* meiotic mutations in gametes due to errors in the repair of DNA double-strand breaks (DSB). In particular, non-homologous end-joining (NHEJ) (Kuo et al., 2021; Wang et al., 2021) can lead to a variety of *de novo* genomic rearrangements (Cai and Xu, 2007). According to the theoretical framework of Mendel's laws of inheritance (Mendel, 1866), meiotic mutations should be inherited in heterozygous form by all F1 offspring whose parents are highly homozygous. In plants, examples of non-Mendelian inheritance have been identified, for example template-directed extra-genomic sequence insertions in *A. thaliana* (Lolle et al., 2005) or selfish genetic elements leading to segregation distortion in rice (Yu et al., 2018).

In this study, we used long read sequencing to detect unexpected inheritance patterns across a set of F1 sister plants developed from a cross between the genetically diverse *B. napus* accessions Express 617 and *B. napus* G3D001. Express 617 is an inbred (F11), natural winter-type oilseed rape that has been widely used in genetic and genomic analyses (Lee et al., 2020), while G3D001 is an advanced homozygous synthetic *B. napus* line derived from crosses between *B. napus*, *B. rapa* (AA, 2n = 20) and *B. carinata* (BBCC, 2n = 34), as described by Zou et al. (2018).

The important oilseed crop plant *Brassica napus* (genome AACC, 2n = 38) originated only very recently (Chalhoub et al., 2014) from interspecific crosses between the closely related diploid progenitors *B. rapa* and *B. oleracea* (CC, 2n = 18). Its recent origin from a limited number of founders and intensive selection for important seed quality characters in the past several decades represent extreme genetic bottlenecks. Paradoxically, despite its narrow genetic basis, *B. napus* has very quickly become one of the world's most important oilseed crops and profited from tremendous breeding success. Its importance as an oilseed crop and its closeness to *A. thaliana* have made it an interesting model for polyploid crop evolution. A striking feature in this context is the broad prevalence of genomic structural variations (SV), first discovered as large-scale homoeologous chromosome exchanges in genetic mapping studies (Song et al., 1995) and later confirmed by fluorescence *in situ* hybridization (Xiong et al., 2011), transcriptome-based visualization (He et al., 2017), or whole-genome assembly and genome resequencing (Chalhoub et al., 2014; Samans et al., 2017; Hurgobin et al., 2018). In *B. napus*, homoeologous non-reciprocal translocations (HNRT) and other homoeologous recombinations between highly similar chromosomes are particularly prevalent and often very large in synthetic *B. napus*, but also commonly found in naturally-derived accessions (Chalhoub et al., 2014; Higgins et al., 2018).

Samans et al. (2017) postulated that elevated frequencies of HNRT in early generations after *de novo* polyploidization could be an important driver for novel genetic variation to overcome the allopolyploidy bottleneck in evolution and breeding. Subsequently, Higgins et al. (2018) used allele presence-absence data from the *Brassica* 60K SNP array (Mason et al., 2017) to detect *de novo* homoeologous recombination events in test-cross families derived from a panel of 11 *B. napus* cultivars, demonstrating for the first time that *de novo* HNRT indeed generates novel, unexpected genetic diversity during *B. napus* breeding. Using transcriptome sequencing, Lloyd et al. (2018) observed *de novo* homoeologous exchanges between individual plants of the same inbred *B. napus* cultivar.

In recent years, advances in long read sequencing techniques achieved highly accurate resolution of homoeologous

chromosome regions in allopolyploid genome assemblies (Lee et al., 2020; Rousseau-Gueutin et al., 2020). Furthermore, resequencing using long-read sequencing provided a technical platform for accurate, routine detection of SV in complex plant genomes, including that of *B. napus* (Mahmoud et al., 2019; Yuan et al., 2021). Unexpectedly, surveys of genome-wide SV extent and patterns using Oxford Nanopore Technology (ONT) and/or Pacific Biosciences (PacBio) long read sequencing techniques suggested that all individual *B. napus* accessions carry many small- to medium-scale SV events within genic regions, with direct functional implications (Gabur et al., 2020; Vollrath et al., 2021a) and a potentially major role as drivers of genetic diversity and phenotypic adaptation (Chawla et al., 2021).

Here, we used ONT long reads to sequence 12 F1 sister plants of *B. napus* derived from a single cross between two strongly homozygous parents. As in the F1 test cross families investigated previously by Higgins et al. (2018), the 12 sister hybrids in our study would be expected to be genetically uniform according to Mendel's law of uniformity were investigated in detail on a genome-wide scale for high-resolution detection of spontaneous genomic rearrangements that were not observed in the genomes of the two parents. This was achieved by combining putative SV alleles with read coverage information after alignment and SV calling. Homoeologous exchanges were linked to large structural rearrangements and methylation patterns were predicted from ONT reads to assess the putative transcriptomic and epigenomic effects from observed spontaneous mutations. We confirmed that genome-wide rearrangements derived from a recent allopolyploid plant can give rise to vastly different new genetic variants in just a single generation. These observations add to the growing body of evidence that homoeologous exchanges can lead to rapid and ongoing diversification of allopolyploids crops during evolution and breeding, despite the enormous bottleneck of a spontaneous interspecific hybridization.

Materials and methods

Development of F1 sister plants

Parental lines G3D001 (Zou et al., 2018) and Express 617 (Lee et al., 2020) were sown in Hawita propagation substrate "F.-E. Typ P" (Hawita Gruppe GmbH) and placed in a greenhouse chamber in Giessen, Germany with a 16:8 hour light/dark photoperiod and an average temperature and relative humidity of 5°C and 65% RH, respectively. After 6 weeks, Express 617 seedlings were transferred to a separate chamber for 10 weeks of vernalization at 5°C and 65% RH and 16:8 light/dark photoperiod. The seedlings were then re-transferred inside the greenhouse and grown along with G3D001 under the same conditions. Pollen from a single plant of the paternal parent G3D001 was used to pollinate a single plant of the maternal parent Express 617 after emasculating immature maternal flower buds with alcohol sterilized tweezers. The crosses were immediately labelled and covered with a plastic bag to prevent cross-pollination.

Material sampling and phenotyping

F1 seeds were harvested, sown and the resulting F1 sister plants were grown and vernalized under the same conditions as described for the cross parents. Twelve sister F1 plants, along with one plant each from the maternal and paternal parents, were transplanted to 120 litre large plant containers (Hohmann et al., 2016) containing a homogenized 60/40 sand-soil mix and grown side-by-side, alongside one plant per parental line, under semi-controlled conditions in a tunnel greenhouse at Rauischholzhausen, Germany to ensure a uniform growing environment in a large soil volume for all 14 plants. All plants in the greenhouse unit were phenotyped with a 3D PlantEye Dual-Scanner F500 (Phenospex) for 11 weeks from the seedling stage to full flowering stage to evaluate their morphological uniformity, and an identical watering regime was applied to all plants. The second or third youngest leaf of each plant was harvested at 11:00 am on the same day and then frozen with liquid nitrogen and stored at -80°C until further processing. Leaves from each F1 plant and the two parental plants used as crossing parents were subsequently ground in liquid nitrogen using a sterilized mortar and pestle.

DNA isolation and long read sequencing

High-molecular-weight (HMW) DNA was extracted following a previously described protocol (Chawla et al., 2021). DNA quality and length were evaluated with a Nanodrop spectrophotometer (Thermo Fisher), a Qubit 2.0 fluorometer (Thermo Fisher) and gel electrophoresis. Libraries were prepared using SQK-LSK109 ligation sequencing kits (Oxford Nanopore Technology) and were afterwards loaded on Oxford Nanopore R9.4.1 flow cells in a MinION sequencing device (Oxford Nanopore Technology) for the G3D001 plant used to develop the F1 sister plants, and in a PromethION (Oxford Nanopore Technology) sequencing platform for all other samples.

Base calling and long read data filtering

Raw electrical signals from plants grown in Rauischholzhausen were base-called using Guppy Basecaller v.4.0.11 (Oxford Nanopore Technology), in a virtual machine operating with Ubuntu 20.04.1 LTS with two NVIDIA Tesla 4 TU104GL

(NVIDIA Corporation) Graphic Processor Units (GPU) and using the following options: *–device cuda: 0,1:50% –kit SQK-LSK109 –num_callers 16 – disable_pings* and *–flowcell FLO-PRO002*. Reads from the G3D001 plant used as pollen donor for the F1 generation were base-called with the same settings except for *FLO-MIN106* with Guppy Basecaller v.5.0.7. Only reads with a quality score above 7 and length above 5000 nucleotides were kept using NanoFilt v.2.8.0 (de Coster et al., 2018). The filtered library quality was evaluated with NanoStat v.1.5.0 (de Coster et al., 2018) and genome-wide coverages were estimated.

Structural variation calling

Filtered long reads from G3D001 and all F1 sister plants were aligned against the Express 617 reference genome (Lee et al., 2020) using minimap2 v. 2.20 (Li, 2018) *map-ont* function with *-ax* settings. The output file of each alignment was then filtered using samtools v.1.12 (Li et al., 2009) *view* function, so that only reads with an alignment score above 50 were kept. Mid-sized structural variations longer than 30 bp and supported by at least 25 reads were called using sniffles v.1.0.12 (Sedlazeck et al., 2018). Only insertions and deletions detected through aligned and/or split reads, having precise breakpoints and with resolved lengths for insertions were kept and merged with the forced calling pipeline from SURVIVOR v.1.0.7 (Jeffares et al., 2017) to allow SV comparison across samples. Moreover, only SVs where G3D001 had a homozygous alternate allele, smaller than 50 kbp and without miscalled alleles from any sample were selected. Furthermore, SVs which had less than 90% of reads supporting the predicted allele were discarded, in order to reduce false positives due to residual heterozygosity. SVs having different alleles across F1 sister plants were identified based on the variant calling files and further visualized with the Integrative Genomics Viewer (IGV) tool (Robinson et al., 2011). Selected insertions were assessed with polymerase chain reaction (PCR) and agarose gel electrophoresis using DNA from all 12 F1 sister plants and their two parents.

Detections of large genomic rearrangements above 1 Mbp

1Mbp windows were prepared for each chromosome using bedtools (Quinlan and Hall, 2010) *makewindows* and *coverage* functions. 1Mb window coverages were then combined with the allele information and position of SVs that were putatively different across F1 sister plants in tab-delimited files based on the allele type: homozygous reference, homozygous alternate and heterozygous. The SV coverage and allele genotype were visualized using the circlize package (Gu et al., 2014). Regions larger than 1 Mbp in which one or more F1 genotype showed no heterozygous SV alleles were further visualized with IGV to estimate the rearrangement start and end based on the genomic positions at which the coverage halved. Due to the large memory requirements to display coverages from all genotypes in large regions at the same time, plots for each genotype were saved as images separately instead and then merged with GIMP for easier display. Moreover, the number of genes within each large rearrangement were found by using the *intersect* function from bedtools against the Express 617 gene annotation. The genes functions were assessed by blasting their complementary DNA (cDNA) sequences with BLASTn (Altschul et al., 1990) against the Araport 11 A. *thaliana* representative gene model cDNA sequences (Cheng et al., 2017). Only the hit with lowest e-value was kept, on the condition that it had an e-value lower than 1×10^{-4}, no opening gaps and a percentage of identity equal or higher than 90%. In addition, gene ontology (GO) enrichment was analyzed using ShinyGo v.0.76.2 (Ge et al., 2020) with a false discovery rate (FDR) of 0.05 as specified in Orantes-Bonilla et al. (2022a).

Centromere prediction

Centromeres were predicted to define their distance to detected large genomic rearrangements. Briefly, two centromere-specific repeat sequences CentBr1 (GenBank accession CW978699) and CentBr2 (GenBank accession CW978837) were used to estimate the approximate positions of the centromeric regions for each chromosome. The methods were based on (Mason et al., 2016), where the two sequences where aligned to the Express 617 chromosomal assembly using BLASTn (Altschul et al., 1990) with a cut off of at least 90% sequence similarity. Since these sequences are satellite repeats that are not limited to only centromeric regions, further refinement of the centromeric positions was required. Using the approximate position range obtained through the alignment results representing the centromere boundaries, we traced them back to the scaffolding process of Express 617 assembly to find breakpoints. Breakpoints were defined as the positions where two non-overlapping scaffolds were merged together through genetic maps (Lee et al., 2020). These breakpoints were then set as the refined version of the centromeric boundaries and were used to estimate the relative position of large rearrangements to centromeres.

Identification of homoeologous exchanges

Recent assemblies from *B. oleracea* (Lv et al., 2020) and *B. rapa* (Zhang et al., 2018) were concatenated and used for homoeologous exchange identification. Homologous gene pairs between the A and C subgenomes were located with inparanoid v.4.2 (O'Brien et al., 2005) using bootstrap, a BLOSUM80

(BLOcks SUbstitution Matrix) and an initial cut-off score of 60. Inparalogs with a similarity score equal or greater than 70 were selected for each gene. Only pairs with the highest similarity score were kept and only the first reported homologous gene pair was selected in cases where two or more gene pairs had the same similarity score. Quality filtered long reads from plants grown in Rauischholzhausen were aligned with minimap2 against the concatenated *B. napus* reference. Coverage across chromosomes was calculated using the *bamtobed* and *genomecov* functions from bedtools and used as input in a modified deletion-duplication pipeline previously described (Stein et al., 2017). Briefly, outlier regions with a coverage above 150 were discarded and segments equal or larger than 25000 bp, with a coverage that deviated by at least one standard deviation above or below the mean coverage, were called as duplication or deletion, respectively. Those segments in which a gene homolog was deleted and its reciprocal homolog was duplicated in a homoeologous chromosome, were considered as putative non-reciprocal homoeologous exchanges; these were further searched within large genomic rearrangements (> 1 Mbp) to determine if such large-scale rearrangements were indeed homoeologous exchanges. In cases where a large rearrangement was a deletion from a NRHE, then the corresponding duplication length was defined by the common genomic positions in which the coverage showed a 1.25-fold increase compared to its mean chromosome coverage and in which the coverage increased one standard deviation from the mean. For this purpose, the coverage was calculated in 100 kbp bins with bedtools *coverage* and bins with coverage above 100 were discarded to reduce mean bias due to outliers.

Long read DNA methylation analyses

Raw reads (fast5 files) from the plants grown in Rauischholzhausen were converted from multi read to single read format using the ont_fast5_api package (Oxford Nanopore Technology) while basecalled reads were concatenated and used to annotate raw reads with tombo v.1.5.1 (Stoiber et al., 2016) using first the *annotate_raw_with_fastq* followed by the *resquiggle* functions with the *overwrite* option. Modified cytosines in CpG, CHG and CHH methylation contexts were predicted with DeepSignal-plant v. 0.1.2 (Ni et al., 2021) *call_mods* function and the *model.dp2.CNN.arabnrice2-1_120m_R9.4plus_tem.bn13_sn16.both_bilstm.epoch6.ckpt* model from the same package. The log files were then examined and only samples where the estimated coverage surpassed 30x were selected for further analyses. The frequency of methylated cytosines was calculated using the *call_freq* function and split with the *split_freq_file_by_5mC_motif.py* script from DeepSignal-plant. The output files were then re-merged so that they could be compatible with DMRCaller v. 1.22.0 (Catoni et al., 2018) for further differentially methylated region identification using a custom bash script (Supplementary Data 1). The number of methylated cytosines and methylation level (proportion of reads supporting a methylated cytosine) in the genomic rearrangements larger than 1 Mb were calculated based on the output files and plot as heatmaps with the ComplexHeatmap (Gu et al., 2016) package. Differentially Methylated Regions (DMRs) were identified by comparing each F1 sister plant against G3D001. For this purpose, the DMRCaller *computeDMRs* function was employed to find DMRs in 1000 bp bins in chromosomes using the *bins* method, score testing and a 0.01 p value threshold. A minimum cytosine count of 4, methylation proportion difference of 0.4 and gap between bins of 0 were simultaneously set at as running parameters. Lastly, DMRs were intersected with exons, introns, repeats and 1 kbp upstream promoter regions from Express 617 using bedtools Furthermore, repetitive elements were assigned into repeat families with RepeatModeler (Smit and Hubley, 2008) to gain insights into linkages between genomic rearrangements, methylation and transposable element (TE) composition.

Results

Large genomic rearrangements diverge across F1 sister plants

Twelve F1 sister plants derived from a single cross between Express 617 (female recipient) and G3D001 (pollen donor) were sequenced using Oxford Nanopore Technology. Sequenced reads were aligned against the Express 617 reference assembly (Lee et al., 2020). The average depth of read coverage and N50 value after read filtering were approximately 42x and 34.9 kbp, respectively (Supplementary Table 1). A total of 3309 putative insertions and 1727 deletions, longer than 30 bp and supported by at least 25 reads, were detected before quality filtering as having a distinct SV allele in at least one F1 plant in comparison to the remainder of the F1 sister plants. SVs with low read allele support and low coverage were discarded. This resulted in a set of 189 and 338 high-confidence insertions and deletions respectively. PCR amplification of selected insertions from this filtered set (Supplementary Table 2) confirmed that they occurred only in one or a few of the F1 sister plants (Supplementary Figures 1-3). However, closer inspection of the sequence coverage and alleles in chromosome regions surrounding these putative SVs revealed that the detected SVs clustered in larger segmental rearrangements (Table 1) in specific chromosome regions (Supplementary Table 3).

The detected rearrangements spanned a range from 1.2 to 9.95 Mbp in length. A prominent rearrangement on chromosome C03 is displayed in Figure 1 as an example. Visualizations from large-scale rearrangements in other chromosomes are included in Supplementary Figures 4-7. As

TABLE 1 Genomic features from large segmental deletions and non-reciprocal homoeologous exchanges (NRHE).

Chromosome	F1 biological replicate	Chromosome coverage (x)	Segment coverage (x)	Start (Mbp)	End Mbp)	Length (Mbp)	Mid-sized SV alleles			Genes		Structural rearrangement type
							AA	AB	BB	DEL	DUP	
A09	6	40.35	22.07	39	40.20	1.2	0	0	64	261	422	NRHE
A10	5	40.46	22.07	18.82	20.49	1.67	0	0	107	447	261	NRHE
C01	2,3,4	46.06	19.84	18.70	20.70	2	0	0	31	108	-	Segmental deletion
C03	2	38.45	22.58	3.65	13.60	9.95	0	0	228	1852	-	Segmental deletion
C08	1	40.05	21.11	48.20	52.06	3.87	0	0	87	754	736	NRHE

Coverage is based on the Express 617 *Brassica napus* reference genome (Lee et al., 2020). Allele information is shown as AA, AB or BB for homozygous alternate, heterozygous or homozygous reference alleles respectively. Deleted and duplicated genes within large rearrangements are displayed by the DEL and DUP abbreviations accordingly. F1 sample numbers correspond to the 12 single F1 plants.

observed in Figure 1, the read coverage of chromosome C03 is mostly halved for *F1 biological replicate 2* in comparison to the rest of the genotypes and lacks insertions in this chromosome region that are specific to the paternal genotype. Interestingly, all large-scale rearrangements had a high frequency of homozygous reference alleles and halved read coverages, indicating that the segments were deleted from the inherited G3D001 chromosomes (Table 1, Supplementary Figures 4-7) in the respective F1 individuals.

Analysis of homoeologous exchanges showed that the putative segmental deletions detected in chromosomes A09, A10 and C08 in individual F1 plants are clustered in larger-scale, non-reciprocal homoeologous exchanges (Supplementary Figures 8-10, Supplementary Tables 4-5). These NRHEs include a deleted segment from chromosome C08 that has been replaced by a duplicated segment from chromosome A09 in *F1 biological replicate 6*, a deleted segment from chromosome A10 that has been substituted by a duplicated segment from chromosome C09 in *F1 biological replicate 5*, and a deleted segment from chromosome C08 that has been replaced by a duplicated segment from chromosome A09 in *F1 biological replicate 1*.

Impact of *de novo* SV on gene presence-absence

A total of 3422 genes were deleted and 1419 duplicated by segmental rearrangements across the 12 F1 sister offspring. Details of SV-induced gene copy number variation (CNV) are outlined in Table 1. The high rate of *de novo* genetic variation in a single, small family of F1 sister plants, reflecting the results of Higgins et al. (2018) in test-cross families, highlights the putative functional impact of chromosomal rearrangements *via* gene copy number variation. A clear validation of phenotype-genotype relationships is outside the scope in this study because each genotype is represented by only a single individual plant which prevents biological replicates to validate phenotypes. Nevertheless, preliminary phenotypic observations revealed large, unexpected phenological and developmental differences between individual plants. For example, 3D scanning-based phenotyping from the seedling to the full flowering stage revealed differences in plant height, leaf area and digital biomass between the *F1 biological replicate 1* and all other F1 sister plants (Supplementary Figures 11-13, Supplementary Table 6). Furthermore, this plant showed a similar phenology and development to that of Express 617, which was not the case for the other F1 sister plants. Although this might be an effect of the segmental C08 deletion and C09 duplication present in *F1 replicate 1*, nonetheless, additional F1 plants having the exact rearrangement would be required as replicates to validate the proposed hypothesis.

Additional gene copies found within the NHREs and segmental deletions in this study include *B. napus* orthologs of well-known flowering regulatory genes (*FLC, TFL1, ELF6*),

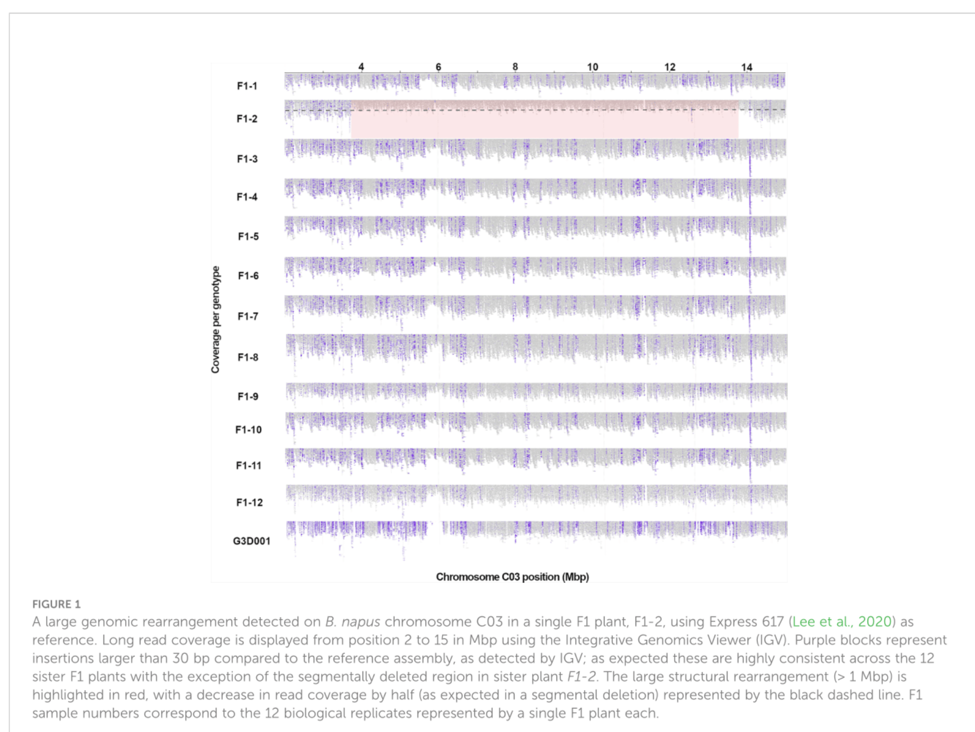

FIGURE 1
A large genomic rearrangement detected on *B. napus* chromosome C03 in a single F1 plant, F1-2, using Express 617 (Lee et al., 2020) as reference. Long read coverage is displayed from position 2 to 15 in Mbp using the Integrative Genomics Viewer (IGV). Purple blocks represent insertions larger than 30 bp compared to the reference assembly, as detected by IGV; as expected these are highly consistent across the 12 sister F1 plants with the exception of the segmentally deleted region in sister plant *F1-2*. The large structural rearrangement (> 1 Mbp) is highlighted in red, with a decrease in read coverage by half (as expected in a segmental deletion) represented by the black dashed line. F1 sample numbers correspond to the 12 biological replicates represented by a single F1 plant each.

along with genes corresponding to a variety of other functions (Supplementary Table 7) such as disease resistance (*WRKY-4, RVB1, EDR1, EDR4, EDR8*), embryo development (*EMB1873, EMB2107, EDA22, LEA4-5*), growth and development (*DWARF4, DWARF3, OPL1, PEAR2, ATSRG1*) or abiotic stress responses (*ATHSP70-1, ATHSP90-3, ATHMP44, ATPHB2, RCI3*). Although no common GO term was shared among all F1 sister plants, genes associated to auxin metabolism, oxidation processes, histone methylation, cell organelles and enzyme activity were enriched in chromosomes A09, A10, C01 and C08 (Supplementary Table 8).

Chromosome coverage plots showed that G3D001 lacks chromosome C02 and has two copies of chromosome A02, which in turn leads to their F1 offspring having three copies of A02 and one copy of C02 (Supplementary Figure 14). The full sequences and roles in meiosis from the different A02 chromosomes in G3D001 and F1 offspring are not yet clear. Although no large-scale genomic rearrangements occurred in either of those chromosomes, further studies are still required to elucidate their impact on inheritance patterns.

Chromosome rearrangements relate to DNA methylation patterns in F1 offspring

Genome-wide CpG, CHH, and CHG methylation was analyzed in F1 plants and their parents as described in Materials and Methods to investigate potential associations of methylation patterns with large rearrangements. The average read coverage after methylation prediction was approximately 36x (Supplementary Table 9). The number of methylated cytosines was higher in the CHH context, yet the methylation level was higher in the CpG and CHG contexts (Supplementary Tables 10-11), as reported in previous studies in oilseed rape and other plants (Shen et al., 2017; Bartels et al., 2018). Overall, the number of methylated cytosines were lower in genotypes with segmental deletions as expected. Nevertheless, the methylation level was more evenly distributed among all F1 sister plants despite the presence of large segmental chromosome rearrangements (Supplementary Figures 15-18) which could be due to uneven coverage distribution as outlined in figures (Supplementary Figures 19-23).

The large segmental deletion on chromosome C03 (Figure 1) was selected to illustrate methylation patterns within a large-scale structural variant. A lower number of methylated cytosines was observed in the F1 plant with the segmental deletion in the CpG methylation context in comparison to all other F1 sister plants as expected due to the deletion. Despite this, no large differences were observed in the overall methylation levels among F1 plants (Figures 2A, B). The same methylated cytosines and methylation level patterns are observed in the CHG and CHH contexts (Figures 2C, D). Differentially

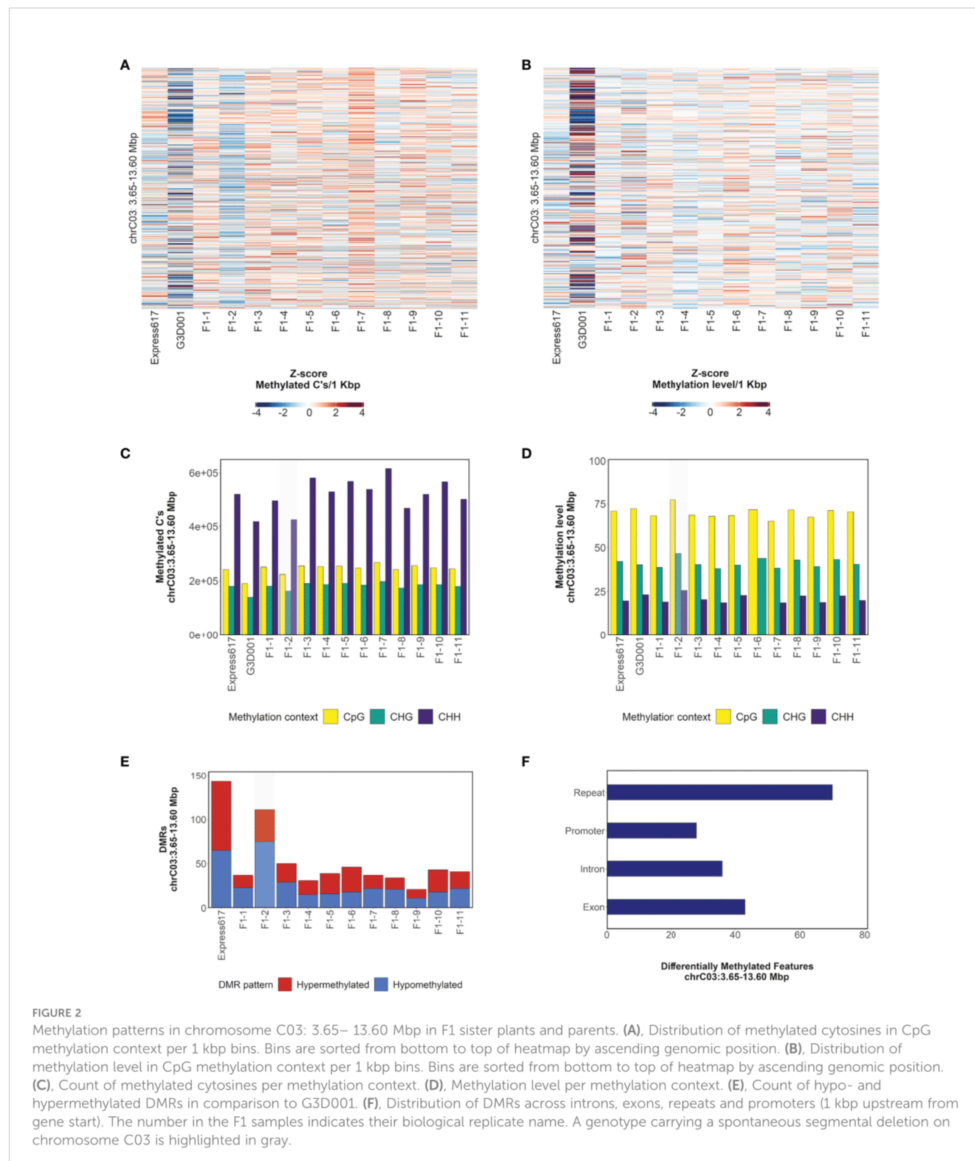

FIGURE 2
Methylation patterns in chromosome C03: 3.65– 13.60 Mbp in F1 sister plants and parents. **(A)**, Distribution of methylated cytosines in CpG methylation context per 1 kbp bins. Bins are sorted from bottom to top of heatmap by ascending genomic position. **(B)**, Distribution of methylation level in CpG methylation context per 1 kbp bins. Bins are sorted from bottom to top of heatmap by ascending genomic position. **(C)**, Count of methylated cytosines per methylation context. **(D)**, Methylation level per methylation context. **(E)**, Count of hypo- and hypermethylated DMRs in comparison to G3D001. **(F)**, Distribution of DMRs across introns, exons, repeats and promoters (1 kbp upstream from gene start). The number in the F1 samples indicates their biological replicate name. A genotype carrying a spontaneous segmental deletion on chromosome C03 is highlighted in gray.

methylated regions were more abundant in the F1 plant with the deletion, with 36 and 75 hyper- and hypomethylated DMRs accordingly (Figure 2E). Although genes and promoters were differentially methylated, most of the genomic methylated features inside DMRs corresponded to repetitive elements (Figure 2F). Closer evaluation of the repetitive element composition revealed that most belong to long terminal repeat (LTR) retrotransposons and diverse DNA transposons groups (Supplementary Table 12).

Overall, the number of DMRs in genotypes with segmental deletions was higher, and repeats were the most prevalent methylated feature within DMRs although no repeat family enrichment was found (Supplementary Figures 15-18). Furthermore, the average distance from DMRs to the closest gene was 2746 bp in genotypes with large-scale spontaneous rearrangements, making it feasible that they potentially play a role in transcriptomic regulation (Supplementary Table 13).

In contrast to deleted chromosome segments, duplicated segments did not show any divergent methylation pattern in comparison to F1 sister plants without the corresponding duplication (Supplementary Figures 24-29). For both deleted and duplicated regions, is noticeable that the overall methylation levels were not drastically changed. The mechanisms behind this phenomenon are still unknown.

Discussion

A total of five large-scale, spontaneous chromosome rearrangements were observed in distinct chromosomes of different F1 sister plants. All of these rearrangements could be shown to be caused by segmental deletions occurring in inherited paternal chromosomes. The size of the rearrangements ranged from 1.2 Mbp to 9.95 Mbp and resulted not only in gene losses, but also gene duplications *via* non-reciprocal homoeologous exchanges. Homoeologous exchanges are known to contribute to gene loss and duplication and influence flowering time, seed lignin content and seeds per silique in *B. napus* (Stein et al., 2017; Lloyd et al., 2018). Genetic diversity within populations through other large SV such as presence-absence variation (PAV) has been previously reported in Arabidopsis, maize, sorghum and chickpea (Pucker et al., 2016; Huang et al., 2021; Tao et al., 2021; Varshney et al., 2021) as well as in oilseed rape (Gabur et al., 2018; Vollrath et al., 2021b). However, studies of PAV normally present genomic variation across genetically divergent populations of a species rather than somatic or meiotic mutations within single genotypes. Frequently, PAV analyses focus on the concept of core and disposable genes, which despite the value for pangenomic studies does not illustrate the potential regulatory role of non-coding genomic regions (Zanini et al., 2022). In this study, spontaneous exchanges and segmental deletions covering both coding and non-coding regions could be related to unexpected genetic diversity within F1 offspring from two homozygous parents.

Based on the observed parental and F1 hybrid alleles, the segmental deletions observed in the sister F1 plants most likely arose during meiosis in the pollen donor. Although they could theoretically be due to spontaneous somatic mutations, these tend to be smaller in size than the large chromosomal segments seen here. Genomic features being inherited in unexpected patterns in early generations have been reported in the form of paramutations in maize, green pea, barley grass and Arabidopsis (Lolle et al., 2005; Hollick, 2017; Adu-Yeboah et al., 2021; Bente et al., 2021; Pereira and Leitão, 2021; Cao et al., 2022), and as selfish genetic elements in rice (Lolle et al., 2005; Hollick, 2017; Yu et al., 2018). In all cases, the reported mutations were limited to a gene-size scale and not to larger genomic features. In contrast, genomic rearrangements in allotetraploid *B.napus* (Higgins et al., 2018) and allohexaploid Brassica hybrids (Quezada-Martinez et al., 2022) showed large scale genomic rearrangements and evidence for *de novo* SV in cross offspring that were not observed in parental lines. Given the widely reported observation of homoeologous exchanges in both synthetic and natural *B. napus* genotypes (Song et al., 1995; Szadkowski et al., 2010; Xiong et al., 2011; Samans et al., 2017; Higgins et al., 2018; Hurgobin et al., 2018), the rearrangements observed in the current study are not altogether unexpected. In contrast to previous studies, however, which used more or less densely spaced genetic markers and segregation patterns to infer positions of large-scale segmental exchanges among homoeologous chromosomes, the use of long-read sequencing enables (i) detection of SV in regions with few genetic markers, (ii) higher-resolution definition of SV breakpoints, and (iii) direct determination of the gene content and allelic composition of genes impacted by duplication and deletion events.

Species are expected to have low to intermediate mutation rates to avoid loss of required biological information (Lesaffre, 2021) and retain fitness across generations. Nevertheless, mutations within populations can lead to significant functional changes. A previous report based on sequencing of 754 plant genomes showed that annual plants carry less somatic mutations in comparison to perennials, and that the average number of mutations per biological replicate ranged from to 0.69 to 23.9 in leaf samples (Wang et al., 2019). Another comprehensive study carried out on the 25th generation of a population generated by single-seed descent (SSD) in *A. thaliana* demonstrated that genomic mutations occurred randomly, and accounted for 90% of variance in gene bodies, along with accompanying epigenomic mutations (Monroe et al., 2022). In our study, a total of 3422 gene copies were deleted and 1419 were duplicated due to genomic rearrangements. These included key flowering time orthologues of *FLC*, *TLF1* and *ELF6* genes that drive photoperiod responses and diversification in *B.napus* (Guo et al., 2014; Raman et al., 2019; Schiessl et al., 2019a). Moreover, gene copy number variation covered stress response orthologues of the *WRKY* and *Hsp* gene families which influence

susceptibility to fungal pathogens and heat stress in oilseed rape (Yang et al., 2009; Yu et al., 2014). Although no common GO enrichment term was found among all chromosomes with spontaneous large-scale rearrangements; genes in chromosomes A09, A10, C01 and C08 were significantly enriched in diverse biological functions including auxin metabolism, oxidation processes and histone methylation. Gene copy number variation shapes post-polyploidization by altering overall genome dosage and triggering gene neo-/ subfunctionalization (Schiessl et al., 2017; Birchler and Veitia, 2021). Since the presence of homoeologous genes in allopolyploids influences the prediction of balanced or unbalanced gene expression after copy number variation events (Lloyd et al., 2018), detailed transcriptomic studies evaluating multiple tissues and environments might further elucidate the impacts of spontaneous structural rearrangements on genetic diversity and genome dosage.

Because we investigated individual, heterozygous F1 plants which cannot be biologically replicated for detailed phenotypic comparisons of seed-grown plants, a detailed analysis of phenotypic consequences from the genomic rearrangements is not possible. In general, the individual F1 sister plants showed a very uniform phenology and morphology, as would be expected in genetically identical F1 offspring from a Mendelian cross between two largely homozygous inbred parents. However, the F1 plant *F1 replicate 1*, which was found to carry a unique NHRE between chromosomes C08 and A09 (Supplementary Figures 11-13), was similar to the maternal line Express 617 in terms of height, digital biomass and leaf area throughout its life cycle, and dissimilar to the other 11 sister plants for these characters despite growing side-by-side in the same controlled environment. Because many genes were impacted by the various SV events, it is likely that other macro and micro-phenotypic traits could be affected by the spontaneous rearrangements in individual plants, although gene redundancy in the allopolyploid *B. napus* genome likely balances or buffers many effects from gene loss or inactivation due to rearrangements (Lesaffre, 2021). Nevertheless, the plant with a putative SV-driven impact on height, leaf area and biomass demonstrates the potential adaptive implications of frequent, spontaneous structural rearrangements as a source of novel genetic variation in a recent allopolyploid species with a narrow genetic diversity due to polyploidization and breeding bottlenecks.

Interestingly, the rearrangements on chromosomes A09, A10, C03 and C08 were located at or near telomeres, while pericentromeric regions were rearranged in chromosome C01 (Supplementary Table 14). This matches corresponding observations by Higgins et al. (2018), who also observed a higher frequency of homoeologous exchanges near the ends of *B. napus* chromosomes. Distal chromosome regions tend to have a higher frequency of crossovers (Aguilar and Prieto, 2021; Kuo et al., 2021), supporting the hypothesis of Samans et al. (2017) that homoeologous rearrangements in *B. napus* are driven by meiotic crossovers between homoeologous chromosomes. This is of high relevance in a breeding context since CO occurring during meiosis results in genomic exchange, and hence, population diversity (Samans et al., 2017; Lambing and Heckmann, 2018; Mason and Wendel, 2020; Blasio et al., 2022). Although most chromosomes only exhibit one CO per meiosis in most species (Fernandes et al., 2018), it might be expected that the unusual paternal ploidy and genomic structure (Supplementary Figure 14) could have played a role in the observed F1 patterns. It has been reported that the increase or loss of specific chromosomes can alter the number of CO in *B. napus* (Suay et al., 2014).

The methylation patterns were similar to results from previous studies in oilseed rape, where methylated cytosine counts were higher in the CpG context and lower in CHH context, while methylation levels displayed the opposite trend. Likewise, DMRs were mostly found in the CpG and CHG contexts and abundantly in upstream promoter regions, as also reported previously in *B. napus* (Shen et al., 2017; Wang et al., 2018). Differentially methylated repetitive elements included multiple simple and unclassified repeats from which conclusion drawing might be more limited; nonetheless, diverse transposons and LTR retrotransposons were also differentially methylated in large-scale rearrangements. LTRs are abundantly present in the plant kingdom and have been associated to polyplodization in *Brassica napus* (Vicient and Casacuberta, 2017; Chen et al., 2021; Song et al., 2020). The classification and functional evaluation of TEs remain challenging due to their repetitive sequences, especially in polyploids; nevertheless, the large-scale rearrangements reported in our study comprehended members of the LTR *Copia* and *Gypsy* families which also differed in methylation patterns. TEs are characterized by their role in speciation and adaptation (Serrato-Capuchina and Matute, 2018; Gill et al., 2021); hence, comprehensive TE studies can help assessing the effects of genomic rearrangements and methylation on TEs, and their contribution to genomic stability and transcriptome expression. In addition, recent advances in long-read sequencing technology have allowed the prediction of epigenomic features like DNA cytosine methylation in Arabidopsis and triticale (Kirov et al., 2021; Naish et al., 2021). In the present study, F1 plants carrying segmental deletions displayed consistently reduced methylated cytosine counts. This is expected since the number of available cytosines that can be methylated is reduced by the deletions. Despite this, their overall methylation levels remained similar to the rest of the offspring (Figure 2, Supplementary Figures 15-18). Furthermore, methylation levels in F1 individuals with duplicated segments was not higher than the rest of the F1 plants (Supplementary Figures 24-26). Although DMRs were still found mostly in F1s with rearrangements, it appears that the methylation levels were maintained to similar levels across all F1

sister plants. This suggests the presence of a mechanism which maintains overall balance in methylation levels despite genomic rearrangements, for example a maternal dominance which compensates methylation losses due to deleted regions in chromosomes inherited from the paternal parent. Methylation dominance has been previously reported in resynthesized *B. napus* at the subgenome level, but the mechanisms behind this phenomenon are still not clear (Bird et al., 2021). Future work is still needed to characterize methylation patterns and the role of genomic variants and epigenomics. The implications of methylation, however, are overall key to generating diversity as it has been reported for traits such as flowering time, plant height and stress resistance (Mercé et al., 2020; Omony et al., 2020).

As in previous, related studies, we showed that genomic diversity in *B. napus* can become rapidly increased within a single generation by large scale, spontaneous chromosome rearrangements. The adaptation and survival of natural polyploids after whole genome duplication (WGD) and putative genomic shock is still not elucidated. For instance, polyploidization might lead to genomically instable offspring and reproductive isolation; however, it is also recognized as an speciation mechanism (van de Peer et al., 2017; Pelé et al., 2018; Hörandl, 2022) and believed to contribute to environmental stress adaptation. Many species that underwent WGD have outperformed their progenitors and thrived, whereas their sister taxa did not (van de Peer et al., 2017). Interesting examples are further described by Edger et al. (2015), who found that WGD increased genetic diversity among glucosinolate genes in Brassicales to counter herbivore predation, and by Estep et al. (2014) who discovered a considerable increase in polyploid C_4 grasslands in the Late Miocene period. Further research is required to determine whether post-polyploidization occurs mainly through spontaneous genomic rearrangements or through environmental changes. Recent studies revealed that not all polyploidizations are linked to drastic genomic reshuffling and transcriptomic shocks, as reported in allotetraploids *A. suecica* and *B.rapa x Raphanus sativus* species (Burns et al., 2021; Shin et al., 2022).

The high rate of spontaneous rearrangements in the present study might lie in the synthetic nature of the paternal line, since resynthesized *B. napus* is associated with genomic instability (Szadkowski et al., 2010; Xiong et al., 2011). However, the frequency of large-scale SV is comparable to that reported by Higgins et al. (2018) in natural *B. napus*. Nevertheless, the parentage of this cross reflects potential scenarios of accelerated genomic diversity after formation of natural *B. napus*, representing an important source to enrich species diversity in a new polyploid. Our results underline previous findings showing that post-polyploidization genome restructuring can drastically expand gene diversity among offspring in just a single self-fertilized generation. Although genetic engineering has already shown great advantages in agriculture (Sedeek et al., 2019), sudden variation generated by spontaneous chromosomes rearrangements might be an alternative method to disrupt genetic bottlenecks in scenarios where genetic engineering is not feasible.

Epigenetic modifications and structural variations altogether have contributed not only to generate diversity in the formation of allopolyploid *B. napus* (Mason and Wendel, 2020; He et al., 2021) but also in modern ecotypes (Lu et al., 2019; Song et al., 2020). Genomic rearrangements have also been associated to changes in flowering time (Schiessl et al., 2019b; Chawla et al., 2021; Vollrath et al., 2021a), seed quality (Stein et al., 2017) and disease resistance (Gabur et al., 2020; Vollrath et al., 2021b) in *B. napus* cultivars. Intragenic structural variations within cultivars have also been reported in maize and wheat (Lesaffre, 2021). The present study adds a new example for rapid generation of novel genetic diversity through genome restructuring during meiosis in *B. napus*.

Data availability statement

The data presented in the study are deposited in the National Center for Biotechnology Information (NCBI) repository, accession number PRJNA837580. Please find the repository data link below: https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA837580.

Author contributions

RJS and JZ conceived and supervised the study. MOB drafted the manuscript and conducted the genomic and methylation bioinformatic analyses. HL contributed to the bioinformatic pipelines and carried the centromere prediction. MM contributed to the structural variation validation, data analyses, primer design and greenhouse trials. HSC and PV contributed to the generation and interpretation of long read data and structural variations and FJS to the genomic rearrangement analysis. AL performed the 3D scanning-based phenotyping and contributed to the greenhouse trial management. RJS revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was carried under the Joint Sino-German Research (2018) frame and was sponsored by the German Research Foundation (DFG grant number SN14/22-1 to RJS) and the National Natural Science Foundation of China (NSFC, grant number 31861133016 to JZ).

Acknowledgments

Greenhouse and laboratory experiments were supported by Regina Illgner, Liane Renno, Stavros Tzigos, Juliette Kellermann,

Birgit Keiner, Annette Plank, Daniela Quezada-Martinez and Andreas Eckert. Computational analysis was supported by the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) and the Bioinformatics Core Facility at JLU Giessen. The present research is available as a preprint at Biorxviv (Orantes-Bonilla et al., 2022b).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1057953/full#supplementary-material

References

Adu-Yeboah, P., Malone, J. M., Gill, G., and Preston, C. (2021). Non-mendelian inheritance of gene amplification-based resistance to glyphosate in *Hordeum glaucum* (barley grass) from South Australia. *Pest Manag. Sci.* 77, 4298–4302. doi: 10.1002/ps.6518

Aguilar, M., and Prieto, P. (2021). Telomeres and subtelomeres dynamics in the context of early chromosome interactions during meiosis and their implications in plant breeding. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.672489

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Bartels, A., Han, Q., Nair, P., Stacey, L., Gaynier, H., Mosley, M., et al. (2018). Dynamic DNA methylation in plant growth and development. *Int. J. Mol. Sci.* 19, 2144. doi: 10.3390/ijms19072144

Bente, H., Foerster, A. M., Lettner, N., and Mittelsten Scheid, O. (2021). Polyploidy-associated paramutation in arabidopsis is determined by small RNAs, temperature, and allele structure. *PloS Genet.* 17, e1009444. doi: 10.1371/journal.pgen.1009444

Birchler, J. A., and Veitia, R. A. (2021). One hundred years of gene balance: How stoichiometric issues affect gene expression, genome evolution, and quantitative traits. *Cytogenet. Genome Res.* 161, 529–550. doi: 10.1159/000519592

Bird, K. A., Niederhuth, C. E., Ou, S., Gehan, M., Pires, J. C., Xiong, Z., et al. (2021). Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid *Brassica napus*. *New Phytol.* 230, 354–371. doi: 10.1111/nph.17137

Blasio, F., Prieto, P., Pradillo, M., and Naranjo, T. (2022). Genomic and meiotic changes accompanying polyploidization. *Plants (Basel)* 11, 125. doi: 10.3390/plants11010125

Burns, R., Mandáková, T., Gunis, J., Soto-Jiménez, L. M., Liu, C., Lysak, M. A., et al. (2021). Gradual evolution of allopolyploidy in *Arabidopsis suecica*. *Nat. Ecol. Evol.* 5, 1367–1381. doi: 10.1038/s41559-021-01525-w

Cai, X., and Xu, S. S. (2007). Meiosis-driven genome variation in plants. *Curr. Genomics* 8, 151–161. doi: 10.2174/138920207780833847

Cao, S., Wang, L., Han, T., Ye, W., Liu, Y., Sun, Y., et al. (2022). Small RNAs mediate transgenerational inheritance of genome-wide trans-acting epialleles in maize. *Genome Biol.* 23, 53. doi: 10.1186/s13059-022-02614-0

Catoni, M., Tsang, J. M., Greco, A. P., and Zabet, N. R. (2018). DMRcaller: A versatile R/Bioconductor package for detection and visualization of differentially methylated regions in CpG and non-CpG contexts. *Nucleic Acids Res.* 46, e114. doi: 10.1093/nar/gky602

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Plant genetics. early allopolyploid evolution in the post-neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Chawla, H. S., Lee, H., Gabur, I., Vollrath, P., Tamilselvan-Nattar-Amutha, S., Obermeier, C., et al. (2021). Long-read sequencing reveals widespread intragenic structural variants in a recent allopolyploid crop plant. *Plant Biotechnol. J.* 19, 240–250. doi: 10.1111/pbi.13456

Cheng, C.-Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., and Town, C. D. (2017). Araport11: A complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J.* 89, 789–804. doi: 10.1111/tpj.13415

Chen, X., Tong, C., Zhang, X., Song, A., Hu, M., Dong, W., et al. (2021). A high-quality *Brassica napus* genome reveals expansion of transposable elements, subgenome evolution and disease resistance. *Plant Biotechnol. J.* 19, 615–630. doi: 10.1111/pbi.13493

de Coster, W., D'Hert, S., Schultzm, D. T., Cruts, M., and Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34, 2666–2669. doi: 10.1093/bioinformatics/bty149

Edger, P. P., Heidel-Fischer, H. M., Bekaert, M., Rota, J., Glöckner, G., Platts, A. E., et al. (2015). The butterfly plant arms-race escalated by gene and genome duplications. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8362–8366. doi: 10.1073/pnas.1503926112

Estep, M. C., McKain, M. R., Vela Diaz, D., Zhong, J., Hodge, J. G., Hodkinson, T. R., et al. (2014). Allopolyploidy, diversification, and the Miocene grassland expansion. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15149–15154. doi: 10.1073/pnas.1404177111

Fernandes, J. B., Séguéla-Arnaud, M., Larchevêque, C., Lloyd, A. H., and Mercier, R. (2018). Unleashing meiotic crossovers in hybrid plants. *Proc. Natl. Acad. Sci. U.S.A.* 115, 2431–2436. doi: 10.1073/pnas.1713078114

Gabur, I., Chawla, H. S., Liu, X., Kumar, V., Faure, S., von Tiedemann, A., et al. (2018). Finding invisible quantitative trait loci with missing data. *Plant Biotechnol. J.* 16, 2102–2112. doi: 10.1111/pbi.12942

Gabur, I., Chawla, H. S., Lopisso, D. T., von Tiedemann, A., Snowdon, R. J., and Obermeier, C. (2020). Gene presence-absence variation associates with quantitative *Verticillium longisporum* disease resistance in *Brassica napus*. *Sci. Rep.* 10, 4131. doi: 10.1038/s41598-020-61228-3

Ge, S. X., Jung, D., and Yao, R. (2020). ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36, 2628–2629. doi: 10.1093/bioinformatics/btz931

Gill, R. A., Scossa, F., King, G. J., Golicz, A. A., Tong, C., Snowdon, R. J., et al. (2021). On the role of transposable elements in the regulation of gene expression and subgenomic interactions in crop genomes. *Crit. Rev. Plant Sci.* 40, 157–189. doi: 10.1080/07352689.2021.1920731

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847–2849. doi: 10.1093/bioinformatics/btw313

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics* 30, 2811–2812. doi: 10.1093/bioinformatics/btu393

Guo, Y., Harloff, H. J., Jung, C., and Molina, C. (2014). Mutations in single FT- and TFL1-paralogs of rapeseed (*Brassica napus* l.) and their impact on flowering time and yield components. *Front. Plant Sci.* 17;5. doi: 10.3389/fpls.2014.00282

He, Z., Ji, R., Havlickova, L., Wang, L., Li, Y., Lee, H. T., et al. (2021). Genome structural evolution in *Brassica* crops. *Nat. Plants* 7, 757–765. doi: 10.1038/s41477-021-00928-8

He, Z., Wang, L., Harper, A. L., Havlickova, L., Pradhan, A. K., Parkin, I. A. P., et al. (2017). Extensive homoeologous genome exchanges in allopolyploid crops revealed by mRNAseq-based visualization. *Plant Biotechnol. J.* 15, 594–604. doi: 10.1111/pbi.12657

Hickey, L. T., Hafeez, N., Robinson, H., Jackson, S. A., Leal-Bertioli, S. C. M., Tester, M., et al. (2019). Breeding crops to feed 10 billion. *Nat. Biotechnol.* 37, 744–754. doi: 10.1038/s41587-019-0152-9

Higgins, E. E., Clarke, W. E., Howell, E. C., Armstrong, S. J., and Parkin, I. A. P. (2018). Detecting de Novo homoeologous recombination events in cultivated *Brassica napus* using a genome-wide SNP array. *G3 (Bethesda)* 8, 2673–2683. doi: 10.1534/g3.118.200118

Hohmann, M., Stahl, A., Rudloff, J., Wittkop, B., and Snowdon, R. J. (2016). Not a load of rubbish: simulated field trials in large-scale containers. *Plant Cell Environ.* 39, 2064–2073. doi: 10.1111/pce.12737

Hollick, J. B. (2017). Paramutation and related phenomena in diverse species. *Nat. Rev. Genet.* 18, 5–23. doi: 10.1038/nrg.2016.115

Hörandl, E. (2022). Novel approaches for species concepts and delimitation in polyploids and hybrids. *Plants (Basel)* 11, 204. doi: 10.3390/plants11020204

Huang, Y., Huang, W., Meng, Z., Braz, G. T., Li, Y., Wang, K., et al. (2021). Megabase-scale presence-absence variation with *Tripsacum* origin was under selection during maize domestication and adaptation. *Genome Biol.* 22, 237. doi: 10.1186/s13059-021-02448-2

Hurgobin, B., Golicz, A. A., Bayer, P. E., Chan, C.-K. K., Tirnaz, S., Dolatabadian, A., et al. (2018). Homoeologous exchange is a major cause of gene presence/absence variation in the amphidiploid *Brassica napus*. *Plant Biotechnol. J.* 16, 1265–1274. doi: 10.1111/pbi.12867

Jeffares, D. C., Jolly, C., Hoti, M., Speed, D., Shaw, L., Rallis, C., et al. (2017). Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nat. Commun.* 8, 14061. doi: 10.1038/ncomms14061

Kirov, I., Polkhovskaya, E., Dudnikov, M., Merkulov, P., Vlasova, A., Karlov, G., et al. (2021). Searching for a needle in a haystack: Cas9-targeted Nanopore sequencing and DNA methylation profiling of full-length glutenin genes in a big cereal genome. *Plants (Basel)* 11, 5. doi: 10.3390/plants11010005

Kuo, P., Da Ines, O., and Lambing, C. (2021). Rewiring meiosis for crop improvement. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.708948

Labroo, M. R., Studer, A. J., and Rutkoski, J. E. (2021). Heterosis and hybrid crop breeding: A multidisciplinary review. *Front. Genet.* 12. doi: 10.3389/fgene.2021.643761

Lambing, C., and Heckmann, S. (2018). Tackling plant meiosis: From model research to crop improvement. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00829

Lee, H., Chawla, H. S., Obermeier, C., Dreyer, F., Abbadi, A., and Snowdon, R. (2020). Chromosome-scale assembly of winter oilseed rape *Brassica napus*. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00496

Lesaffre, T. (2021). Population-level consequences of inheritable somatic mutations and the evolution of mutation rates in plants. *Proc. Biol. Sci.* 288, 20211127. doi: 10.1098/rspb.2021.1127

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Lloyd, A., Blary, A., Charif, D., Charpentier, C., Tran, J., Balzergue, S., et al. (2018). Homoeologous exchanges cause extensive dosage-dependent gene expression changes in an allopolyploid crop. *New Phytol.* 217, 367–377. doi: 10.1111/nph.14836

Lolle, S. J., Victor, J. L., Young, J. M., and Pruitt, R. E. (2005). Genome-wide non-mendelian inheritance of extra-genomic information in arabidopsis. *Nature* 434, 505–509. doi: 10.1038/nature03380

Louwaars, N. P. (2018). Plant breeding and diversity: A troubled relationship? *Euphytica* 214, 114. doi: 10.1007/s10681-018-2192-5

Lu, K., Wei, L., Li, X., Wang, Y., Wu, J., Liu, M., et al. (2019). Whole-genome resequencing reveals *Brassica napus* origin and genetic loci involved in its improvement. *Nat. Commun.* 10, 1154. doi: 10.1038/s41467-019-09134-9

Lv, H., Wang, Y., Han, F., Ji, J., Fang, Z., Zhuang, M., et al. (2020). A high-quality reference genome for cabbage obtained with SMRT reveals novel genomic features and evolutionary characteristics. *Sci. Rep.* 10, 12394. doi: 10.1038/s41598-020-69389-x

Mahmoud, M., Gobet, N., Cruz-Dávalos, D. I., Mounier, N., Dessimoz, C., and Sedlazeck, F. J. (2019). Structural variant calling: the long and the short of it. *Genome Biol.* 20, 246. doi: 10.1186/s13059-019-1828-7

Mason, A. S., Higgins, E. E., Snowdon, R. J., Batley, J., Stein, A., Werner, C., et al. (2017). A user guide to the *Brassica* 60K Illumina Infinium™ SNP genotyping array. *Theor. Appl. Genet.* 130, 621–633. doi: 10.1007/s00122-016-2849-1

Mason, A. S., Rousseau-Gueutin, M., Morice, J., Bayer, P. E., Besharat, N., Cousin, A., et al. (2016). Centromere locations in brassica a and c genomes revealed through half-tetrad analysis. *Genetics* 202, 513–523. doi: 10.1534/genetics.115.183210

Mason, A. S., and Wendel, J. F. (2020). Homoeologous exchanges, segmental allopolyploidy, and polyploid genome evolution. *Front. Genet.* 11. doi: 10.3389/fgene.2020.01014

Mayrose, I., Zhan, S. H., Rothfels, C. J., Arrigo, N., Barker, M. S., Rieseberg, L. H., et al. (2015). Methods for studying polyploid diversification and the dead end hypothesis: a reply to soltis et al., (2014). *New Phytol.* 206, 27–35. doi: 10.1111/nph.13192

Mendel, G. J. (1866). Versuche über pflanzenhybriden. *Verhandlungen Des. Naturforschenden Vereines Brünn* 4, 3–47.

Mercé, C., Bayer, P. E., Tay Fernandez, C., Batley, J., and Edwards, D. (2020). Induced methylation in plants as a crop improvement tool: Progress and perspectives. *Agronomy* 10, 1484. doi: 10.3390/agronomy10101484

Monroe, J. G., Srikant, T., Carbonell-Bejerano, P., Becker, C., Lensink, M., Exposito-Alonso, M., et al. (2022). Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature* 602, 101–105. doi: 10.1038/s41586-021-04269-6

Naish, M., Alonge, M., Wlodzimierz, P., Tock, A. J., Abramson, B. W., Schmücker, A., et al. (2021). The genetic and epigenetic landscape of the *Arabidopsis* centromeres. *Science* 374, eabi7489. doi: 10.1126/science.abi7489

Ni, P., Huang, N., Nie, F., Zhang, J., Zhang, Z., Wu, B., et al. (2021). Genome-wide detection of cytosine methylations in plant from Nanopore data using deep learning. *Nat. Commun.* 12, 5976. doi: 10.1038/s41467-021-26278-9

O'Brien, K. P., Remm, M., and Sonnhammer, E. L. L. (2005). Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res.* 33, D476–D480. doi: 10.1093/nar/gki107

Omony, J., Nussbaumer, T., and Gutzat, R. (2020). DNA Methylation analysis in plants: review of computational tools and future perspectives. *Brief Bioinform.* 21, 906–918. doi: 10.1093/bib/bbz039

Orantes-Bonilla, M., Wang, H., Lee, H., Golicz, A. A., Hu, D., Li, W., et al. (2022a). Transgressive and parental dominant gene expression and cytosine methylation during seed development in *Brassica napus* hybrids. *bioRxiv* 497715. doi: 10.1101/2022.09.05.506610

Orantes-Bonilla, M., Makhoul, M., Lee, H., Chawla, H. S., Vollrath, P., Langstroff, A., et al. (2022b). Frequent spontaneous structural rearrangements promote rapid genome diversification in a *Brassica napus* F1 generation. *bioRxiv*.. doi: 10.1101/2022.06.27.497715

Pelé, A., Rousseau-Gueutin, M., and Chèvre, A.-M. (2018). Speciation success of polyploid plants closely relates to the regulation of meiotic recombination. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00907

Pereira, R., and Leitão, J. M. (2021). A non-rogue mutant line induced by ENU mutagenesis in paramutated rogue peas (Pisum sativum l.) is still sensitive to the rogue paramutation. *Genes (Basel)* 12, 1680. doi: 10.3390/genes12111680

Pucker, B., Holtgräwe, D., Rosleff Sörensen, T., Stracke, R., Viehöver, P., and Weisshaar, B. (2016). A *de novo* genome sequence assembly of the *Arabidopsis thaliana* accession Niederzenz-1 displays presence/absence variation and strong synteny. *PLoS One* 11, e0164321. doi: 10.1371/journal.pone.0164321

Quezada-Martinez, D., Zou, J., Zhang, W., Meng, J., Batley, J., and Mason, A. S. (2022). Allele segregation analysis of F1 hybrids between independent brassica allohexaploid lineages. *Chromosoma* 131, 147–161. doi: 10.1007/s00412-022-00774-3

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033

Raman, H., Raman, R., Qiu, Y., Yadav, A. S., Sureshkumar, S., Borg, L., et al. (2019). GWAS hints at pleiotropic roles for FLOWERING LOCUS T in flowering time and yield-related traits in canola. *BMC Genomics* 6 20, 1–18 . doi: 10.1186/s12864-019-5964-y

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26. doi: 10.1038/nbt.1754

Rousseau-Gueutin, M., Belser, C., Da Silva, C., Richard, G., Istace, B., Cruaud, C., et al. (2020). Long-read assembly of the *Brassica napus* reference genome darmor-bzh. *Gigascience* 9, giaa137. doi: 10.1093/gigascience/giaa137

Samans, B., Chalhoub, B., and Snowdon, R. J. (2017). Surviving a genome collision: Genomic signatures of allopolyploidization in the recent crop species *Brassica napus*. *Plant Genome* 10, plantgenome2017.02.0013. doi: 10.3835/plantgenome2017.02.0013

Schiessl, S., Huettel, B., Kuehn, D., Reinhardt, R., and Snowdon, R. (2017). Post-polyploidisation morphotype diversification associates with gene copy number variation. *Sci. Rep.* 7, 41845. doi: 10.1038/srep41845

Schiessl, S.-V., Katche, E., Ihien, E., Chawla, H. S., and Mason, A. S. (2019b). The role of genomic structural variation in the genetic improvement of polyploid crops. *Crop J.* 7, 127–140. doi: 10.1016/j.cj.2018.07.006

Schiessl, S. V., Quezada-Martinez, D., Tebartz, E., Snowdon, R. J., and Qian, L. (2019a). The vernalisation regulator FLOWERING LOCUS c is differentially expressed in biennial and annual *Brassica napus*. *Sci. Rep.* 9, 14911. doi: 10.1038/s41598-019-51212-x

Schoen, D. J., and Schultz, S. T. (2019). Somatic mutation and evolution in plants. *Annu. Rev. Ecol. Evol. Syst.* 50, 49–73. doi: 10.1146/annurev-ecolsys-110218-024955

Sedeek, K. E. M., Mahas, A., and Mahfouz, M. (2019). Plant genome engineering for targeted improvement of crop traits. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00114

Sedlazeck, F. J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A., et al. (2018). Accurate detection of complex structural variations using single-molecule sequencing. *Nat. Methods* 15, 461–468. doi: 10.1038/s41592-018-0001-7

Serrato-Capuchina, A., and Matute, D. R. (2018). The role of transposable elements in speciation. *Genes (Basel)* 9, 254. doi: 10.3390/genes9050254

Shen, Y., Sun, S., Hua, S., Shen, E., Ye, C.-Y., Cai, D., et al. (2017). Analysis of transcriptional and epigenetic changes in hybrid vigor of allopolyploid *Brassica napus* uncovers key roles for small RNAs. *Plant J.* 91, 874–893. doi: 10.1111/tpj.13605

Shin, H., Park, J. E., Park, H. R., Choi, W. L., Yu, S. H., Koh, W., et al. (2022). Admixture of divergent genomes facilitates hybridization across species in the family brassicaceae. *New Phytol* 235, 743–758. doi: 10.1111/nph.18155

Singer, S. D., Laurie, J. D., Bilichak, A., Kumar, S., and Singh, J. (2021). Genetic variation and unintended risk in the context of old and new breeding techniques. *Crit. Rev. Plant Sci.* 40, 68–108. doi: 10.1080/07352689.2021.1883826

Smit, A. F. A., and Hubley, R. (2008) *RepeatModeler open-1.0*. Available at: http://www.repeatmasker.org (Accessed 2018).

Soltis, D. E., Segovia-Salcedo, M. C., Jordon-Thaden, I., Majure, L., Miles, N. M., Mavrodiev, E. V., et al. (2014). Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of mayrose et al. *New Phytol.* 202, 1105–1117. doi: 10.1111/nph.12756

Song, J.-M., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., et al. (2020). Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nat. Plants* 6, 34–45. doi: 10.1038/s41477-019-0577-7

Song, K., Lu, P., Tang, K., and Osborn, T. C. (1995). Rapid genome change in synthetic polyploids of *brassica* and its implications for polyploid evolution. *Proc. Natl. Acad. Sci. U.S.A.* 92, 7719–7723. doi: 10.1073/pnas.92.17.7719

Stein, A., Coriton, O., Rousseau-Gueutin, M., Samans, B., Schiessl, S. V., Obermeier, C., et al. (2017). Mapping of homoeologous chromosome exchanges influencing quantitative trait variation in *Brassica napus*. *Plant Biotechnol. J.* 15, 1478–1489. doi: 10.1111/pbi.12732

Stoiber, M., Quick, J., Egan, R., Eun Lee, J., Celniker, S., Neely, R. K., et al. (2016). De novo identification of DNA modifications enabled by genome-guided nanopore signal processing. *bioRxiv*. 094672. doi: 10.1101/094672

Suay, L., Zhang, D., Eber, F., Jouy, H., Lodé, M., Huteau, V., et al. (2014). Crossover rate between homologous chromosomes and interference are regulated by the addition of specific unpaired chromosomes in *Brassica*. *New Phytol.* 201, 645–656. doi: 10.1111/nph.12534

Szadkowski, E., Eber, F., Huteau, V., Lodé, M., Huneau, C., Belcram, H., et al. (2010). The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phytol.* 186, 102–112. doi: 10.1111/j.1469-8137.2010.03182.x

Tao, Y., Luo, H., Xu, J., Cruickshank, A., Zhao, X., Teng, F., et al. (2021). Extensive variation within the pan-genome of cultivated and wild sorghum. *Nat. Plants* 7, 766–773. doi: 10.1038/s41477-021-00925-x

van de Peer, Y., Mizrachi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. *Nat. Rev. Genet.* 18, 411–424. doi: 10.1038/nrg.2017.26

Varshney, R. K., Roorkiwal, M., Sun, S., Bajaj, P., Chitikineni, A., Thudi, M., et al. (2021). A chickpea genetic variation map based on the sequencing of 3,366 genomes. *Nature* 599, 622–627. doi: 10.1038/s41586-021-04066-1

Vicient, C. M., and Casacuberta, J. M. (2017). Impact of transposable elements on polyploid plant genomes. *Ann. Bot.* 120, 195–207. doi: 10.1093/aob/mcx078

Vollrath, P., Chawla, H. S., Alnajar, D., Gabur, I., Lee, H., Weber, S., et al. (2021b). Dissection of quantitative blackleg resistance reveals novel variants of resistance gene Rlm9 in elite *Brassica napus*. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.749491

Vollrath, P., Chawla, H. S., Schiessl, S. V., Gabur, I., Lee, H., Snowdon, R. J., et al. (2021a). A novel deletion in FLOWERING LOCUS T modulates flowering time in winter oilseed rape. *Theor. Appl. Genet.* 134, 1217–1231. doi: 10.1007/s00122-021-03768-4

Wang, L., Ji, Y., Hu, Y., Hu, H., Jia, X., Jiang, M., et al. (2019). The architecture of intra-organism mutation rate variation in plants. *PloS Biol.* 17, e3000191. doi: 10.1371/journal.pbio.3000191

Wang, Y., van Rengs, W. M. J., Zaidan, M. W. A. M., and Underwood, C. J. (2021). Meiosis in crops: from genes to genomes. *J. Exp. Bot.* 72, 6091–6109. doi: 10.1093/jxb/erab217

Wang, Z., Wu, X., Wu, Z., An, H., Yi, B., Wen, J., et al. (2018). Genome-wide DNA methylation comparison between *Brassica napus* genic Male sterile line and restorer line. *Int. J. Mol. Sci.* 19, 2689. doi: 10.3390/ijms19092689

Xiong, Z., Gaeta, R. T., and Pires, J. C. (2011). Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7908–7913. doi: 10.1073/pnas.1014138108

Yang, B., Jiang, Y., Rahman, M. H., Deyholos, M. K., and Kav, N. N. (2009). Identification and expression analysis of WRKY transcription factor genes in canola (*Brassica napus* l.) in response to fungal pathogens and hormone treatments. *BMC Plant Biol.* 9, 68. doi: 10.1186/1471-2229-9-68

Yuan, Y., Bayer, P. E., Batley, J., and Edwards, D. (2021). Current status of structural variation studies in plants. *Plant Biotechnol. J.* 19, 2153–2163. doi: 10.1111/pbi.13646

Yu, E., Fan, C., Yang, Q., Li, X., Wan, B., Dong, Y., et al. (2014). Identification of heat responsive genes in *Brassica napus* siliques at the seed-filling stage through transcriptional profiling. *PLoS One* 9, e101914. doi: 10.1371/journal.pone.0101914

Yu, X., Zhao, Z., Zheng, X., Zhou, J., Kong, W., Wang, P., et al. (2018). A selfish genetic element confers non-Mendelian inheritance in rice. *Science* 360, 1130–1132. doi: 10.1126/science.aar4279

Zanini, S. F., Bayer, P. E., Wells, R., Snowdon, R. J., Batley, J., Varshney, R. K., et al. (2022). Pangenomics in crop improvement-from coding structural variations to finding regulatory variants with pangenome graphs. *Plant Genome* 15, e20177. doi: 10.1002/tpg2.20177

Zhang, L., Cai, X., Wu, J., Liu, M., Grob, S., Cheng, F., et al. (2018). Improved *Brassica rapa* reference genome by single-molecule sequencing and chromosome conformation capture technologies. *Hortic. Res.* 5, 50. doi: 10.1038/s41438-018-0071-9

Zou, J., Hu, D., Mason, A. S., Shen, X., Wang, X., Wang, N., et al. (2018). Genetic changes in a novel breeding population of *Brassica napus* synthesized from hundreds of crosses between *b. rapa* and *b. carinata*. *Plant Biotechnol. J.* 16, 507–519. doi: 10.1111/pbi.12791

3. Transgressive and parental dominant gene expression and cytosine methylation during seed development in *Brassica napus* hybrids

Mauricio Orantes-Bonilla, Hao Wang, HueyTyng Lee, Agnieszka A. Golicz, Dandan Hu, Wenwen Li, Jun Zou and Rod J. Snowdon

Theoretical and Applied Genetics 136:113 (2023)

doi: 10.1007/s00122-023-04345-7

Theoretical and Applied Genetics (2023) 136:113
https://doi.org/10.1007/s00122-023-04345-7

ORIGINAL ARTICLE

Transgressive and parental dominant gene expression and cytosine methylation during seed development in *Brassica napus* hybrids

Mauricio Orantes-Bonilla[1] · Hao Wang[2] · Huey Tyng Lee[1] · Agnieszka A. Golicz[1] · Dandan Hu[2] · Wenwen Li[2] · Jun Zou[2] · Rod J. Snowdon[1]

Received: 3 November 2022 / Accepted: 12 March 2023 / Published online: 18 April 2023
© The Author(s) 2023

Abstract

Key message **Transcriptomic and epigenomic profiling of gene expression and small RNAs during seed and seedling development reveals expression and methylation dominance levels with implications on early stage heterosis in oilseed rape.**

Abstract The enhanced performance of hybrids through heterosis remains a key aspect in plant breeding; however, the underlying mechanisms are still not fully elucidated. To investigate the potential role of transcriptomic and epigenomic patterns in early expression of hybrid vigor, we investigated gene expression, small RNA abundance and genome-wide methylation in hybrids from two distant *Brassica napus* ecotypes during seed and seedling developmental stages using next-generation sequencing. A total of 31117, 344, 36229 and 7399 differentially expressed genes, microRNAs, small interfering RNAs and differentially methylated regions were identified, respectively. Approximately 70% of the differentially expressed or methylated features displayed parental dominance levels where the hybrid followed the same patterns as the parents. Via gene ontology enrichment and microRNA-target association analyses during seed development, we found copies of reproductive, developmental and meiotic genes with transgressive and paternal dominance patterns. Interestingly, maternal dominance was more prominent in hypermethylated and downregulated features during seed formation, contrasting to the general maternal gamete demethylation reported during gametogenesis in angiosperms. Associations between methylation and gene expression allowed identification of putative epialleles with diverse pivotal biological functions during seed formation. Furthermore, most differentially methylated regions, differentially expressed siRNAs and transposable elements were in regions that flanked genes without differential expression. This suggests that differential expression and methylation of epigenomic features may help maintain expression of pivotal genes in a hybrid context. Differential expression and methylation patterns during seed formation in an F_1 hybrid provide novel insights into genes and mechanisms with potential roles in early heterosis.

Introduction

Heterosis refers to the enhanced performance observed in F_1 hybrids derived from two genetically distant, homozygous parents. The improved performance of hybrids compared to their inbred parents was described as early as the late nineteenth century by Charles Darwin during his studies on maize and other plants (Darwin 1876). Since Shull first coined the term "heterosis" in 1914 (Shull 1914; 1948), this phenomenon has been applied in plant breeding to develop hybrids that outperform their inbred parents. However, despite the success of hybrid breeding in many major crops with selfing/outcrossing mating systems, for example maize, sunflower, tomato, sugarbeet and oilseed rape (Steeg et al. 2022), the mechanisms driving heterosis have yet to be fully elucidated. Several factors contributing to heterosis

Communicated by Yiqun Weng.

Mauricio Orantes-Bonilla and Hao Wang have contributed equally to this work.

✉ Rod J. Snowdon
Rod.Snowdon@agrar.uni-giessen.de

[1] Department of Plant Breeding, Land Use and Nutrition, IFZ Research Centre for Biosystems, Justus Liebig University, Giessen, Germany

[2] National Key Laboratory of Crop Genetic Improvement, College of Plant Science & Technology, Huazhong Agricultural University, Wuhan, People's Republic of China

have been proposed: For extensive reviews see for example Wu et al. (2021) and Mackay et al. (2021). Classical quantitative genetics remains core to understanding heterosis as a product of allele interactions through dominance, overdominance, or epistasis (Fujimoto et al. 2018). Nevertheless, through recent research, the cumulative understanding of molecular biology features has raised the question of how non-genetic and non-genomic features are also associated with heterotic patterns and to which extent. Recent studies suggest that gene networks, allele bias, epigenomic and transcriptomic factors play a key role in heterosis (Wu et al. 2021; Yu et al. 2021).

Brassica napus (oilseed rape, canola; AACC, $2n = 38$) is not only an important crop where hybrid breeding has been implemented successfully, but is also a model crop of research interest due to its polyploid nature and phylogenetic proximity to *Arabidopsis thaliana*. Oilseed rape is the second most widely grown oilseed crop and has been the third most important food oil source worldwide in the last decade (FAO 2022). The economic importance of hybrid breeding in oilseed rape is evident from the number and relevance of hybrid varieties in key producing countries like Canada, China and Germany. It is estimated that in 2015–2016, at least 80% of oilseed rape grown in China were hybrid varieties (Bonjean et al. 2016), whereas in Canada, the world's largest producer of spring-type canola, herbicide-tolerant hybrid varieties have contributed to significant yield increases during the past decade (Malla and Brewin 2019; FAO 2022). The highest worldwide yields in winter-type oilseed rape in the past ten years were recorded in Germany (FAO 2022), where the percentage of hybrid cultivars registered in the German National List increased from 74% in 2016 to more than 90% in 2022 (Friedt et al. 2018; BSA 2022). These figures highlight the relevance of heterosis in current oilseed rape/canola breeding worldwide. The economic importance increases the need for a more refined understanding of the underlying molecular mechanisms behind heterosis in *B. napus*.

Technological advances in transcriptomic and epigenomic profiling in recent decades have increased awareness of regulatory and epigenetic factors in crop improvement and hybrid breeding (Scossa et al. 2021; Yang et al. 2021). "Omics" technologies not only help to describe and expand genetic diversity in crop species (Louwaars 2018), but can also contribute to elucidating the role of regulatory and non-coding features in plants (Zanini et al. 2022). Transcriptomic and epigenomic features have been widely used to determine molecular and biological functions associated with improved performance in plant hybrids (Yu et al. 2021). For instance, RNA sequencing (RNA-Seq) data developed through microarrays and next-generation sequencing (NGS) has been used to find differentially expressed genes (DEGs) linked to heterosis during diverse growth stages (Wang et al. 2017a; Zhu et al. 2020). Small RNAs (sRNAs) derived from endogenous genomic loci or exogeneous sources are known to regulate various functions and responses in plants. Classification and characterization of microRNAs (miRNAs) and small interfering RNAs (siRNAs) provide valuable information to investigate regulatory factors involved in modulation of gene and trait expression (Griffiths-Jones et al. 2006; Lunardon et al. 2020). For example, sRNAs have been associated with changes in performance in maize, rice and wheat (Zhang et al. 2014a; Li et al. 2014; Seifert et al. 2018a), while epigenomic features, including chromatin interaction, histone modification and DNA methylation, can cause phenotypical changes without alterations in DNA sequences (Fitz-James and Cavalli 2022). Genome-wide methylation differences in various plant species have been associated with phenotypic consequences (Muyle et al. 2022) and linked to heterosis (Kawanabe et al. 2016; Lauss et al. 2018).

Differential gene expression studies in *B.napus* revealed key genes regulating flowering time, disease resistance and abiotic stress (Wu et al. 2016; Wang et al. 2017b; Jian et al. 2019), while small RNA profiling identified microRNA and siRNA sequences associated with pathogen response, abiotic stress and lipid metabolism in oilseed rape (Wang et al. 2017c; Jian et al. 2018; Martinez Palacios et al. 2019; Regmi et al. 2021). Furthermore, DNA methylation patterns were found to contribute to heat response, DNA repair and fertility in *B.napus* (Li et al. 2016; Ran et al. 2016; Wang et al. 2018; Yin et al. 2021).

Nevertheless, few studies have integrated multiple omics strategies to obtain a detailed scenario of expression and methylation patterns in oilseed rape (Shen et al. 2017; Wang et al. 2018). Interestingly, Shen et al. (2017) found specific expression and methylation patterns associated with heterosis in a commercial *B. napus* hybrid cultivar. Enhanced performance due to heterosis has been mostly evaluated at the genomic level and explained through allele interactions (Fujimoto et al. 2018) and introgressions of genomic regions between genetically and genomically distant parents (Hu et al. 2021a; Quezada-Martinez et al. 2021). Nevertheless, the transcriptomic and epigenomic networks involved in heterosis have not been fully elucidated, and the potential to include information on coding and non-coding features in hybrid breeding has been barely explored.

Comprehensive studies in maize and Arabidopsis demonstrated that heterosis can be observed in various developmental stages (van Hulten et al. 2018; Zhou et al. 2019). Heterosis during seed development can contribute directly to grain yield, seed biomass, germination and early vigor (Hochholdinger and Hoecker 2007; Jahnke et al. 2010). Since seed formation is characterized by the merging of parental genomes, parent-specific epigenomic effects and genomic imprinting (Thiemann et al. 2009; Castillo-Bravo

et al. 2022), it is an ideal stage for transcriptomic and epigenomic assessments in relation to heterosis. RNA-Seq and methylation-based studies have dissected putative heterotic loci in embryo and seed developmental stages in hybrids of Arabidopsis (Meyer et al. 2012; Kawanabe et al. 2016; Alonso-Peral et al. 2017; Chen et al. 2022). In *A. thaliana* and maize, early heterosis was associated with increases in cell size and number, seed yield and biomass (Jahnke et al. 2010; Wang et al. 2017a; Zhu et al. 2020). Groszmann et al. (2014) found that the maternal genotype was the major determinant of heterosis at early developmental stages in *A. thaliana*. Seed development is also well characterized for enriched epigenomic mechanisms through methylation and transcriptomic regulation, with pollen cells being hypermethylated and ovule cells demethylated in most plants (Batista and Köhler 2020; Montgomery and Berger 2021). Parental dominance effects are attributed with a key role during seed formation through diverging gamete methylation patterns (Weigel and Colot 2012; Lauss et al. 2018). Moreover, the merging of parental genomes during embryogenesis can cause a genomic shock that can further alter the hybrid transcriptome (Bird et al. 2018).

Diverse studies in parent–offspring trios have compared parental dominant and transgressive gene expression patterns via expression level dominance (ELD) analyses in polyploids including *B. napus* (Yoo et al. 2013; Wu et al. 2018; Li et al. 2020) to elucidate the parental genotype effects on gene expression in interspecific hybrids between the diploid species progenitors. In the present study, we analyze transcriptomic and epigenomic differences during seed and seedling development in homozygous parental lines and their F_1 hybrid from a cross between the winter-type *B.napus* accession Express 617 (Lee et al. 2020) and the semi-winter accession *B. napus* G3D001 (Zou et al. 2018).

Preliminary observations by Hu et al. (2021b) showed significant heterosis in hybrids from Express 617 and newly resynthesized oilseed rape lines, including the Express 617 × G3D001 hybrid (encoded in that publication as genotype number T4_N22) under environmental conditions in China. Numerous studies during the last decades have also shown the heterotic advantages of crossing genetically distant oilseed rape varieties (Qian et al. 2009; Basunanda et al. 2010; Girke et al. 2012; Hu et al. 2021a). Hence, determining differentially expressed and methylated heterotic features between distant germplasm sources can potentially improve our understanding of the molecular mechanisms of heterosis. The overall aim of the present article is to provide an atlas of transcriptomic and epigenomic features associated with heterosis and contribute to dissecting relevant multiomics loci in oilseed rape as a model crop. For this purpose, mRNA, sRNA and whole-genome bisulfite sequencing were carried in the two parental inbreds and their F_1 hybrid. Differential features were identified and classified by their respective expression or methylation dominance levels to detect parental and hybrid-specific patterns associated with early developmental stages. Gene ontology enrichment (GO) and integration of omics features were performed to find putative interactions between the detected features and consequently evaluate their epigenomic and transcriptomic impact on early heterosis.

Material and methods

Experimental design and growing conditions

Seeds from homozygous, advanced inbred lines of winter-type oilseed rape Express 617 (maternal line), semi-winter semi-synthetic oilseed rape G3D001 (paternal line) and their F_1 hybrid offspring were planted in the 2020–2021 growing season at Huazhong Agricultural University of Wuhan field station. The third youngest leaf from each genotype were sampled from seedlings having six unfolded leaves (BBCH16) at 10:00 am under liquid nitrogen. Flower buds with similar sizes were selected on the fifth day after reaching full flowering (BBCH65) to perform manual selfing on a defined day in all genotypes along with crosses between Express 617 (female pollen recipient) and G3D001 (male pollen donor). The newly generated F_1 crosses were employed to analyze the transcriptomic and epigenomic differences during seed formation between ovules pollinated from selfed F_1 plants and those pollinated by outcrossing from Express 617 and G3D001 (referred to hereinafter as F_0). Pollinated ovules were excised with forceps 15 and 30 days after pollination (DAP) at 10:00 am and immediately transferred to liquid nitrogen. Biological replicates consisted of pooled samples from the third youngest leaf from seven individual plants for leaf samples, and from four pollinated ovules from four different plants. The sampled tissue was aliquoted and used for all sequencing types described in this study. Lastly, three independent biological replicates were used for messenger RNA (mRNA) and small RNA expression experiments. Due to low material availability for some samples, only two biological replicates have been used for the methylation studies. Phenotypic measurements were recorded for plant height and dry seed weight for each pooled biological replicate. Significant differences between genotypes were assessed using a one-way analysis of variance (ANOVA) followed by a Tukey test ($p < 0.05$). The experimental designs are summarized in Figs. 1 and S1.

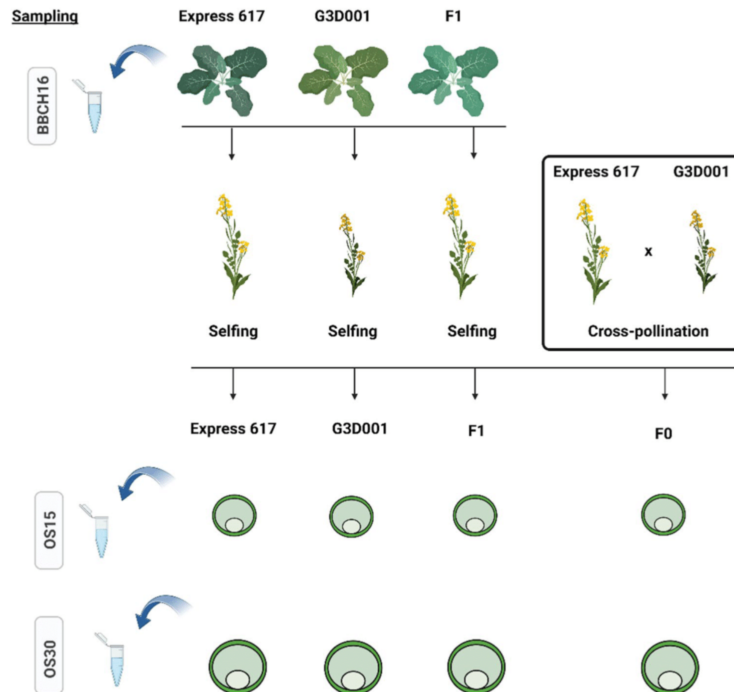

Fig. 1 Transcriptomic and epigenomic experimental design. Leaves samples were taken at the six-leaves stages (BBCH16) from each biological replicate. Homozygous inbred plants of Express 617 and G3D001 along with their heterozygous F_1 hybrid were self-pollinated to generate selfed ovules from each genotype. The two inbreed parents were also crossed during the experiment to develop cross-pollinated ovules (F_0). Pollinated ovules were sampled from each biological replicate and sequenced at 15 (OS15) and 30 (OS30) days after pollination

mRNA, small RNA and whole-genome bisulfite sequencing

mRNA was extracted using TRIzol™ Reagent (Thermo Fisher). A total of 0.5 µg of total RNA per biological replicate were used for preparing 150 bp paired-end (PE) read libraries using the NEBNext® Ultra™ II RNA Library Prep Kit (New England Biolabs, Inc.). Small RNA was extracted using a Plant miRNA kit (Omega Bio-tek Inc.). One microgram of total RNA per biological replicate was employed for the construction of 50 bp single-end (SE) reads using NEBNext® Multiplex Small RNA Library Prep Set for Illumina™ (New England Biolabs, Inc.). Lastly, 2.5 µg of CTAB extracted-DNA per biological replicate were first treated with sodium-bisulfite using the Zymo EZ DNA Methylation-Lightning™ Kit (Zymo Research Corp.) and then built into 150 bp PE read libraries with the TruSeq Nano DNA LT Sample Prep Kit (Illumina Inc.) for whole-genome bisulfite sequencing (WGBS). All libraries were sequenced using an Illumina NovaSeq 6000 platform (Illumina Inc.). Read quality was evaluated with FastQC v.0.11.9 (Andrews. 2010) and multiqc v.1.9 (Ewels et al. 2016) for all sequencing types. Principal component analyses (PCA) were carried for all libraries per developmental stage using the plotPCA function from DESEQ2 for RNA data (Love et al. 2014) and the *prcomp* R function for WGBS libraries.

mRNA and sRNA alignments

mRNA libraries were first filtered by selecting reads with an exact length of 150 bp, minimum base quality phred value of 5, no unqualified bases and less than 15% N bases using fastp v.0.23.1 *-q 5 -u 0 -n 15 -l 150* settings (Chen et al. 2018). Splice sites in the Express 617 reference genome assembly (Lee et al. 2020) were identified by first converting the gene annotation file format (Express617_v1_gene.gff3; MD5: cf26ec54823f348a0e23f027dc386a16) from a general feature format v.3 (GFF3) to a general transfer format (GTF) using the *agat_convert_sp_gff2gtf.pl* script from AGAT v.0.5.0 (Dainat 2019). The output was then employed to find splice sites with the *hisat2_extract_splice_sites.py* script

from HISAT2 (Kim et al. 2019). An index from the same Express 617 reference was built with *hisat2-build* function, and libraries were then aligned with HISAT2 using the *sensitive* preset and the *known-splicesite-infile* setting with the *hisat2_extract_splice_sites.py* previously generated file as input. Alignments were sorted and converted to a binary alignment map (BAM) format with samtools (Li et al. 2009) *view* and *sort* functions. The number of fragments in genes was counted with featureCounts 2.0.1 (Liao et al. 2014) using the AGAT GTF annotation file and the following settings *-p -B -C -Q 50 -t "exon" -g "gene_id"*, so that only read pairs having a minimum mapping quality of 50 and both reads aligned to the same strand and chromosome were counted. Genes without any counts in all genotypes were removed. Small RNA libraries were first filtered by removing reads shorter than 18 bp with seqtk v.1.3 (Li 2016). Subsequently, sRNA libraries were aligned against the Express 617 reference using ShortStack v.3.8.5 (Johnson et al. 2016). Only sRNA in which at least 80% of the primary reads had a length between 20–24 nucleotides, with less than 5 unpaired bases in secondary structure, and which were contained in predicted hairpin structures (i.e. only small RNAs clusters with *Y, N15, N14* or *N13* flags.) were considered as miRNA candidates. Small RNA sequences in which in which 80% of the primary reads had an exact length of 24 nucleotides and without miRNAs selection flags were regarded as putative siRNA based on similar thresholds employed by Lunardon et al. (2020). miRNA and siRNA clusters without any coverage in all biological samples were discarded prior to differential expression analysis.

Expression level dominance analysis

Differential mRNA, miRNA and siRNA expression patterns between the F_1 hybrid and its parents were assessed by comparing tissues within genotype trios in five tissues/stages: leaves from parental and F_1 plants at six-leaf stage (BBCH16); ovules 15 days after pollination from selfed parents and F_1 (OS15-F_1) or F_0 (OS15-F_0); and ovules 30 days after pollination from selfed parents and F_1 (OS30-F_1) or F_0 (OS30-F_0), respectively. Differentially expressed genes, differentially expressed miRNAs (DE-miRNAs) and differentially expressed siRNAs (DE-siRNAs) between genotypes for each stage were identified using the counts from each biological replicate with DESEQ2 (Love et al. 2014) with a padj value threshold < 0.05. The DESEQ2 built-in *estimateSizeFactors* and *counts* functions were used to extract the normalized counts which were then used for expression level dominance analyses. Briefly, student's t test ($p < 0.05$) from normalized counts of DEGs and DE-miRNAs identified in DESEQ2 were run between all genotypes for each comparison stage and gene. Tukey tests ($p < 0.05$) were then carried to rank each genotype by expression level. Finally, the resulting patterns were divided based on Yoo et al. (2013) as additive (I, XII), paternal dominant (II, XI), maternal dominant (IV, IX) and transgressively up- (III, VII, X) or downregulated (V, VI, VIII). Gene expression heatmaps were generated with idep93 (Ge et al. 2018) using correlation distances and average linkages, and differentially expressed genes or sRNAs shared between all stages were detected using the *Venn Diagrams* tool (VIB-UGent 2021). In addition, the percentages of upregulated and downregulated DEGs from all genes per subgenome, genotype and stage were calculated to evaluate subgenomic expression bias and normalized gene transcript values were summarized as heatmap values for easier comparison.

Gene ontology enrichment

Gene models in the Express 617 reference assembly were functionally annotated through synteny comparison against the Darmor v.4.1 genome (Chalhoub et al. 2014) with inparanoid v.4.2 (O'Brien et al. 2005) using bootstrap, a BLOSUM80 (BLOcks SUbstitution Matrix) and an initial cutoff score of 60. Inparalogs with a similarity score equal or greater than 70 were selected for each gene. Pairs with only one homolog and with the highest similarity score were kept. The homologs were used for GO enrichment of biological processes based on expression level dominance for each stage, as well in comparisons between the F_1 and F_0 genotypes, using ShinyGo v.0.76 (Ge et al. 2020) with a 0.05 false discovery rate (FDR) cutoff. Only biological functions with more than one gene per biological pathway and with at least two GO groups were selected.

DE-miRNA target prediction and mRNA interaction

Differentially expressed miRNAs sequences were extracted and used to predict their corresponding targets in Express 617 gene models using psRNATarget (Dai et al. 2018) with the version 2 scoring schema (Axtell 2013). Maximum unpaired energy (UPE) of 25 and a flank length between 13 to 17 nucleotides in up/downstream region were set as target accessibility cutoffs. All possible targets for DE-miRNAs were reported since each miRNA can have multiple mRNA targets due to isomiRs formation. The DE-miRNAs were classified into putative miRNA families by blasting their sequences with BLAST (Altschul et al. 1990) against the mature miRNAs from the Brassicaceae family available at the miRBase sequence database release version 22.1 (Griffiths-Jones et al. 2006). Only the top five matches with the highest alignment scores and lowest expect values for each DE-miRNA were retained. Stem-loop sequences from the Brassicaceae family were used as BLAST targets when no mature miRNAs matches were found. Alternatively, if

no Brassicaceae matches were found, then mature miRNAs and stem-loop sequences from the Viridiplantae clade were employed. The expression patterns from miRNA targets that were DEGs were compared with their associated targeting DE-miRNA expression to evaluate possible interactions between miRNA and mRNA target. The DEGs target functions were estimated by comparing their coding sequences against the Araport v.11 *Arabidopsis thaliana* coding sequences model (Cheng et al. 2017) via BLAST. Only the hit with the lowest expect value and not greater than 1.0×10^{-4}, lowest identity percentage equal or above 90% and without gaps were selected.

Bisulfite sequencing alignment and methylation level dominance

Reads with a minimum base quality phred value of 5, unqualified base percent limit of 50 and less than 15% N bases were selected from WGBS libraries using fastp v.0.23.1 *-q 5 -u 50 -n 15* settings (Chen et al. 2018). TrimGalore (Krueger et al. 2021) was then employed for trimming 8 basepairs from both 5' and 3' ends for each library as recommended for TruSeq libraries in the Bismark documentation. The Express 617 reference genome was bisulfite converted and indexed with Bismark v.0.23 (Krueger and Andrews 2011) *bismark_genome_preparation* tool. Filtered reads for each biological replicate were aligned to the bisulfite converted genome using *bismark* under default settings. Duplicates were afterward removed with *deduplicate_bismark* and methylated cytosines (mC) were extracted using *bismark_methylation_extractor* while ignoring the first 2 basepairs from both 5' and 3' ends for both reads of a pair. Every mC in a CpG, CHG or CHH methylation context was selected and converted to a browser extensible data (BED) format with *bismark2bedGrap*h using the *–cutoff 3 –CX* –and *–scaffolds* settings to select all nucleotides in which the methylation state was reported at least thrice.

The coverage for each mC in every methylation context per biological replicate was calculated with the *coverage2cytosine* from the Bismark package. The mC coverage in assigned chromosomes was then used as input for DMRCaller v. 1.22.0 (Catoni et al. 2018) to detect differentially methylated regions (DMRs). Each genotype within a trio was compared to each other using the *computeDMRs* function in 1000 bp bins with the *bins* method and the following settings: score test, a 0.01 p value threshold, and minimum cytosine count, methylation proportion difference and gap between bins of 4, 0.4 and 0 accordingly. The DMR methylation levels (i.e. the number of reads supporting methylation) were extracted from DMR output files and student's t test ($p < 0.05$) was run between all genotypes for each stage and DMR. Tukey tests ($p < 0.05$) were then used to rank the methylation within DMRs between genotypes and classified them by methylation level dominance (MLD) following the same categorization employed for ELD by Yoo et al. (2013) as additive (I, XII), paternal dominant (II, XI), maternal dominant (IV, IX) and transgressively hyper- (III, VII, X) or hypomethylated (V, VI, VIII). Shared and unique DMR across all stages were found with the *Venn Diagrams* tools (VIB-UGent 2021). Moreover, heatmaps displaying the genome-wide methylation levels between biological replicates per methylation context in 100 kbp bins were generated using the *circlize* and *ComplexHeatmap* packages (Gu et al. 2014, 2016).

Cytosine methylation statistics and identification of methylated features

The number of mC nucleotides and the cytosine methylation level per 1 kbp bin (i.e., numbers of reads supporting cytosine methylation per bin) in each methylation context, genotype and stage were determined based on Bismark's *coverage2cytosine* generated files using bedtools *makewindows* and *intersect* functions (Quinlan and Hall 2010). In addition, DMRs were intersected with exons, introns, repeats and 1 kbp upstream promoter regions from Express 617 using bedtools *intersect* function. GO enrichment was analyzed for differentially expressed genes having DMRs for all stages and genotypes. If no enrichment was detected, then the most frequent biological functions found in Ensembl Biomart (Cunningham et al. 2022) *B. napus* reference (Chalhoub et al. 2014) were reported. Detected differentially expressed genes having an additive or dominant expression level dominance pattern were defined as putative genetic epialleles if their loci coincided with corresponding additive or dominant methylation patterns in DMRs. Furthermore, the correlation between gene expression and methylation in either the gene body or promoter were assessed with a Kendall τ test using the *cor.test* function in *R*, since the pre-evaluation of our gene expression data showed a non-Gaussian distribution, as also observed in other transcriptomic studies (Robinson et al. 2010; Di et al. 2011; Love et al. 2014; Church et al. 2019).

Heatmaps comparing the gene methylation in gene bodies and promoters and gene expression in transgressive DEGs were made with Heatmapper (Babicki et al. 2016) using Euclidean distances and average linkages to analyze the putative interaction between expression and methylation. Moreover, repeats in the Express 617 assembly were assigned to repeat families using RepeatModeler (Smit and Hubley 2008) and CpG islands were identified with *cpgplot* from the EMBOSS v.6.6.0 package (Rice et al. 2000). CpG islands were called if the GC% was equal or greater

than 50%, length greater than 200 bp and a minimum 0.6 observed to expected CpG dinucleotides ratio as described by Gardiner-Garden and Frommer (1987). Additionally, plots showing DEGs and methylation levels for each chromosome and stage, centromere loci and repeat density were made using the *circlize* package (Gu et al. 2014). Repeat density for each 1 kbp bin within each chromosome was calculated using bedtools while predicted Express 617 centromere loci were added based on Orantes-Bonilla et al. (2022).

Lastly, DMRs were intersected with DE-siRNAs, CpG islands and transposable elements (TEs) in 5 kbp upstream and downstream gene and DEG-flanking regions in assigned chromosomes using bedtools to evaluate putative interactions between differentially methylated features and gene expression during seed development. The threshold was selected based on previous work on transposable elements and genomic imprinting in *B.napus* by Rong et al. (2021) and the fact that the average distance between genes in assigned chromosomes of the Express 617 reference is approximately 7.5 kbp. Chi-square tests followed by an FDR post-hoc adjustment ($p < 0.05$) were performed to find significant associations between differentially methylated and non-methylated features and determine their respective distances to genes or DEGs across all stages.

Segmental expression assessment

Clustering of DEGs across chromosomal segments observed on *circlize* generated plots was further investigated. To assess the presence of expression clusters, segments that had more than 20 DEGs over a 500 kbp window were considered as putative differentially expressed segments. The threshold was selected on the basis that the Express 617 genome assembly has an average of 200 genes per 500 kbp and hence 20 genes would correspond to 10% of genes in the segment. The ratio of upregulated to downregulated DEGs per genotype and stage in each segment was calculated and normalized to *Z*-scores. Only segments showing clear differential patterns between genotypes based on *Z*-score heatmap clustering were retained. Such segments could either be a result of parental expression bias or due to commonly observed genomic rearrangements in allopolyploid *B. napus*. To investigate both possibilities, short read genomic sequence data from a G3D001 biological replicate was used for calling copy number variation (CNV) and investigating putative linkages between structural rearrangements and expression patterns. For this purpose, genomic DNA from a G3D001 ovule biological replicate taken 30 days after pollination was extracted using a CTAB protocol (Doyle and Doyle 1987). Paired-end libraries were built with KAPA HyperPlus Kit (KAPA Biosystems) and sequenced with an Illumina NovaSeq 6000 platform (Illumina Inc.). Read quality was evaluated with FastQC v.0.11.9 and libraries were afterward aligned with minimap2 (Li 2018) against the Express 617 genomic reference (Lee et al. 2020). Alignments with both forward and reverse reads properly mapped (flags 99, 163, 147 and 83) were selected with samtools *view* and used to calculate coverage across chromosomes using the *bamtobed* and *genomecov* functions from bedtools. The coverage was used as input in a previously described deletion-duplication pipeline (Stein et al. 2017), modified by excluding outliers if the depth was above 100 and by defining deletions and duplications as 25 kbp length segments that are one standard deviation above or below the mean coverage. Deletions and duplications were recorded in tab-separated files and intersected with differentially expressed segments using the *intersect* function in bedtools.

Results

Maternal dominant expression and methylation increases during seed development

Gene and small RNA expression as well as genome-wide methylation patterns from Express 617, G3D001 and their hybrid were compared during seed and seedling developmental stages. The parents were crossed during the experiment to evaluate developmental differences between selfed-F_1 plants and Express 617 × G3D001 pollinated ovules that would develop into F_1 plants (referred here as F_0). Figure 1 provides an overview of the plant materials, sampling tissues/timepoints and sample nomenclature and Fig. S1 summarizes the bioinformatic pipeline used. Next-generation sequencing yielded abundant coverage for each biological replicate (Tables S1, S2, S3). Approximately 6.8 Gbp of mRNA sequences per biological replicate were aligned against the Express 617 genome assembly using HISAT2 v. 2.2.1 splice site aware aligner (Kim et al. 2019), producing mean alignment rates of 98.2% (Table S1). In addition, an average of 31 million sRNA reads per biological replicate were used to find putative miRNA and siRNA sequences with ShortStack v.3.8.5 (Johnson et al. 2016). Overall, each sRNA cluster had an average coverage depth of 186 (Table S2). Moreover, whole-genome bisulfite treated reads having a 31 × genome coverage per biological replicate were aligned and processed with Bismark v.0.23 (Krueger and Andrews 2011), as reported in Table S3. PCA plots from mRNA and sRNA libraries showed an overall clustering of biological replicates (Fig. S2, S3, S4). PCA plots from methylation levels displayed grouping of biological replicates to a slightly lesser extend since methylation level is calculated in the whole genome unlike RNA and sRNA counts that are only derived from a specific set of features. Nonetheless, genome-wide methylation level heatmaps

suggested an overall agreement of methylation level between biological replicates for each stage and context (Fig. S5, S6, S7, S8, S9, S10).

All alignments were then employed to find features that were differentially expressed or differentially methylated between genotypes across all stages. In summary, a total of 31,117 DEGs, 344 DE-miRNAs, 36,229 DE-siRNAs and 7399 DMRs in both CpG and CHG methylation contexts were identified across all possible parents-hybrid comparisons per stage (Tables S4, S5, S6, S7, S8, S9). The detected features were evenly distributed across all chromosomes (Tables S10, S11, S12, S13, S14). Differential features were further classified by their expression and methylation level dominance (Fig. 2). More than 90% of the differentially expressed and methylated features corresponded to parental dominant (II, XI, IV, IX) and additivity (I, XII) models based on the dominance classification proposed by Yoo et al. (2013). Moreover, maternal dominance accounted for approximately 89%, 85%, 83% and 60% from all detected DEGs, DE-miRNAs, DE-siRNAS and DMRs in the F_0, respectively, whereas paternal dominance was more prevalent in the F_1-selfed offspring (Table S15). Furthermore, most maternal dominant DMRs in the F_0 were hypermethylated, whereas DEGs were downregulated. This contrasts to the expected female gamete demethylation observed in seed formation in other plants (Batista and Köhler 2020).

Transgressive upregulated features, in which the hybrid has a higher expression than the parents, were more frequent in seeds from selfed-F_1 plants compared to those from the recently formed F_0. Maternal dominance from Express 617 accounts for most of the DEG and DE-siRNAs patterns observed, suggesting a potential maternal relevance in seed development. Interestingly, no gene expression bias was found between the A and C subgenomes (Table S16, Figure S11, S12, S13, S14, S15); nevertheless, more upregulation was observed in the paternal line, while the maternal line displayed more downregulation during seed development. This contrasts with the expected gene silencing in the maternal genome that is attributed to maternal demethylation during seed formation. Moreover, a slightly higher number of

Feature	Stage	Additivity		G3D001 dominance		Express 617 dominance		Transgressive-down regulation/hypomethylation			Transgressive up regulation/hypermethylation		
		I	XII	II	XI	IV	IX	III	VII	X	V	VI	VIII
		E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G
DEG	BBCH16	25.19	8.68	9.01	20.34	15.05	18.75	0.08	0.77	0.19	0.13	0.18	1.62
	OS15-F1	9.79	6.44	9.59	18.16	13.20	24.58	0.89	3.64	0.26	0.49	1.53	11.43
	OS15-F0	0.55	0.08	0.09	1.04	46.34	50.77	0.09	0.13	0.00	0.00	0.50	0.40
	OS30-F1	8.75	5.44	5.49	11.71	20.68	37.93	0.81	1.73	0.06	0.28	2.32	4.81
	OS30-F0	7.37	2.18	1.93	3.21	32.68	48.52	0.34	0.53	0.01	0.05	1.38	1.81
DE-miRNA	BBCH16	13.85	9.23	9.23	27.69	9.23	20.00	0.00	9.23	0.00	1.54	0.00	0.00
	OS15-F1	5.38	5.38	8.60	23.66	18.28	23.66	0.00	0.00	0.00	3.23	0.00	11.83
	OS15-F0	1.96	0.00	0.00	2.94	42.16	50.00	0.00	0.00	0.00	0.00	1.96	0.98
	OS30-F1	14.74	1.99	2.39	8.37	23.90	45.02	0.40	0.40	0.00	0.00	0.40	2.39
	OS30-F0	14.12	1.15	0.76	3.44	33.21	45.04	0.38	0.38	0.00	0.00	0.00	1.53
DE-siRNA	BBCH16	16.30	7.54	8.83	30.81	15.39	19.82	0.05	0.56	0.12	0.02	0.12	0.44
	OS15-F1	6.69	8.67	16.20	25.78	11.78	17.36	0.03	0.26	0.01	2.50	2.46	8.25
	OS15-F0	1.06	0.16	0.09	1.30	44.77	49.02	0.10	0.11	0.00	0.00	2.16	1.23
	OS30-F1	12.58	2.58	3.58	13.20	23.07	40.56	0.27	0.76	0.05	0.16	0.60	2.60
	OS30-F0	12.91	1.46	2.12	7.73	28.63	44.42	0.19	0.50	0.04	0.08	0.48	1.46
DMR	BBCH16	42.94	26.46	5.51	13.31	11.11	0.00	0.00	0.00	0.03	0.06	0.39	0.19
	OS15-F1	49.10	9.16	13.21	9.33	17.74	0.00	0.00	0.00	0.03	0.49	0.55	0.39
	OS15-F0	5.08	2.29	0.00	0.15	46.47	15.04	0.03	0.00	0.00	0.03	29.37	1.55
	OS30-F1	39.17	14.23	15.77	8.24	21.13	0.00	0.00	0.00	0.00	0.49	0.52	0.44
	OS30-F0	21.99	14.88	2.21	0.69	49.12	8.99	0.00	0.00	0.00	0.03	2.03	0.06

Fig. 2 Percentages of differentially expressed genes (DEGs), differentially expressed miRNAs (DE-miRNAS), differentially expressed siRNAs (DE-siRNAS) and differentially methylated regions (DMRs) in CpG and CHG methylation contexts by expression level dominance (ELD) and methylation level dominance (MLD) patterns per stage. Increase and decrease in expression and methylation per patterrn are displayed by the dot-ended lines showing the relative expression or methylation levels for the parental genotypes Express 617 (*E*) and G3D001 (*G*) along with their F_1 hybrid (*F*). Differential expression and methylation are displayed for leaf samples at stage BBCH16 and for ovules at 15 (OS15) and 30 (OS30) days after pollination by selfing (F_1 ovules) or cross-pollination between the two parental lines (F_0 ovules). Percentages are displayed with colored backgrounds to represent high (red) or low (blue) abundance

differentially expressed features following maternal expression patterns were found in the F_0 than in the selfed-F_1; however, this might be due to the allele segregation in the selfed-F_1 plants that would lead to the maternal parent being heterozygote and putatively reducing the number of features with maternal dominant expression.

A total of 1565 DEGs, 12 DE-miRNAs, 1111 DE-siRNAs, 896 DMRs in CpG context and 650 DMRs in CHG context were consistently detected across all stages (Tables S17). Altogether, differential features present consistently in all stages and genotypes corresponded to 3% of all detected features, whereas differential features unique to a single stage compromised approximately 2% of all detected features (Table S17). Differential features with consistent dominance level patterns across sampling stages are presented in Fig. 3. Interestingly, features with consistent dominance patterns among stages tended to exhibit maternal dominance. These features were mostly shared between early and late pollinated ovule stages in the F_1 and F_0 (Table S18, S19, S20, S21), suggesting dominance of the maternal genotype during seed formation for both genotypes. Parental effects have been reported to play a role in heterosis in maize and Arabidopsis (Ma et al. 2018; Castillo-Bravo et al. 2022) and are further addressed in the discussion.

DEGs and differential miRNA expression regulate F_1 seed development

Gene ontology enrichment for biological processes was carried out for all stages based on their expression level dominance. Significant enrichment for pivotal biological functions such as amino acid and carbohydrate synthesis, stress response, photosynthesis, protein transport and DNA repair and replication were found in 15 and 30 days after pollination ovules (Table S22). No significant enrichment was identified in leaves during the seedling stage. Only transgressively upregulated genes in F_1 ovules after 15 days of pollination displayed terms associated with reproduction and meiosis (Fig. 4, Table S23, Figures S16-S20). Differential gene expression and gene ontology between the F_1 and F_0 at 15 days after pollination showed an increase in photosynthesis-related functions in the F_1 hybrid, whereas the F_0 showed increased accumulation of energy reserve compounds and cell mobilization (Table S22). GO terms linked to carbohydrate metabolism, photosynthesis, stress response and cell development have been linked to heterosis in maize, rice, sunflower and oilseed rape (Bao et al. 2005; Lai et al. 2006; Ma et al. 2018; Zhu et al. 2020). Phenotypic measurements recorded during the growing of the parents and the hybrid genotypes used in this study revealed that hybrids had a more robust plant architecture and higher dry seed weight than their parents (Table S24, Fig. 5), in concordance, with

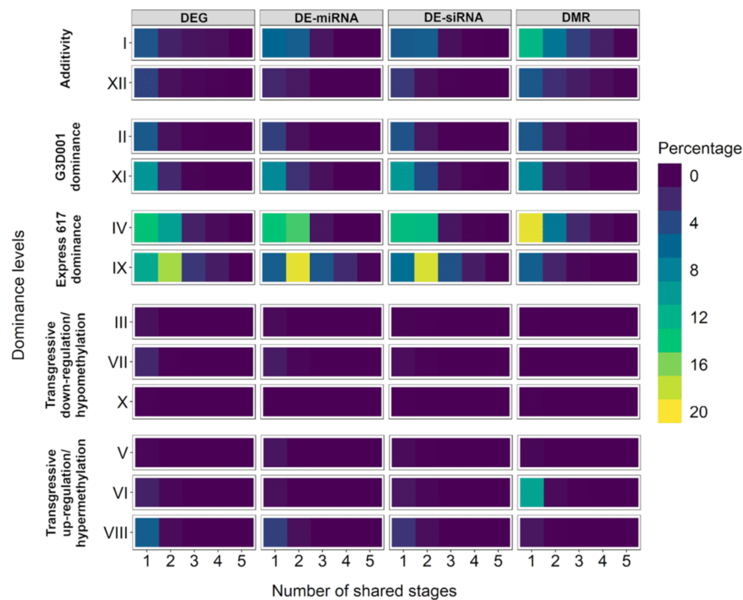

Fig. 3 Percentage of shared differential features between stages based on dominance level patterns displaying differentially expressed genes (DEGs), differentially expressed miRNAs (DE-miRNAs), differentially expressed siRNAs (DE-siRNAs) and differentially methylated regions (DMRs) in CpG and CHG contexts

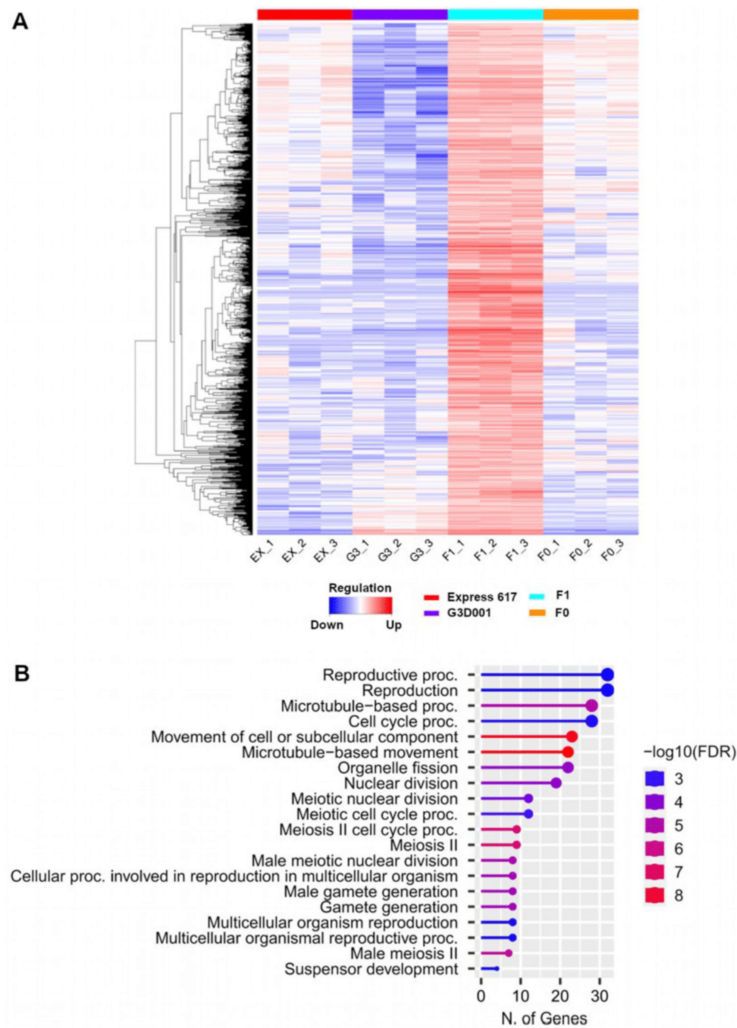

Fig. 4 (**a**) Gene expression heatmap and (**b**) gene ontology (GO) enrichment of biological processes from 15 days after pollination ovules with transgressive upregulation patterns in the *F*1

previous observations on this and other related hybrids (Hu et al. 2021b; Orantes-Bonilla et al. 2022).

Additionally, 51 putative mRNA targets from all DE-miRNAs were detected across all stages (Table S25). Interactions between DE-miRNAs and DEG mRNA targets are reported in Table 1 and Table S26. Most DE-miRNAs associated with DEG targets had downregulated expression in the parents and F_1 (ELD group IX) and were more abundant during the late seed developmental stage. Expression from *B. napus* orthologs of *PHABULOSA (PHB)*, *REVOLUTA* (*REV*) and *TARGET OF EARLY ACTIVATION TAGGED 2* (*TOE2*), which are involved in plant growth and development, was not increased despite the low expression of miRNAs known to target these genes. Likewise, positive proportional expression interactions were observed in a *B. napus* ortholog of *EMBRYO DEFECTIVE 2204* (*EMB2204*) on chromosome *A*02, whereas an inversely proportional interaction between the miRNA and mRNA target was found for a *B. napus* ortholog of *EMBRYO DEFECTIVE 2016 (EMB2016)* on chromosome *A*03 (Fig. 6). Both genes are involved in embryo development, yet they appear to be regulated in an opposite manner, although further research

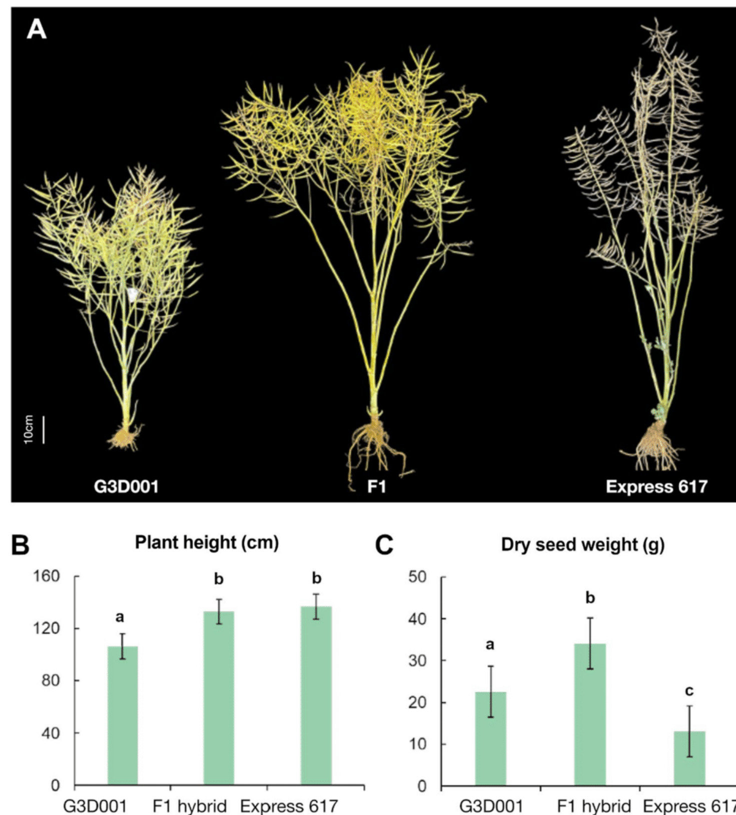

Fig. 5 Phenotypes for field-grown plants of inbred *B. napus* parents G3D001, Express 617 and their F_1 hybrid showing (**a**) plant architecture, (**b**) plant height and (**c**) dry seed weight during an experimental trial in Wuhan, Central China. Biological replicate averages and standard deviations are shown in (**b**) and (**c**) for plant height and dry seed weight, and genotypes showing significant differences ($p < 0.05$) detected by Tukey tests are indicated with letters above the bars

is required to elucidate their role in *B. napus* seed formation. *PHB*, and possibly *PHAVOLUTA (PHV)*, are positive regulators of the *LEAFY COTYLEDON 2* (*LEC2*), a well-known regulator of seed maturation (Tang et al. 2012). A BLAST search from the *A. thaliana* Araport 11 assembly (Cheng et al. 2017) for the *LEC2* coding sequence (*AT1G28300.1*) revealed a single hit that passed the filtering criteria specified for miRNA targets in Materials and Methods. The ortholog corresponded to the *C05p022870.1_BnaEXP* gene model in the Express 617 genome assembly (Lee et al. 2020), which was found to be differentially expressed in late seed development (Table S18 and S26). This example further highlights the broader and indirect impact of miRNAs through gene network interactions. Such transcriptomic networks not only play critical roles in heterosis (Wu et al. 2021) but are also regulated partially by miRNAs (Dong et al. 2022).

Methylated features in early seed formation

Methylation levels were highest in the CpG context, with an average of 80% across all stages and genotypes (Fig. 7). Methylation levels in CHH context were the lowest, ranging from 20 to 27% despite having the highest number of methylated cytosines (Table S27, Fig S21, S22, S23, S24). No methylation bias per chromosome was observed (Table S27). Approximately 12%, 14% and 10% of DMRs were in promoters, exons and introns, respectively, whereas a high percentage of DMRs (43%) were inside repeat motifs (Table S28). Repetitive sequences account for 37.5% of the Express 617 genome (Lee et al. 2020), and although 66% of repeats were methylated with an average 41% methylation level, less than 1% were differentially methylated (Table S29). Most differentially methylated transposable element (TE) families and superfamilies coincided with those that are most frequent in the Express 617 reference genome, such as LTR (long terminal repeat) Copia and Gypsy families. Approximately 70%

Table 1 Predicted mRNA target from differentially expressed miRNAs (DE-miRNAs) in 15 and 30 days after pollination ovules in F_1 and parents by expression level dominance (ELD)

Stage	Predicted DE-miRNA family	DE-miRNA-ELD	miRNA Target	miRNA target-ELD	*A. thaliana* homolog ID	*A. thaliana* homolog name
OS15-F_1	miRNA 165/166 A	IX	C04p041520.1_BnaEXP	IX	AT2G34710	PHABULOSA
	miRNA 165/166 B					
	miRNA 165/166 C					
OS30-F_1	miRNA 3629 A	VIII	A02p005120.1_BnaEXP	VIII	AT1G22090	EMB2204
	miRNA 166 A	IX	C04p041520.1_BnaEXP	IX	AT2G34710	PHABULOSA
	miRNA 165/166 B					
	miRNA 166 B					
	miRNA 166 C					
	miRNA 166 A	IX	C05p024820.1_BnaEXP	XII	AT1G30490	PHAVOLUTA
	miRNA 165/166 B					
	miRNA 166 B					
	miRNA 166 C					
	miRNA 9410/9411 A	IX	A06p017700.1_BnaEXP	VII	AT3G47060	FTSH PROTEASE 7
	miRNA 9410 A					
	miRNA 165/166 B	IX	A02p009610.1_BnaEXP	IX	AT5G60690	REVOLUTA
	miRNA 169 A	IX	A03p029900.3_BnaEXP	XI	AT3G05680	EMB2016
	miRNA 172 A	IX	C09p033660.1_BnaEXP	IX	AT5G60120	TOE2

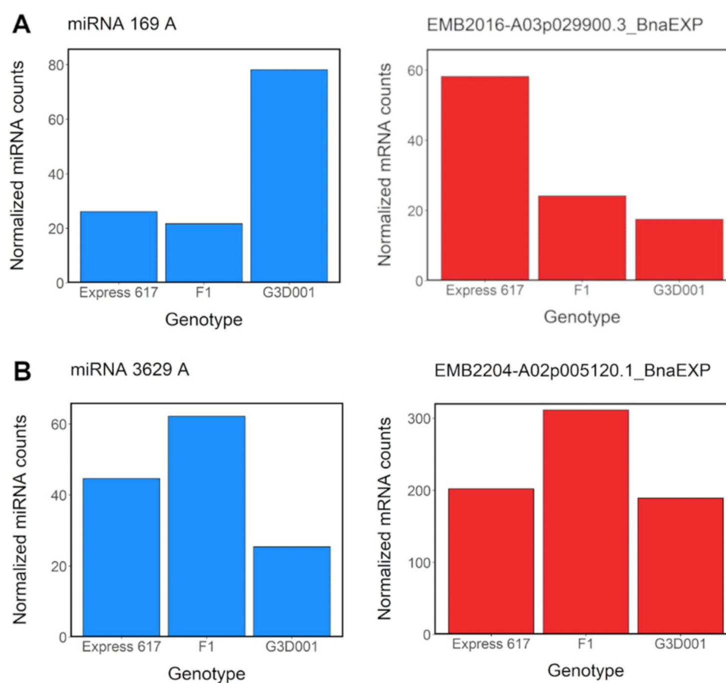

Fig. 6 Normalized expression levels from selected differentially expressed miRNAs (DE-miRNA) and their respective differentially expressed target genes (DEG) in ovules 30 days after pollination in the F_1 and parental genotypes, respectively. (**a**) Inversely proportional miRNA-mRNA target expression of miRNA 169*A* and a *B. napus* ortholog of its target gene *EMB2016* on chromosome *A*03 (A03p029900.3_BnaEXP). (**b**) Proportional miRNA-mRNA target expression of miRNA 3629A and a *B. napus* ortholog of its target gene *EMB2204* on chromosome *A*03 (A02p005120.1_BnaEXP)

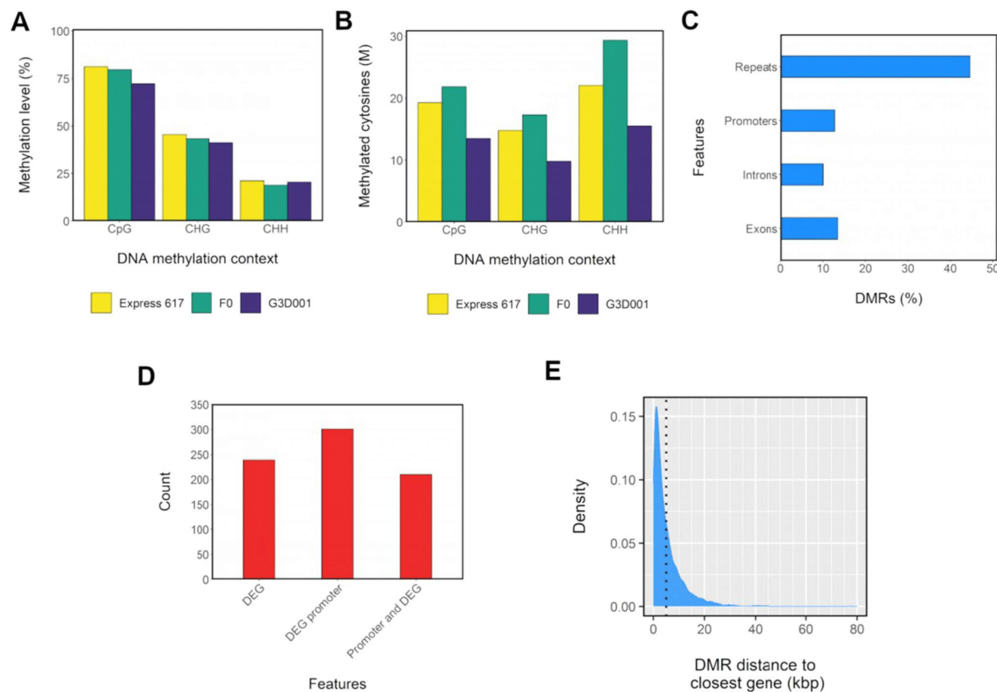

Fig. 7 Methylation patterns in 15 days after pollination ovules from F_0 and parents. (**a**) Methylation level per genotype and DNA methylation context. (**b**) Count of methylated cytosines in million (*M*) scale per genotype and DNA methylation context. (**c**) Distribution of differentially methylated regions (DMRs) across introns, exons, repeats and promoters (1 kbp upstream from gene start). (**d**) Distribution of methylated differential expressed genes (DEGs) and their promoters. (**e**) Kernel density estimation (KED)-based distribution of DMRs distance to closest gene. A dotted line is used to delimit DMRs located 5 kbp from a gene

of these were found within 5 kbp flanking regions of genes (Table S30, S31). Chi-square tests followed by FDR adjusted post-hoc testing at $p > 0.05$ revealed a significant association between the analyzed genomic features (DMRs, TEs, and differentially methylated or non-methylated DE-siRNAs) in terms of their distance to genes and DEGs (Table S32). Interestingly, around 70% of the detected features were in 5 kbp gene flanking regions; nevertheless, only less than 20% of them were found in 5 kbp DEG-flanking regions (Table S32). Moreover, most DMRs inside gene flanking regions converged more around non-DEGs in comparison to DEGs(Tables S32-S33), which suggests an overall conserved pattern of gene regulation despite differential methylation across genotypes.

A total of 392 genes that were both differentially expressed and differentially methylated were regarded as putative epialleles (Table S34). No gene ontology enrichment was found in regard to these putative epialleles. Instead, they covered diverse biological functions such as DNA transcription, carbohydrate and lipid metabolic processes and photosynthesis (Table S35). Interestingly, both the gene body and its promoter were methylated in most putative genetic epialleles (Table S36). Most DMRs were less than 5 kbp away from a gene, suggesting a potential regulatory role (Fig. 7, Fig. S21, S22, S23, S24). Significant correlations using a 0.05 p value threshold were detected when the influence of gene body methylation ($\tau = -0.14$) and gene promoter methylation ($\tau = -0.24$) on gene expression based on Kendall τ tests. Both interactions show a negative correlation coefficient indicating that expression tended to decrease as methylation levels rose. The coefficients nonetheless do not imply a strictly strong relationship as expected due to the diversity of factors affecting gene expression and as also observed in the spectrum of gene expression and methylation interactions during seed formation (Fig. S25, S26, S27) and as also observed in other studies in tomatoes and strawberries (Lang et al. 2017; Cheng et al. 2018). However, during early seed development in

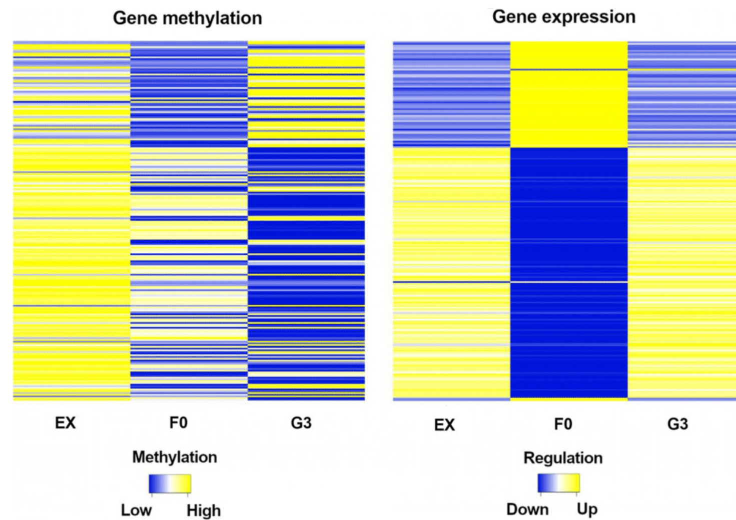

Fig. 8 Gene expression and gene body and promoter methylation in CpG and CHG contexts from 15 days after pollination ovules displaying transgressive patterns in the F_0 and its parents. Genes are sorted in the same order in both heatmaps

the hybrid, proportional interactions with upregulation of hypomethylated genes were most prevalent (Fig. 8). Studies linking methylation and expression and evaluating epialleles have proved beneficial in detecting heterotic patterns in other crops like maize, rice and Arabidopsis (Greaves et al. 2015; Cao et al. 2022; Wang and Wang 2022). In addition, 112,635 CpG islands were detected in all assigned chromosomes in the Express 617 reference genome, with an average length of 363 bp and strong differences in frequency in centromeric regions of different chromosomes (Table S37). Although 86% of all identified CpG islands were methylated, with an average methylation level of 62%, only 1.35% of these were differentially methylated (Table S38).

Segmental and subgenome expression bias in hybrids

Segmental patterns of differential gene expression were visualized by circos plots displaying expression patterns for each chromosome, genotype and stage. The presence of putative expression clusters was assessed more precisely through a 500 kbp genome-wide binning approach where consistent DEGs patterns per segment, chromosome, genotype and stage were grouped as described in Material and Methods. Consequently, 144 differentially expressed segments across genotypes and stages were determined (Table S39). More differentially expressed segments were found in the A subgenome than the C subgenome and most segments found in F_0 comparisons mimicked the expression patterns of the maternal parent Express 617 (Table S39). An example on chromosome *A*03 is shown in Fig. 9. Sequence reads from G3D001 pollinated ovules were employed to exclude the possibility that the observed patterns were due to genomic rearrangements (Table S40). Large-scale deletions were found only in chromosome C01 in G3D001, which accounts for the low expression found on the deleted segments in that chromosome (Fig. S28, S29, S30, Table S39). However, no large-scale rearrangements were found in chromosome A03 in G3D001 (Fig. S31), so that cannot be the reason why a segment on this chromosome showed low expression in both early (15 days) and late (30 days) seed development stages (Fig. 9 and Fig. S32). Moreover, the corresponding chromosome region in Express 617 does not appear to be duplicated, since neither the F_1 nor the F_0 showed a high expression pattern that could have been inherited from a large-scale duplication from the Express 617 parent (Fig. S33, S34). Furthermore, no relationships were observed between methylation level, repeat density or position relative to the centromere. This suggests that specific chromosome segments may correspond more closely to maternal expression patterns than other regions. The mechanisms of such a phenomenon could be associated with parental roles during embryo development, genomic imprinting or chromatin activity and/or genome accessibility for transcription. However, more detailed investigations are necessary to validate these hypotheses.

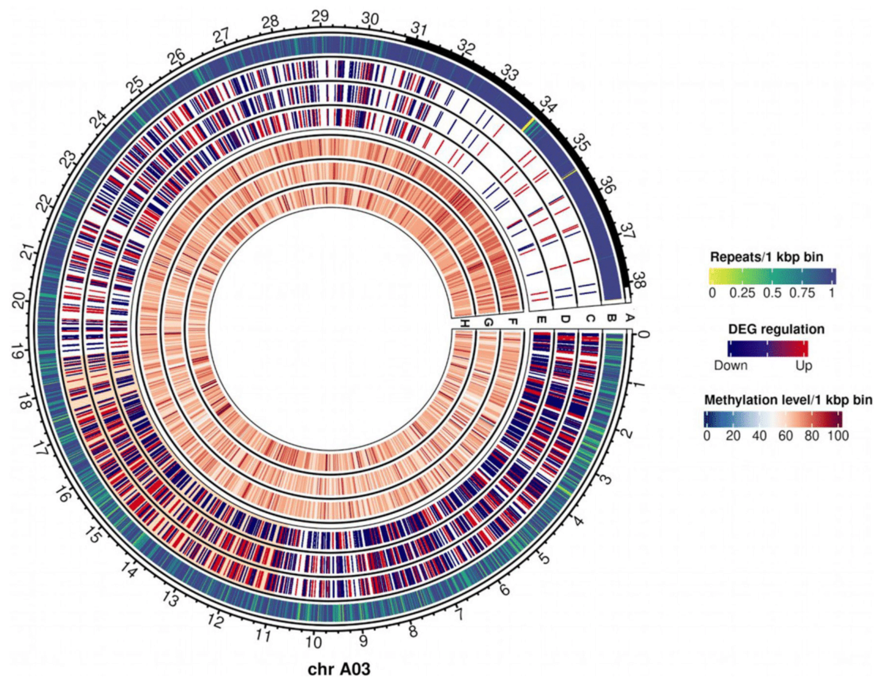

Fig. 9 Differentially expressed genes (DEGs) and methylation levels from 15 days after pollination ovules from F_0 and parents in chromosome *A*03. Outer to inner tracks correspond to: (**a**) Predicted centromere positions in black; (**b**) Repeat density per 1 kbp bin; (**c**–**e**) DEG regulation in (**c**) Express 617, (**d**) F_0 and (**e**) G3D001; (SS**f**–**h**): Methylation levels per 1 kbp bin in (**f**) Express 617, (**g**) F_0 and (**h**) G3D001. A differentially expressed chromosome segment between around 11 Mbp and 18.8 Mbp is highlighted in orange in tracks **c**–**e**

Discussion

The results of this study demonstrate that differential expression and methylation patterns potentially associated with heterotic patterns are already detectable during seed development and seedling stages in an F_1 hybrid. Most DEGs showed maternal and paternal dominances regardless of the tissue and stage, indicating that some of these differential features represent stable regulatory patterns with a general involvement in heterosis. Similar kinds of parental gene expression dominance have been reported previously in oilseed rape and cotton (Yoo et al. 2013; Wu et al. 2018; Wei et al. 2021). Li et al. (2020) demonstrated that expression dominance levels in interspecific hybrids can vary based on the sampled tissue, with stems and leaves showing more additive gene expression in allopolyploid *B. napus* compared to the expression from its diploid progenitors *B. rapa* and *B. oleracea*. Gene expression additivity was also reported by Zhang et al. (2021a) as a main pattern of expression dominance levels in excised pod sections in crosses between *Raphanus sativus* (RR, $2n=18$) and *B. oleracea*, whereas seeds and pods from the homozygous diploids displayed predominantly paternal dominance. The diversity of sampled tissues, species and genotypes in the previous studies and ours could account for the contrasting expression dominance levels observed. The segmental patterns of differential gene expression on a chromosome scale which we observed reflect similar results reported in *B. napus* by Lloyd et al. (2018) and He et al. (2017), who found that large-scale rearrangements induced large segmental expression differences. Interestingly, not all differentially expressed chromosome segments analyzed in our study were associated with large-scale rearrangements. Although the underlying reasons for this observation require further elucidation, this may indicate that chromosome-level patterns of chromatin rearrangement and transcription accessibility may be involved.

Additionally, pollinated F_0 ovules that would develop later into F_1 seeds and plants showed a higher similarity to the maternal genotype Express 617 in terms of gene

expression, small RNA expression and methylation. As noted by Jahnke et al. (2010), this might be due to the triploid nature of the endosperm, which arises from the union of a duplicated maternal gamete and a paternal gamete via double-fertilization. Seeds are composed of a seed coat, embryo and endosperm, with the proportions of the three components varying depending on species and developmental age. Transcriptomic profiling of different seed tissues using laser microdissection has been employed to characterize the transcriptomic profiles during seed formation in *A. thaliana* and *B. napus* (Kirkbride et al. 2019; Ziegler et al. 2019; Khan et al. 2022), and could provide further insights into heterotic patterns during F_1 seed development; however, this technically demanding task was outside of the scope of the present study.

Subgenomic expression bias has been reported earlier in *Brassica* species (Bird et al. 2018, 2021a; Zhang et al. 2021b). Hence, we also investigated expression bias of differentially expressed up- and downregulated genes on a subgenomic basis in each genotype and stage. Although we did not detect any subgenome bias in gene expression, on the whole, more genes were differentially upregulated in G3D001 than in Express 617 (Table S16, Figures S11, S12, S13, S14, S15). The observed genotype-specific bias is potentially a result of different genomic, transcriptomic and epigenomic factors. Firstly, genomic rearrangements gene copy number variations and other structural variants are known to affect various traits in *B. napus* and other polyploid plants (Schiessl et al. 2017; Vollrath et al. 2021; Makhoul et al. 2022) and could have led to potential biases in expression patterns. Transcriptomic aspects such as gene isoforms, gene network interactions and allele expression bias might also be involved in favoring the up- or downregulation from a certain genotype or haplotype (Fan et al. 2020; Schiessl et al. 2020; Golicz et al. 2021). Lastly, epigenomic factors like parental gamete methylation mechanisms, genomic imprinting or differences between parental *cis–trans* regulating factors, miRNA isoforms (isomiRs) and TE families and densities could all result in potential genotype-biased or haplotype-biased expression (Jain et al. 2018; Go and Civetta 2020; Gill et al. 2021).

Around 12–18% of features shared in at least 2–3 stages followed maternal dominant patterns, highlighting the potential relevance of maternal effects on transcriptomic and epigenomic regulation of early development in this hybrid. Furthermore, less than 3% of all expression and methylation features had the same expression and methylation patterns across all stages, suggesting that the role of those features might be more essential throughout seed and early seedling development. Our results in regard to GO enrichment of differentially regulated features also underlined the role of heterotic patterns in driving key biological functions involved in photosynthesis, stress response, growth and development This also highlights the potential of RNA-Seq for global transcriptomic profiling as also supported by its high accuracy and robustness (Everaert et al. 2017; Coenye 2021). Furthermore, heterosis has been associated with a combination of similar functions like photosynthetic activity and cell division, which already help to enhance performance during early developmental stages (Liu et al. 2021). Here, DEGs involved in reproduction and meiotic functions already showed transgressive upregulated expression in ovules from selfed-F_1 plants at 15 days after pollination. Furthermore, the hybrids revealed a more robust structure and higher dry seed weight in comparison to their parents. Enrichment of GO groups linked to cell division, stress response and development functions, such as the ones detected in the hybrids of this study, have also led to similar phenotypes in hybrids of oilseed rape, cotton and Arabidopsis (Shen et al. 2017; Yang et al. 2017; Shahzad et al. 2020; Zhu et al. 2020; Rong et al. 2021). Information on these genes and their expression patterns could be of potential interest for proteomic validation and approaches using transcriptomic data for hybrid performance prediction.

Interestingly, we observed differentially expressed miRNAs in early and late seed development among miRNA families which are normally involved in plant growth and development (Plotnikova et al. 2019; Dong et al. 2022; Verma et al. 2022); thus, broadening the range and data availability of miRNAs in oilseed rape development, for example, miR172 regulates not only the flowering time pathway, but also embryo development by controlling *APETALA 2* (*AP2*) and *AP2-like* genes such as *TOE2* (Boutilier et al. 2002; Shivaraj et al. 2018; Nowak et al. 2022) was found in our study. miR165/166 families control leaf adaxial/abaxial development and embryogenesis by targeting the class III homeodomain leucine zipper (HD-ZIP III) transcription factor gene family which includes the *REV*, *PHV* and *PHB* (Wang et al. 2007; Tang et al. 2012). Both *PHB* and *PHV* were identified in the present study and have been described to indirectly regulate *LEC2*, a gene that promotes embryo formation in Arabidopsis and seed size and seed lipid biosynthesis in *B. napus* (Braybrook et al. 2006; Tang et al. 2012; Wójcik et al. 2017; Miller et al. 2019). As noted by Dong et al. (2022), miR169 targets *C-REPEAT BINDING FACTOR* (*CBF*) and *NUCLEAR FACTOR YA* (*NF-YA*) genes. During late seed development in *B. napus*, we found that miR169 targets *EMB2016,* a member of the EMB gene family critical for embryo development (Tzafrir et al. 2004; Růžička et al. 2017; Meinke 2020), while *EMB2204* was targeted by miR3629. mir3629 was first reported in *Vitis vinifera* cv. Pinot Noir by Pantaleo et al. (2010) and has since been reported in *Camellia azalea*, in response to chilling in *Prunus persica* and in disease susceptibility in *V. vinifera* cv. Bosco and *V. vinifera* cv. Chardonnay (Barakat et al. 2012; Pantaleo et al. 2016; Yin

et al. 2016; Snyman et al. 2017). mir9410 has been detected in *B. oleracea* and *B. rapa* (Lukasik et al. 2013; Zhang et al. 2018), yet no clear function information for mir9410 exists for Brassica species. In our study, miR9410 targeted a *filamentation temperature sensitive protein H 1* (*FtsH7*) gene copy encoding a protease that in turns degrades *D*1 protein in photosystem II. *FtsH* genes have been reported in tomato, sorghum, Arabidopsis and *B. napus* (Xu et al. 2021; Yi et al. 2022).

The study identified multiple differential miRNA sequences and their putative targets with implications on plant development and performance. Further validation of targets associated with DE-miRNAs can potentially be achieved through degradome sequencing (German et al. 2008), precise isomiRs classification (Morin et al. 2008; Sablok et al. 2015; Yang et al. 2019), target knock-out experiments (Jain et al. 2018; Wei et al. 2018; Li et al. 2021) or gene co-expression networks (Schiessl et al. 2020) in order to delimit their role in seed and embryo formation in *B. napus*.

Overall, the number of methylated cytosines in the CHH context was higher in all genotypes compared to other contexts; nonetheless, methylation levels were higher in CpG and CHG contexts, as observed previously in multiple plants species (e.g. Niederhuth et al. 2016; Bartels et al. 2018). Methylation is generally associated with gene downregulation through transcription inhibition. Nevertheless, hypermethylation and hypomethylation were also linked with up- and downregulation, respectively. Proportional gene hypermethylation and upregulation were observed in mice and human cells (Arechederra et al. 2018; Rauluseviciute et al. 2020) as well as in strawberry and tomatoes (Lang et al. 2017; Cheng et al. 2018); however, no mechanisms explaining gene activation through hypermethylation are fully known so far; thus, further research would elucidate the interactions between methylation and gene regulation, particularly in relation to heterozygosity and heterosis.

In our study, we evaluated methylation during seed development because parental asymmetric methylation and genomic imprinting occurs mostly at that stage in flowering plants (Batista and Köhler 2020). DNA hypomethylation of the female gamete and paternal gamete hypermethylation has been reported in many flowering plants, including Arabidopsis, rice and maize (Gehring et al. 2009; Zemach et al. 2010; Zhang et al. 2014b). Interestingly, we observed contrasting patterns of maternal hypermethylation and paternal hypomethylation in the F_0. Similar parental methylation trends were also observed in the F_1 despite allele segregation. Such patterns were also reported by Liu et al. (2018) in *B. napus* and by Grover et al. (2020) in *B.rapa*. As a possible explanation for maternal hypermethylation, Grover et al. (2020) proposed that a high expression of so-called siren siRNAs in the seed coat could trigger maternal DNA methylation during seed development. The molecular mechanisms and effects of genomic imprinting, where an allele follows a parental expression pattern due to inherited epigenomic modifications, are restricted mostly to the endosperm rather than the embryo in flowering plants and have been extensively discussed by Weigel and Colot (2012) and Batista and Köhler (2020), respectively. The role of imprinted genes has been linked to chromatin modification, hormone biosynthesis, nutrient transfer, endosperm proliferation and seed size regulation (reviewed by Jiang and Köhler 2012; Batista and Köhler 2020). Furthermore, Rong et al. (2021) reported the enrichment of transposable elements located in 5 kbp flanking regions of imprinted genes in *B. napus*. Cao et al. (2022) analyzed imprinted genes in six backcrossing generations of maize as well as in three selfing generations derived from the 6th backcross. They proposed that the divergence between TEs derived from 24-nt siRNAs in the parental maize genomes might have led to transgenerational inheritance of imprinted genes. Putative imprinted genes were also found in the seedling and seed development stages in our study. Epigenetic changes have been reported as relevant heterotic factors which are influenced by allele diversity, parental effects and environmental conditions (Botet and Keurentjes 2020). Epigenomic parental effects are more likely to occur during seed formation, when gametes fuse to form a zygote, given that this stage is marked by epigenomic features involving siRNA, DNA methylation, imprinting and chromatin activity. Because these features shape the epigenomic and transcriptomic landscape of the zygote and, potentially, its future development as a seedling, hybrids can potentially benefit strongly from a heterotic advantage imparted by these features in these very early developmental stages.

Most frequently, differentially methylated transposable elements corresponded to abundant Copia and Gypsy families. Transposable elements are key factors in speciation and subgenome expression patterns (Bird et al. 2018, 2021a; Bottani et al. 2018) and are known for their high variability across plant species (Novák et al. 2020; Mhiri et al. 2022). TEs can also alter the epigenetic landscape in relation to hybrid fitness (Serrato-Capuchina and Matute 2018). Therefore, detailed assessment of transposable element densities and compositions between hybrids and their parents could be beneficial. In addition, siRNAs are known to mediate silencing of transposable elements via the RNA-directed DNA methylation (RdDM) pathway. At the same time, TEs are a source of sRNAs, including siRNAs, that could potentially silence TEs through a post-transcriptional gene silencing (PTGS) process (Matzke and Mosher 2014; Gill et al. 2021). As also described by Rong et al. (2021), most differentially methylated TEs were found

in 5 kbp gene flanking regions. Because most TEs, DMRs and DE-siRNAS converged into regions directly flanking genes, whereas there was no abundance in regions flanking DEGs, gene regulation in the hybrid appears to be associated with conservation of key genetic functions by (i) maintaining similar gene expression patterns between genotypes despite differential methylation in gene flanking regions for most genes and (ii) reducing the number of DMRs, DE-siRNAs and differentially methylated TEs in the proximity of DEGs similarly to other reports in poplar, mammals and invertebrates (Blake et al. 2020; Zhang et al. 2020; Cardoso-Júnior et al. 2021; Dixon and Matz 2022). This could be one of the most interesting and significant results of the present study. Methylation data have been employed in plant breeding to investigate and improve flowering time, disease susceptibility and abiotic stress response (Mercé et al. 2020; Shaikh et al. 2022), and have enabled genomic diversity expansion and discovery of cis-regulatory elements (Xu et al. 2019; Crisp et al. 2020). The present study highlights the spectrum of methylation patterns and putative epialleles in oilseed rape with special implications on seed formation. Additionally, most CpG islands were not differentially methylated between genotypes, indicating a putative role in regulating gene expression, as reported in rice and Arabidopsis (Ashikawa 2001).

The sequencing data gathered from single tissues in the present study allowed an integrated view of gene expression, small RNA interactions and genomic methylation during early developmental stages in an oilseed rape hybrid developed from two distant genotypes. Coding and non-coding features which were differentially expressed or methylated in this study provide new insights into early expression of heterosis in oilseed rape seeds and seedlings from a molecular viewpoint and constitute an extensive multiomics atlas for oilseed rape breeding. The extent of these features in an allopolyploid model crop like *B. napus* also have potential implications in other polyploid crops where heterosis still remains to be exploited, such as wheat and potatoes (Steeg et al. 2022). Patterns of expression and methylation dominance levels could also contribute a new level of understand regarding allele-specific gene expression (Fan et al. 2020; Sands et al. 2021), isoform expression (Vitting-Seerup and Sandelin 2019; Yao et al. 2020; Golicz et al. 2021), gene fusion and dosage (Mahmoud et al. 2019; Serin Harmanci et al. 2020; Bird et al. 2021b) as well as non-germline omics variations among F_1 plants and populations (Higgins et al. 2018; Cortijo et al. 2019; Orantes-Bonilla et al. 2022; Quezada-Martinez et al. 2022). Their role in heterotic gene expression patterns is ultimately also of interest for transcriptome-based genomic selection or hybrid performance prediction (e.g. Frisch et al. 2010). Defining the roles of differentially expression regulatory features in early developmental stages of hybrids could be used to enhance expression-based prediction model (Seifert et al. 2018b; Zrimec et al. 2020; Cheng et al. 2021; Hu et al. 2021c; Knoch et al. 2021). Altogether our findings highlight transcriptomics and epigenomic differences between early developmental stages in F_1 and F_0 in terms of methylation level as well as in gene and small RNA expression. The contribution of differential coding and non-coding features to early hybrid seed formation is of key interest for hybrid breeding and deserves further evaluation using more diverse genotypes, heterotic pools and species. Future developments in sequencing and bioinformatics will also aid in elucidating the role and interactions among transcriptomic and epigenomic features at higher resolution, helping to expand current knowledge and applications of heterosis in polyploid crops.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-023-04345-7.

Acknowledgements The authors acknowledge support for bioinformatic resources from the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) and the Bioinformatics Core Facility at JLU. The experimental design was created with BioRender.com.s

Author contributions statement R.J.S. and J.Z. conceived and supervised the study. M.O.B. drafted the manuscript and designed the bioinformatic analyses. H.W. conducted bioinformatic studies, generated crosses, extracted and sampled pollinated ovules for sequencing and contributed to data analysis. D.H. carried and supervised the field experiment. W.L. contributed to experimental trials. H.T.L. performed genome synteny analyses and contributed to whole-genome bisulfite and mRNA analyses. A.A.G. contributed to analyses of transcriptomic and epigenomic features. All authors read and approved the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was performed within the framework of the Joint Sino-German Research Program (2018) with support from grant SN14/22–1 to RJS from the German Research Foundation (DFG) and grant 31861133016 to JZ from the National Natural Science Foundation of China (NSFC).

Data availability The mRNA, sRNA and WGBS libraries and fragment count datasets generated in this study are found in the GEO data repository under accession GSE202610. G3D001 genomic reads from self-pollinated ovules are found in NCBI Bioproject PRJNA850551.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended

use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alonso-Peral MM, Trigueros M, Sherman B, Ying H, Taylor JM, Peacock WJ et al (2017) Patterns of gene expression in developing embryos of Arabidopsis hybrids. Plant J 89:927–939. https://doi.org/10.1111/tpj.13432

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. https://www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed 2021

Arechederra M, Daian F, Yim A, Bazai SK, Richelme S, Dono R et al (2018) Hypermethylation of gene body CpG islands predicts high dosage of functional oncogenes in liver cancer. Nat Commun 9:3164. https://doi.org/10.1038/s41467-018-05550-5

Ashikawa I (2001) Gene-associated CpG islands in plants as revealed by analyses of genomic sequences. Plant J 26:617–625. https://doi.org/10.1046/j.1365-313x.2001.01062.x

Axtell MJ (2013) Classification and comparison of small RNAs from plants. Annu Rev Plant Biol 64:137–159. https://doi.org/10.1146/annurev-arplant-050312-120043

Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A et al (2016) Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res 44:W147–W153. https://doi.org/10.1093/nar/gkw419

Bao J, Lee S, Chen C, Zhang X, Zhang Y, Liu S et al (2005) Serial analysis of gene expression study of a hybrid rice strain (LYP9) and its parental cultivars. Plant Physiol 138:1216–1231. https://doi.org/10.1104/pp.105.060988

Barakat A, Sriram A, Park J, Zhebentyayeva T, Main D, Abbott A (2012) Genome wide identification of chilling responsive microRNAs in *Prunus persica*. BMC Genomics 13:481. https://doi.org/10.1186/1471-2164-13-481

Bartels A, Han Q, Nair P, Stacey L, Gaynier H, Mosley M et al (2018) Dynamic DNA methylation in plant growth and development. Int J Mol Sci 19:2144. https://doi.org/10.3390/ijms19072144

Basunanda P, Radoev M, Ecke W, Friedt W, Becker HC, Snowdon RJ (2010) Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). Theor Appl Genet 120:271–281. https://doi.org/10.1007/s00122-009-1133-z

Batista RA, Köhler C (2020) Genomic imprinting in plants-revisiting existing models. Genes Dev 34:24–36. https://doi.org/10.1101/gad.332924.119

Bird KA, VanBuren R, Puzey JR, Edger PP (2018) The causes and consequences of subgenome dominance in hybrids and recent polyploids. New Phytol 220:87–93. https://doi.org/10.1111/nph.15256

Bird KA, Niederhuth CE, Ou S, Gehan M, Pires JC, Xiong Z et al (2021a) Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid *Brassica napus*. New Phytol 230:354–371. https://doi.org/10.1111/nph.17137

Bird KA, Pires JC, VanBuren R, Xiong Z, Edger PP (2021b) Gene balance in allopolyploids: homoeologous exchanges show signs of dosage constraint and dosage constraint of biased homoeologs differs between subgenomes. BioRxiv. https://doi.org/10.1101/2021.11.16.468838

Blake LE, Roux J, Hernando-Herraez I, Banovich NE, Perez RG, Hsiao CJ et al (2020) A comparison of gene expression and DNA methylation patterns across tissues and species. Genome Res 30:250–262. https://doi.org/10.1101/gr.254904.119

Bonjean AP, Dequidt C, Sang T (2016) Rapeseed in China. OCL 23:D605. https://doi.org/10.1051/ocl/2016045

Botet R, Keurentjes JJB (2020) The role of transcriptional regulation in hybrid vigor. Front Plant Sci 11:410. https://doi.org/10.3389/fpls.2020.00410

Bottani S, Zabet NR, Wendel JF, Veitia RA (2018) Gene Expression dominance in Allopolyploids: hypotheses and models. Trends Plant Sci 23:393–402. https://doi.org/10.1016/j.tplants.2018.01.002

Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L et al (2002) Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. Plant Cell 14:1737–1749. https://doi.org/10.1105/tpc.001941

Braybrook SA, Stone SL, Park S, Bui AQ, Le BH, Fischer RL et al (2006) Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. Proc Natl Acad Sci U S A 103:3468–3473. https://doi.org/10.1073/pnas.0511331103

BSA (2022) Federal Plant Variety Office (BSA) descriptive variety 2022 list: cereals, maize, oil and fibre plants, legumes, beets, catch crops. https://www.bundessortenamt.de/bsa/media/Files/BSL/bsl_getreide_2022.pdf. Accessed 2022

Cao S, Wang L, Han T, Ye W, Liu Y, Sun Y et al (2022) Small RNAs mediate transgenerational inheritance of genome-wide trans-acting epialleles in maize. Genome Biol 23:53. https://doi.org/10.1186/s13059-022-02614-0

Cardoso-Júnior CA, Yagound B, Ronai I, Remnant EJ, Hartfelder K, Oldroyd BP (2021) DNA methylation is not a driver of gene expression reprogramming in young honey bee workers. Mol Ecol 30:4804–4818. https://doi.org/10.1111/mec.16098

Castillo-Bravo R, Fort A, Cashell R, Brychkova G, McKeown PC, Spillane C (2022) Parent-of-origin effects on seed size modify heterosis responses in *Arabidopsis thaliana*. Front Plant Sci 13:835219. https://doi.org/10.3389/fpls.2022.835219

Catoni M, Tsang JM, Greco AP, Zabet NR (2018) DMRcaller: a versatile R/bioconductor package for detection and visualization of differentially methylated regions in CpG and non-CpG contexts. Nucleic Acids Res 46:e114. https://doi.org/10.1093/nar/gky602

Chalhoub B, Denoeud F, Liu S, Parkin IAP, Tang H, Wang X et al (2014) Plant genetics. Early allopolyploid evolution in the post-neolithic *Brassica napus* oilseed genome. Science 345:950–953. https://doi.org/10.1126/science.1253435

Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chen L, Zhu Y, Ren X, Yao D, Song Y, Fan S et al (2022) Heterosis and differential DNA methylation in soybean hybrids and their parental lines. Plants Basel 11:1136. https://doi.org/10.3390/plants11091136

Cheng C-Y, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD (2017) Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. Plant J 89:789–804. https://doi.org/10.1111/tpj.13415

Cheng J, Niu Q, Zhang B, Chen K, Yang R, Zhu JK et al (2018) Downregulation of RdDM during strawberry fruit ripening. Genome Biol 19:1–14. https://doi.org/10.1186/s13059-018-1587-x

Cheng C-Y, Li Y, Varala K, Bubert J, Huang J, Kim GJ et al (2021) Evolutionarily informed machine learning enhances the power

of predictive gene-to-phenotype relationships. Nat Commun 12:5627. https://doi.org/10.1038/s41467-021-25893-w
Church BV, Williams HT, Mar JC (2019) Investigating skewness to understand gene expression heterogeneity in large patient cohorts. BMC Bioinform 20:1–14. https://doi.org/10.1186/s12859-019-3252-0
Coenye T (2021) Do results obtained with RNA-sequencing require independent verification? Biofilm 3:100043. https://doi.org/10.1016/j.bioflm.2021.100043
Cortijo S, Aydin Z, Ahnert S, Locke JC (2019) Widespread inter-individual gene expression variability in Arabidopsis thaliana. Mol Syst Biol 15:e8591
Crisp PA, Marand AP, Noshay JM, Zhou P, Lu Z, Schmitz RJ et al (2020) Stable unmethylated DNA demarcates expressed genes and their cis-regulatory space in plant genomes. Proc Natl Acad Sci 117:23991–24000. https://doi.org/10.1073/pnas.2010250117
Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM et al (2022) Ensembl 2022. Nucleic Acids Res 50:D988–D995. https://doi.org/10.1093/nar/gkab1049
Dai X, Zhuang Z, Zhao PX (2018) psRNATarget: a plant small RNA target analysis server (2017 release). Nucleic Acids Res 46:W49–W54. https://doi.org/10.1093/nar/gky316
Dainat J (2019) *AGAT: Another Gff analysis toolkit to handle annotations in any GTF/GFF format. (Version v0.5.0)*. 10.5281/zenodo.3552717. Accessed 2020
Darwin C (1876) The effects of cross and self fertilisation in the vegetable kingdom. John Murry, London
Di Y, Schafer DW, Cumbie JS, Chang JH (2011) The NBP negative binomial model for assessing differential gene expression from RNA-Seq. Stat Appl Genet Mol Biol 10:1–28. https://doi.org/10.2202/1544-6115.1637
Dixon G, Matz M (2022) Changes in gene body methylation do not correlate with changes in gene expression in Anthozoa or Hexapoda. BMC Genomics 23:1–11. https://doi.org/10.1186/s12864-022-08474-z
Dong Q, Hu B, Zhang C (2022) microRNAs and their roles in plant development. Front Plant Sci 13:824240. https://doi.org/10.3389/fpls.2022.824240
Doyle JJ, Doyle J (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15
Everaert C, Luypaert M, Maag JL, Cheng QX, Dinger ME, Hellemans J et al (2017) Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. Sci Rep 7:1–11. https://doi.org/10.1038/s41598-017-01617-3
Ewels P, Magnusson M, Lundin S, Käller M (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. https://doi.org/10.1093/bioinformatics/btw354
Fan J, Hu J, Xue C, Zhang H, Susztak K, Reilly MP et al (2020) ASEP: Gene-based detection of allele-specific expression across individuals in a population by RNA sequencing. PLoS Genet 16:e1008786. https://doi.org/10.1371/journal.pgen.1008786
FAO (2022). FAOSTAT. https://www.fao.org/faostat/en. Accessed 2022.
Fitz-James MH, Cavalli G (2022) Molecular mechanisms of transgenerational epigenetic inheritance. Nat Rev Genet 23:325–341. https://doi.org/10.1038/s41576-021-00438-5
Friedt W, Tu J, Fu T (2018) Academic and economic importance of *Brassica napus* rapeseed. In: Liu S, Snowdon R (eds) The *Brassica napus* genome chalhoub. Springer International Publishing, Cham, pp 1–20
Frisch M, Thiemann A, Fu J, Schrag TA, Scholten S, Melchinger AE (2010) Transcriptome-based distance measures for grouping of germplasm and prediction of hybrid performance in maize. Theor Appl Genet 120:441–450. https://doi.org/10.1007/s00122-009-1204-1
Fujimoto R, Uezono K, Ishikura S, Osabe K, Peacock WJ, Dennis ES (2018) Recent research on the mechanism of heterosis is important for crop and vegetable breeding systems. Breed Sci 68:145–158. https://doi.org/10.1270/jsbbs.17155
Gardiner-Garden M, Frommer M (1987) CpG Islands in vertebrate genomes. J Mol Biol 196:261–282. https://doi.org/10.1016/0022-2836(87)90689-9
Ge SX, Son EW, Yao R (2018) iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. BMC Bioinform 19:534. https://doi.org/10.1186/s12859-018-2486-6
Ge SX, Jung D, Yao R (2020) ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics 36:2628–2629. https://doi.org/10.1093/bioinformatics/btz931
Gehring M, Bubb KL, Henikoff S (2009) Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science 324:1447–1451. https://doi.org/10.1126/science.1171609
German MA, Pillay M, Jeong D-H, Hetawal A, Luo S, Janardhanan P et al (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 26:941–946. https://doi.org/10.1038/nbt1417
Gill RA, Scossa F, King GJ, Golicz AA, Tong C, Snowdon RJ et al (2021) On the role of transposable elements in the regulation of gene expression and subgenomic interactions in crop genomes. Crit Rev Plant Sci 40:157–189. https://doi.org/10.1080/07352689.2021.1920731
Girke A, Schierholt A, Becker HC (2012) Extending the rapeseed gene pool with resynthesized *Brassica napus* II: Heterosis. Theor Appl Genet 124:1017–1026. https://doi.org/10.1007/s00122-011-1765-7
Go AC, Civetta A (2020) Hybrid incompatibilities and transgressive gene expression between two closely related subspecies of Drosophila. Front Genet 11:599292. https://doi.org/10.3389/fgene.2020.599292
Golicz AA, Allu AD, Li W, Lohani N, Singh MB, Bhalla PL (2021) A dynamic intron retention program regulates the expression of several hundred genes during pollen meiosis. Plant Reprod 34:225–242. https://doi.org/10.1007/s00497-021-00411-6
Greaves IK, Gonzalez-Bayon R, Wang L, Zhu A, Liu P-C, Groszmann M et al (2015) Epigenetic changes in hybrids. Plant Physiol 168:1197–1205. https://doi.org/10.1104/pp.15.00231
Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 34:D140–D144. https://doi.org/10.1093/nar/gkj112
Groszmann M, Gonzalez-Bayon R, Greaves IK, Wang L, Huen AK, Peacock WJ et al (2014) Intraspecific Arabidopsis hybrids show different patterns of heterosis despite the close relatedness of the parental genomes. Plant Physiol 166:265–280. https://doi.org/10.1104/pp.114.243998
Grover JW, Burgess D, Kendall T, Baten A, Pokhrel S, King GJ et al (2020) Abundant expression of maternal siRNAs is a conserved feature of seed development. Proc Natl Acad Sci U S A 117:15305–15315. https://doi.org/10.1073/pnas.2001332117
Gu Z, Gu L, Eils R, Schlesner M, Brors B (2014) circlize Implements and enhances circular visualization in R. Bioinformatics 30:2811–2812. https://doi.org/10.1093/bioinformatics/btu393
Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32:2847–2849. https://doi.org/10.1093/bioinformatics/btw313
He Z, Wang L, Harper AL, Havlickova L, Pradhan AK, Parkin IAP et al (2017) Extensive homoeologous genome exchanges in allopolyploid crops revealed by mRNAseq-based visualization. Plant Biotechnol J 15:594–604. https://doi.org/10.1111/pbi.12657

Higgins EE, Clarke WE, Howell EC, Armstrong SJ, Parkin IAP (2018) Detecting de novo homoeologous recombination events in cultivated *Brassica napus* using a genome-wide SNP array. G Bethesda 8:2673–2683. https://doi.org/10.1534/g3.118.200118

Hochholdinger F, Hoecker N (2007) Towards the molecular basis of heterosis. Trends Plant Sci 12:427–432. https://doi.org/10.1016/j.tplants.2007.08.005

Hu D, Jing J, Snowdon RJ, Mason AS, Shen J, Meng J et al (2021a) Exploring the gene pool of *Brassica napus* by genomics-based approaches. Plant Biotechnol J 19:1693–1712. https://doi.org/10.1111/pbi.13636

Hu D, Zhao Y, Shen J, He X, Zhang Y, Jiang Y et al (2021b) Genome-wide prediction for hybrids between parents with distinguished difference on exotic introgressions in *Brassica napus*. Crop J 9:1169–1178. https://doi.org/10.1016/j.cj.2020.11.002

Hu H, Campbell MT, Yeats TH, Zheng X, Runcie DE, Covarrubias-Pazaran G et al (2021c) Multi-omics prediction of oat agronomic and seed nutritional traits across environments and in distantly related populations. Theor Appl Genet 134:4043–4054. https://doi.org/10.1007/s00122-021-03946-4

Jahnke S, Sarholz B, Thiemann A, Kühr V, Gutiérrez-Marcos JF, Geiger HH et al (2010) Heterosis in early seed development: a comparative study of F1 embryo and endosperm tissues 6 days after fertilization. Theor Appl Genet 120:389–400. https://doi.org/10.1007/s00122-009-1207-y

Jain A, Anand S, Singh NK, Das S (2018) Sequence and functional characterization of MIRNA164 promoters from Brassica shows copy number dependent regulatory diversification among homeologs. Funct Integr Genomics 18:369–383. https://doi.org/10.1007/s10142-018-0598-8

Jian H, Yang B, Zhang A, Ma J, Ding Y, Chen Z et al (2018) Genome-wide identification of microRNAs in response to cadmium stress in oilseed rape (*Brassica napus* L.) using high-throughput sequencing. Int J Mol Sci 19:1431. https://doi.org/10.3390/ijms19051431

Jian H, Zhang A, Ma J, Wang T, Yang B, Shuang LS et al (2019) Joint QTL mapping and transcriptome sequencing analysis reveal candidate flowering time genes in *Brassica napus* L. BMC Genomics 20:21. https://doi.org/10.1186/s12864-018-5356-8

Jiang H, Köhler C (2012) Evolution, function, and regulation of genomic imprinting in plant seed development. J Exp Bot 63:4713–4722. https://doi.org/10.1093/jxb/ers145

Johnson NR, Yeoh JM, Coruh C, Axtell MJ (2016) Improved placement of multi-mapping small RNAs. G3 Bethesda 6:2103–2111. https://doi.org/10.1534/g3.116.030452

Kawanabe T, Ishikura S, Miyaji N, Sasaki T, Wu LM, Itabashi E et al (2016) Role of DNA methylation in hybrid vigor in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A 113:E6704–E6711. https://doi.org/10.1073/pnas.1613372113

Khan D, Ziegler DJ, Kalichuk JL, Hoi V, Huynh N, Hajihassani A et al (2022) Gene expression profiling reveals transcription factor networks and subgenome bias during *Brassica napus* seed development. Plant J 109:477–489. https://doi.org/10.1111/tpj.15587

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907–915. https://doi.org/10.1038/s41587-019-0201-4

Kirkbride RC, Lu J, Zhang C, Mosher RA, Baulcombe DC, Chen ZJ (2019) Maternal small RNAs mediate spatial-temporal regulation of gene expression, imprinting, and seed development in Arabidopsis. Proc Natl Acad Sci U S A 116:2761–2766. https://doi.org/10.1073/pnas.1807621116

Knoch D, Werner CR, Meyer RC, Riewe D, Abbadi A, Lücke S et al (2021) Multi-omics-based prediction of hybrid performance in canola. Theor Appl Genet 134:1147–1165. https://doi.org/10.1007/s00122-020-03759-x

Krueger F, Andrews SR (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics 27:1571–1572. https://doi.org/10.1093/bioinformatics/btr167

Krueger F, James F, Ewels P, Afyounian E and Schuster-Boeckler B (2021) *FelixKrueger/TrimGalore: v0.6.7*. *Zenodo*. 10.5281/zenodo.5127899

Lai Z, Gross BL, Zou Y, Andrews J, Rieseberg LH (2006) Microarray analysis reveals differential gene expression in hybrid sunflower species. Mol Ecol 15:1213–1227. https://doi.org/10.1111/j.1365-294X.2006.02775.x

Lang Z, Wang Y, Tang K, Tang D, Datsenka T et al (2017) Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. Proc Natl Acad Sci 114:E4511–E4519. https://doi.org/10.1073/pnas.1705233114

Lauss K, Wardenaar R, Oka R, van Hulten MHA, Guryev V, Keurentjes JJB et al (2018) Parental DNA methylation states are associated with heterosis in epigenetic hybrids. Plant Physiol 176:1627–1645. https://doi.org/10.1104/pp.17.01054

Lee H, Chawla HS, Obermeier C, Dreyer F, Abbadi A, Snowdon R (2020) Chromosome-scale assembly of winter oilseed rape *Brassica napus*. Front Plant Sci 11:496. https://doi.org/10.3389/fpls.2020.00496

Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li A, Liu D, Wu J, Zhao X, Hao M, Geng S et al (2014) mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. Plant Cell 26:1878–1900. https://doi.org/10.1105/tpc.114.124388

Li J, Huang Q, Sun M, Zhang T, Li H, Chen B et al (2016) Global DNA methylation variations after short-term heat shock treatment in cultured microspores of *Brassica napus* cv. Topas Sci Rep 6:38401. https://doi.org/10.1038/srep38401

Li M, Wang R, Wu X, Wang J (2020) Homoeolog expression bias and expression level dominance (ELD) in four tissues of natural allotetraploid *Brassica napus*. BMC Genomics 21:330. https://doi.org/10.1186/s12864-020-6747-1

Li Q, Shah N, Zhou X, Wang H, Yu W, Luo J et al (2021) Identification of micro ribonucleic acids and their targets in response to plasmodiophora brassicae infection in *Brassica napus*. Front Plant Sci 12:734419. https://doi.org/10.3389/fpls.2021.734419

Li H (2016) *Seqtk*. https://github.com/lh3/seqtk. Accessed 2019

Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–930. https://doi.org/10.1093/bioinformatics/btt656

Liu J, Li J, Liu H-F, Fan S-H, Singh S, Zhou X-R et al (2018) Genome-wide screening and analysis of imprinted genes in rapeseed (*Brassica napus* L.) endosperm. DNA Res 25:629–640. https://doi.org/10.1093/dnares/dsy030

Liu W, He G, Deng XW (2021) Biological pathway expression complementation contributes to biomass heterosis in Arabidopsis. Proc Natl Acad Sci U S A 118:e2023278118. https://doi.org/10.1073/pnas.2023278118

Lloyd A, Blary A, Charif D, Charpentier C, Tran J, Balzergue S et al (2018) Homoeologous exchanges cause extensive dosage-dependent gene expression changes in an allopolyploid crop. New Phytol 217:367–377. https://doi.org/10.1111/nph.14836

Louwaars NP (2018) Plant breeding and diversity: a troubled relationship? Euphytica 214:114. https://doi.org/10.1007/s10681-018-2192-5
Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8
Lukasik A, Pietrykowska H, Paczek L, Szweykowska-Kulinska Z, Zielenkiewicz P (2013) High-throughput sequencing identification of novel and conserved miRNAs in the *Brassica oleracea* leaves. BMC Genomics 14:801. https://doi.org/10.1186/1471-2164-14-801
Lunardon A, Johnson NR, Hagerott E, Phifer T, Polydore S, Coruh C et al (2020) Integrated annotations and analyses of small RNA-producing loci from 47 diverse plants. Genome Res 30:497–513. https://doi.org/10.1101/gr.256750.119
Ma J, Zhang D, Cao Y, Wang L, Li J, Lübberstedt T et al (2018) Heterosis-related genes under different planting densities in maize. J Exp Bot 69:5077–5087. https://doi.org/10.1093/jxb/ery282
Mackay IJ, Cockram J, Howell P, Powell W (2021) Understanding the classics: the unifying concepts of transgressive segregation, inbreeding depression and heterosis and their central relevance for crop breeding. Plant Biotechnol J 19:26–34. https://doi.org/10.1111/pbi.13481
Mahmoud M, Gobet N, Cruz-Dávalos DI, Mounier N, Dessimoz C, Sedlazeck FJ (2019) Structural variant calling: the long and the short of it. Genome Biol 20:246. https://doi.org/10.1186/s13059-019-1828-7
Makhoul M, Chawla HS, Wittkop B, Stahl A, Voss-Fels KP, Zetzsche H et al (2022) Long-amplicon single-molecule sequencing reveals novel, trait-associated variants of vernalization1 homoeologs in hexaploid wheat. Front Plant Sci 13:942461. https://doi.org/10.3389/fpls.2022.942461
Malla S, Brewin DG (2019) Crop research, biotech canola, and innovation policy in Canada: challenges, opportunities, and evolution. Can J Agric Econ Revue Can D'agroeconomie 67:135–150. https://doi.org/10.1111/cjag.12195
Martinez Palacios P, Jacquemot M-P, Tapie M, Rousselet A, Diop M, Remoué C et al (2019) Assessing the response of small RNA populations to Allopolyploidy using resynthesized *Brassica napus* Allotetraploids. Mol Biol Evol 36:709–726. https://doi.org/10.1093/molbev/msz007
Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet 15:394–408. https://doi.org/10.1038/nrg3683
Meinke DW (2020) Genome-wide identification of EMBRYO-DEFECTIVE (EMB) genes required for growth and development in Arabidopsis. New Phytol 226:306–325. https://doi.org/10.1111/nph.16071
Mercé C, Bayer PE, Tay Fernandez C, Batley J, Edwards D (2020) Induced methylation in plants as a crop improvement tool: progress and perspectives. Agronomy 10:1484. https://doi.org/10.3390/agronomy10101484
Meyer RC, Witucka-Wall H, Becher M, Blacha A, Boudichevskaia A, Dörmann P et al (2012) Heterosis manifestation during early Arabidopsis seedling development is characterized by intermediate gene expression and enhanced metabolic activity in the hybrids. Plant J 71:669–683. https://doi.org/10.1111/j.1365-313X.2012.05021.x
Mhiri C, Borges F, Grandbastien M-A (2022) Specificities and dynamics of transposable elements in land plants. Biol Basel 11:448. https://doi.org/10.3390/biology11040488
Miller C, Wells R, McKenzie N, Trick M, Ball J, Fatihi A et al (2019) Variation in expression of the HECT E3 ligase UPL3 modulates LEC2 levels, seed size, and crop yields in *Brassica napus*. Plant Cell 31:2370–2385. https://doi.org/10.1105/tpc.18.00577
Montgomery SA, Berger F (2021) The evolution of imprinting in plants: beyond the seed. Plant Reprod 34:373–383. https://doi.org/10.1007/s00497-021-00410-7
Morin RD, Aksay G, Dolgosheina E, Ebhardt HA, Magrini V, Mardis ER et al (2008) Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. Genome Res 18:571–584. https://doi.org/10.1101/gr.6897308
Muyle AM, Seymour DK, Lv Y, Huettel B, Gaut BS (2022) Gene body methylation in plants: mechanisms, functions, and important implications for understanding evolutionary processes. Genome Biol Evol 14:evac038. https://doi.org/10.1093/gbe/evac038
Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim KD, Li Q et al (2016) Widespread natural variation of DNA methylation within angiosperms. Genome Biol 17:194. https://doi.org/10.1186/s13059-016-1059-0
Novák P, Guignard MS, Neumann P, Kelly LJ, Mlinarec J, Koblížková A et al (2020) Repeat-sequence turnover shifts fundamentally in species with large genomes. Nat Plants 6:1325–1329. https://doi.org/10.1038/s41477-020-00785-x
Nowak K, Morończyk J, Grzyb M, Szczygieł-Sommer A, Gaj MD (2022) miR172 regulates WUS during somatic embryogenesis in Arabidopsis via AP2. Cells 11:718. https://doi.org/10.3390/cells11040718
O'Brien KP, Remm M, Sonnhammer ELL (2005) Inparanoid: a comprehensive database of eukaryotic orthologs. Nucleic Acids Res 33:D476–D480. https://doi.org/10.1093/nar/gki107
Orantes-Bonilla M, Makhoul M, Lee H, Chawla HS, Vollrath P, Langstroff A et al (2022) Frequent spontaneous structural rearrangements promote rapid genome diversification in a Brassica napus F1 generation. Front Plant Sci 13:1057953. https://doi.org/10.3389/fpls.2022.1057953
Pantaleo V, Szittya G, Moxon S, Miozzi L, Moulton V, Dalmay T et al (2010) Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. Plant J No No. https://doi.org/10.1111/j.1365-313X.2010.04208.x
Pantaleo V, Vitali M, Boccacci P, Miozzi L, Cuozzo D, Chitarra W et al (2016) Novel functional microRNAs from virus-free and infected *Vitis vinifera* plants under water stress. Sci Rep 6:20167. https://doi.org/10.1038/srep20167
Plotnikova A, Kellner MJ, Schon MA, Mosiolek M, Nodine MD (2019) MicroRNA dynamics and functions during Arabidopsis embryogenesis. Plant Cell 31:2929–2946. https://doi.org/10.1105/tpc.19.00395
Qian W, Li Q, Noack J, Sass O, Meng J, Frauen M et al (2009) Heterotic patterns in rapeseed (*Brassica napus* L.): II. Crosses between European winter and Chinese semi-winter lines. Plant Breed 128:466–470. https://doi.org/10.1111/j.1439-0523.2008.01597.x
Quezada-Martinez D, Addo Nyarko CP, Schiessl SV, Mason AS (2021) Using wild relatives and related species to build climate resilience in Brassica crops. Theor Appl Genet 134:1711–1728. https://doi.org/10.1007/s00122-021-03793-3
Quezada-Martinez D, Zou J, Zhang W, Meng J, Batley J, Mason AS (2022) Allele segregation analysis of F1 hybrids between independent Brassica allohexaploid lineages. Chromosoma. https://doi.org/10.1007/s00412-022-00774-3
Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033
Ran L, Fang T, Rong H, Jiang J, Fang Y, Wang Y (2016) Analysis of cytosine methylation in early generations of resynthesized *Brassica napus*. J Integr Agric 15:1228–1238. https://doi.org/10.1016/S2095-3119(15)61277-1
Rauluseviciute I, Drabløs F, Rye MB (2020) DNA hypermethylation associated with upregulated gene expression in prostate cancer demonstrates the diversity of epigenetic regulation. BMC Med Genomics 13:6. https://doi.org/10.1186/s12920-020-0657-6

Regmi R, Newman TE, Kamphuis LG, Derbyshire MC (2021) Identification of *B napus* small RNAs responsive to infection by a necrotrophic pathogen. BMC Plant Biol 21:366. https://doi.org/10.1186/s12870-021-03148-6

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. Trends Genet 16:276–277. https://doi.org/10.1016/s0168-9525(00)02024-2

Robinson MD, Davis JM, Gordon KS (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140. https://doi.org/10.1093/bioinformatics/btp616

Rong H, Yang W, Zhu H, Jiang B, Jiang J, Wang Y (2021) Genomic imprinted genes in reciprocal hybrid endosperm of *Brassica napus*. BMC Plant Biol 21:140. https://doi.org/10.1186/s12870-021-02908-8

Růžička K, Zhang M, Campilho A, Bodi Z, Kashif M, Saleh M et al (2017) Identification of factors required for m6 A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. New Phytol 215:157–172. https://doi.org/10.1111/nph.14586

Sablok G, Srivastva AK, Suprasanna P, Baev V, Ralph PJ (2015) isomiRs: increasing evidences of isomiRs complexity in plant stress functional biology. Front Plant Sci 6:949. https://doi.org/10.3389/fpls.2015.00949

Sands B, Yun S, Mendenhall AR (2021) Introns control stochastic allele expression bias. Nat Commun 12:6527. https://doi.org/10.1038/s41467-021-26798-4

Schiessl S, Huettel B, Kuehn D, Reinhardt R, Snowdon R (2017) Post-polyploidisation morphotype diversification associates with gene copy number variation. Sci Rep 7:41845. https://doi.org/10.1038/srep41845

Schiessl SV, Quezada-Martinez D, Orantes-Bonilla M, Snowdon RJ (2020) Transcriptomics reveal high regulatory diversity of drought tolerance strategies in a biennial oil crop. Plant Sci 297:110515. https://doi.org/10.1016/j.plantsci.2020.110515

Scossa F, Alseekh S, Fernie AR (2021) Integrating multi-omics data for crop improvement. J Plant Physiol 257:153352. https://doi.org/10.1016/j.jplph.2020.153352

Seifert F, Thiemann A, Grant-Downton R, Edelmann S, Rybka D, Schrag TA et al (2018a) Parental expression variation of small RNAs is negatively correlated with grain yield heterosis in a maize breeding population. Front Plant Sci 9:13. https://doi.org/10.3389/fpls.2018.00013

Seifert F, Thiemann A, Schrag TA, Rybka D, Melchinger AE, Frisch M et al (2018b) Small RNA-based prediction of hybrid performance in maize. BMC Genomics 19:371. https://doi.org/10.1186/s12864-018-4708-8

Serin Harmanci A, Harmanci AO, Zhou X (2020) CaSpER identifies and visualizes CNV events by integrative analysis of single-cell or bulk RNA-sequencing data. Nat Commun 11:89. https://doi.org/10.1038/s41467-019-13779-x

Serrato-Capuchina A, Matute DR (2018) The role of transposable elements in speciation. Genes Basel 9:254. https://doi.org/10.3390/genes9050254

Shahzad K, Zhang X, Guo L, Qi T, Bao L, Zhang M et al (2020) Comparative transcriptome analysis between inbred and hybrids reveals molecular insights into yield heterosis of upland cotton. BMC Plant Biol 20:1–18. https://doi.org/10.1186/s12870-020-02442-z

Shaikh AA, Chachar S, Chachar M, Ahmed N, Guan C, Zhang P (2022) Recent advances in DNA methylation and their potential breeding applications in plants. Horticulturae 8:562. https://doi.org/10.3390/horticulturae8070562

Shen Y, Sun S, Hua S, Shen E, Ye C-Y, Cai D et al (2017) Analysis of transcriptional and epigenetic changes in hybrid vigor of allopolyploid *Brassica napus* uncovers key roles for small RNAs. Plant J 91:874–893. https://doi.org/10.1111/tpj.13605

Shivaraj SM, Jain A, Singh A (2018) Highly preserved roles of Brassica MIR172 in polyploid Brassicas: ectopic expression of variants of Brassica MIR172 accelerates floral transition. Mol Genet Genomics 293:1121–1138. https://doi.org/10.1007/s00438-018-1444-3

Shull GH (1914) Duplicate genes for capsule-form inBursa bursa-pastoris. Z.Ver-erbungslehre 12:97–149. https://doi.org/10.1007/BF01837282

Shull GH (1948) What is “heterosis”? Genetics 33:439–446. https://doi.org/10.1093/genetics/33.5.439

Smit AFA and Hubley, R. (2008). *RepeatModeler Open-1.0*. http://www.repeatmasker.org. Accessed 2018

Snyman MC, Solofoharivelo M-C, Souza-Richards R, Stephan D, Murray S, Burger JT (2017) The use of high-throughput small RNA sequencing reveals differentially expressed microRNAs in response to aster yellows phytoplasma-infection in *Vitis vinifera* cv. “Chardonnay.” PLoS One 12:e018262. https://doi.org/10.1371/journal.pone.0182629

Stein A, Coriton O, Rousseau-Gueutin M, Samans B, Schiessl SV, Obermeier C et al (2017) Mapping of homoeologous chromosome exchanges influencing quantitative trait variation in Brassica napus. Plant Biotechnol J 15:1478–1489. https://doi.org/10.1111/pbi.12732

Tang X, Bian S, Tang M, Lu Q, Li S, Liu X et al (2012) MicroRNA-mediated repression of the seed maturation program during vegetative development in Arabidopsis. PLoS Genet 8:e1003091. https://doi.org/10.1371/journal.pgen.1003091

ter Steeg EMS, Struik PC, Visser RGF, Lindhout P (2022) Crucial factors for the feasibility of commercial hybrid breeding in food crops. Nat Plants 8:463–473. https://doi.org/10.1038/s41477-022-01142-w

Thiemann A, Meyer S, Scholten S (2009) Heterosis in plants: manifestation in early seed development and prediction approaches to assist hybrid breeding. Chin Sci Bull 54:2363–2375. https://doi.org/10.1007/s11434-009-0326-9

Tzafrir I, Pena-Muralla R, Dickerman A, Berg M, Rogers R, Hutchens S et al (2004) Identification of genes required for embryo development in Arabidopsis. Plant Physiol 135:1206–1220. https://doi.org/10.1104/pp.104.045179

van Hulten MHA, Paulo M-J, Kruijer W, Blankestijn-De Vries H, Kemperman B, Becker FFM et al (2018) Assessment of heterosis in two *Arabidopsis thaliana* common-reference mapping populations. PLoS One 13:e0205564. https://doi.org/10.1371/journal.pone.0205564

Verma S, Attuluri VPS, Robert HS (2022) Transcriptional control of Arabidopsis seed development. Planta 255:90. https://doi.org/10.1007/s00425-022-03870-x

VIB-UGent (2021). *Venn*. https://bioinformatics.psb.ugent.be/webtools/Venn. Accessed 2021

Vitting-Seerup K, Sandelin A (2019) IsoformSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. Bioinformatics 35:4469–4471. https://doi.org/10.1093/bioinformatics/btz247

Vollrath P, Chawla HS, Schiessl SV, Gabur I, Lee H, Snowdon RJ et al (2021) A novel deletion in FLOWERING LOCUS T modulates flowering time in winter oilseed rape. Theor Appl Genet 134:1217–1231. https://doi.org/10.1007/s00122-021-03768-4

Wang M, Wang J (2022) Transcriptome and DNA methylome analyses provide insight into the heterosis in flag leaf of inter-subspecific hybrid rice. Plant Mol Biol 108:105–125. https://doi.org/10.1007/s11103-021-01228-7

Wang L, Wang M-B, Tu J-X, Helliwell CA, Waterhouse PM, Dennis ES et al (2007) Cloning and characterization of microRNAs from

Brassica napus. FEBS Lett 581:3848–3856. https://doi.org/10.1016/j.febslet.2007.07.010
Wang L, Wu LM, Greaves IK, Zhu A, Dennis ES, Peacock WJ (2017a) PIF4-controlled auxin pathway contributes to hybrid vigor in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A 114:E3555–E3562. https://doi.org/10.1073/pnas.1703179114
Wang P, Yang C, Chen H, Song C, Zhang X, Wang D (2017b) Transcriptomic basis for drought-resistance in *Brassica napus* L. Sci Rep 7:40532. https://doi.org/10.1038/srep40532
Wang Z, Qiao Y, Zhang J, Shi W, Zhang J (2017c) Genome wide identification of microRNAs involved in fatty acid and lipid metabolism of *Brassica napus* by small RNA and degradome sequencing. Gene 619:61–70. https://doi.org/10.1016/j.gene.2017.03.040
Wang Z, Wu X, Wu Z, An H, Yi B, Wen J et al (2018) Genome-wide DNA methylation comparison between *Brassica napus* genic male sterile line and restorer line. Int J Mol Sci 19:2689. https://doi.org/10.3390/ijms19092689
Wei W, Li G, Jiang X, Wang Y, Ma Z, Niu Z et al (2018) Small RNA and degradome profiling involved in seed development and oil synthesis of Brassica napus. PLoS One 13:e0204998. https://doi.org/10.1371/journal.pone.0204998
Wei Y, Li G, Zhang S, Zhang S, Zhang H, Sun R et al (2021) Analysis of transcriptional changes in different *Brassica napus* synthetic Allopolyploids. Genes Basel 12:82. https://doi.org/10.3390/genes12010082
Weigel D, Colot V (2012) Epialleles in plant evolution. Genome Biol 13:249. https://doi.org/10.1186/gb-2012-13-10-249
Wójcik AM, Nodine MD, Gaj MD (2017) miR160 and miR166/165 contribute to the LEC2-mediated auxin response involved in the somatic embryogenesis induction in Arabidopsis. Front Plant Sci 8:2024. https://doi.org/10.3389/fpls.2017.02024
Wu J, Zhao Q, Yang Q, Liu H, Li Q, Yi X et al (2016) Comparative transcriptomic analysis uncovers the complex genetic network for resistance to *Sclerotinia sclerotiorum* in *Brassica napus*. Sci Rep 6:19007. https://doi.org/10.1038/srep19007
Wu J, Lin L, Xu M, Chen P, Liu D, Sun Q et al (2018) Homoeolog expression bias and expression level dominance in resynthesized allopolyploid *Brassica napus*. BMC Genomics 19:586. https://doi.org/10.1186/s12864-018-4966-5
Wu X, Liu Y, Zhang Y, Gu R (2021) Advances in research on the mechanism of heterosis in plants. Front Plant Sci 12:745726. https://doi.org/10.3389/fpls.2021.745726
Xu J, Chen G, Hermanson PJ, Xu Q, Sun C, Chen W et al (2019) Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize. Genome Biol 20:1–16. https://doi.org/10.1186/s13059-019-1859-0
Xu K, Wu Y, Song J, Hu K, Wu Z, Wen J et al (2021) Fine mapping and identification of BnaC06.FtsH1, a lethal gene that regulates the PSII repair cycle in *Brassica napus*. Int J Mol Sci 22:2087. https://doi.org/10.3390/ijms22042087
Yang M, Wang X, Ren D, Huang H, Xu M, He G et al (2017) Genomic architecture of biomass heterosis in Arabidopsis. Proc Natl Acad Sci 114:8101–8106. https://doi.org/10.1073/pnas.1705423114
Yang K, Wen X, Mudunuri S, Varma GPS, Sablok G (2019) Diff isomiRs: large-scale detection of differential isomiRs for understanding non-coding regulated stress omics in plants. Sci Rep 9:1406. https://doi.org/10.1038/s41598-019-38932-w
Yang Y, Saand MA, Huang L, Abdelaal WB, Zhang J, Wu Y et al (2021) Applications of multi-omics technologies for crop improvement. Front Plant Sci 12:563953. https://doi.org/10.3389/fpls.2021.563953
Yao S, Liang F, Gill RA, Huang J, Cheng X, Liu Y et al (2020) A global survey of the transcriptome of allopolyploid *Brassica napus* based on single-molecule long-read isoform sequencing and Illumina-based RNA sequencing data. Plant J 103:843–857. https://doi.org/10.1111/tpj.14754
Yi L, Liu B, Nixon PJ, Yu J, Chen F (2022) Recent advances in understanding the structural and functional evolution of FtsH proteases. Front Plant Sci 13:837528. https://doi.org/10.3389/fpls.2022.837528
Yin H, Fan Z, Li X, Wang J, Liu W, Wu B et al (2016) Phylogenetic tree-informed microRNAome analysis uncovers conserved and lineage-specific miRNAs in Camellia during floral organ development. J Exp Bot 67:2641–2653. https://doi.org/10.1093/jxb/erw095
Yin L, Zhu Z, Huang L, Luo X, Li Y, Xiao C et al (2021) DNA repair- and nucleotide metabolism-related genes exhibit differential CHG methylation patterns in natural and synthetic polyploids (*Brassica napus* L.). Hortic Res 8:142. https://doi.org/10.1038/s41438-021-00576-1
Yoo M-J, Szadkowski E, Wendel JF (2013) Homoeolog expression bias and expression level dominance in allopolyploid cotton. Hered Edinb 110:171–180. https://doi.org/10.1038/hdy.2012.94
Yu D, Gu X, Zhang S, Dong S, Miao H, Gebretsadik K et al (2021) Molecular basis of heterosis and related breeding strategies reveal its importance in vegetable breeding. Hortic Res 8:120. https://doi.org/10.1038/s41438-021-00552-9
Zanini SF, Bayer PE, Wells R, Snowdon RJ, Batley J, Varshney RK et al (2022) Pangenomics in crop improvement-from coding structural variations to finding regulatory variants with pangenome graphs. Plant Genome 15:e20177. https://doi.org/10.1002/tpg2.20177
Zemach A, Kim MY, Silva P, Rodrigues JA, Dotson B, Brooks MD et al (2010) Local DNA hypomethylation activates genes in rice endosperm. Proc Natl Acad Sci U S A 107:18729–18734. https://doi.org/10.1073/pnas.1009695107
Zhang L, Peng Y, Wei X, Dai Y, Yuan D, Lu Y et al (2014a) Small RNAs as important regulators for the hybrid vigour of super-hybrid rice. J Exp Bot 65:5989–6002. https://doi.org/10.1093/jxb/eru337
Zhang M, Xie S, Dong X, Zhao X, Zeng B, Chen J et al (2014b) Genome-wide high resolution parental-specific DNA and histone methylation maps uncover patterns of imprinting regulation in maize. Genome Res 24:167–176. https://doi.org/10.1101/gr.155879.113
Zhang J, Wei L, Jiang J, Mason AS, Li H, Cui C et al (2018) Genome-wide identification, putative functionality and interactions between lncRNAs and miRNAs in *Brassica* species. Sci Rep 8:4960. https://doi.org/10.1038/s41598-018-23334-1
Zhang Y, Liu C, Cheng H, Tian S, Liu Y, Wang S et al (2020) DNA methylation and its effects on gene expression during primary to secondary growth in poplar stems. BMC Genomics 21:1–18. https://doi.org/10.1186/s12864-020-06902-6
Zhang L, He J, He H, Wu J, Li M (2021a) Genome-wide unbalanced expression bias and expression level dominance toward *Brassica oleracea* in artificially synthesized intergeneric hybrids of *Raphanobrassica*. Hortic Res 8:246. https://doi.org/10.1038/s41438-021-00672-2
Zhang Q, Guan P, Zhao L, Ma M, Xie L, Li Y et al (2021b) Asymmetric epigenome maps of subgenomes reveal imbalanced transcription and distinct evolutionary trends in *Brassica napus*. Mol Plant 14:604–619. https://doi.org/10.1016/j.molp.2020.12.020
Zhou P, Hirsch CN, Briggs SP, Springer NM (2019) Dynamic patterns of gene expression additivity and regulatory variation throughout maize development. Mol Plant 12:410–425. https://doi.org/10.1016/j.molp.2018.12.015
Zhu A, Wang A, Zhang Y, Dennis ES, Peacock WJ, Greaves AIK (2020) Early establishment of photosynthesis and auxin biosynthesis plays a key role in early biomass heterosis in *Brassica*

napus (Canola) Hybrids. Plant Cell Physiol 61:1134–1143. https://doi.org/10.1093/pcp/pcaa038

Ziegler DJ, Khan D, Kalichuk JL, Becker MG, Belmonte MF (2019) Transcriptome landscape of the early *Brassica napus* seed. J Integr Plant Biol 61:639–650. https://doi.org/10.1111/jipb.12812

Zou J, Hu D, Mason AS, Shen X, Wang X, Wang N et al (2018) Genetic changes in a novel breeding population of *Brassica napus* synthesized from hundreds of crosses between *B. rapa* and *B. carinata*. Plant Biotechnol J 16:507–519. https://doi.org/10.1111/pbi.12791

Zrimec J, Börlin CS, Buric F, Muhammad AS, Chen R, Siewers V et al (2020) Deep learning suggests that gene expression is encoded in all parts of a co-evolving interacting gene regulatory structure. Nat Commun 11:6141. https://doi.org/10.1038/s41467-020-19921-4

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

4. Discussion

4.1 Spontaneous rearrangements and genomic diversity

Genomic rearrangements are a main driver of species and population diversification (Yeaman 2013; Stange et al. 2021); however, the frequency of spontaneous *de novo* rearrangements has not been extensively assessed in polyploid plants. Somatic mutations have been reported to be higher in annual plants, based on whole sequencing of more than 700 plant genomes (Wang et al. 2019a), while pedigree analyses in *A. thaliana* over 20 generations found that somatic spontaneous mutations occur mostly and randomly in gene bodies (Monroe et al. 2022). Contrasting to somatic mutations, meiotic derived variants that occur prior to zygote formation have been observed in the form of paramutations, epigenetic alleles, gene gain/loss and selfish-genetic elements in maize, Arabidopsis, peas, northern barley grass and rice (Lolle et al. 2005; Yu et al. 2018; Adu-Yeboah et al. 2021; Bente et al. 2021; Pereira and Leitão 2021; Cao et al. 2022). Previous studies in allotetraploid and allohexaploid *B.napus* using SNP arrays have also demonstrated large scale *de novo* rearrangements in F_1 offspring which were undetected in paternal genotypes (Higgins et al. 2018; Quezada-Martinez et al. 2022).

The large-scale rearrangements reported in *B. napus* in this thesis might not be totally unexpected, since crosses between genetically divergent oilseed cultivars are prone to generate diversity due to multivalent pairing between homoeologous chromosomes and genomic rearrangements (Zou et al. 2010; Higgins et al. 2018; Hu et al. 2021). However, the results in Chapter 2 extend such previous findings to F_1 sister plants from a single cross between two highly homozygous single plants using less genetic markers, higher resolution from rearrangements breakpoints and analyzing high repetitive regions through long read sequencing. Specific associations between spontaneous genomic rearrangements and phenotype are hard to evaluate due to lack of biological replicates having the same genomic rearrangements. Nevertheless, a plant with a large-scale NHRE between chromosomes C08 and A09 showed high resemblance to the maternal line Express 617 in terms of height, digital

biomass and leaf area. Based on long read sequencing data and subgenomic homology, it was identified that chromosome C08 from the paternal line G3D001 had been deleted and replaced by a chromosome A09 segment from the maternal genotype. The same paternal inheritance pattern was observed in all large-scale rearrangements which suggests that the variants likely occurred during the meiosis of the non-traditional paternal line despite its high homozygosity. This result is altogether not highly unexpected since genotypes with exotic backgrounds tend to show variable chromosome number, pairing and instability (Szadkowski et al. 2010; Xiong et al. 2011; Ferreira de Carvalho et al. 2021). Nevertheless, the implications of the results for early and rapid genomic diversification remain fundamental to understanding potential post-polyploidisation scenarios as well as for mining hidden genomic diversity.

Polyploidization has been proposed as a speciation and adaptation mechanism. Nonetheless, the events promoting fitness of polyploid species despite genomic instability and reproductive isolation are not fully understood (van de Peer et al. 2017; Pelé et al. 2018; Hörandl 2022). Structural rearrangements are closely connected to polyploid formation and adaptation, since rearrangements promote gene copy number variation, gene neofunctionalization and genomic diversification (Chen and Ni 2006; Otto 2007). It has been estimated that at least 35% of flowering plants have a polyploid origin with varying geographical distributions and densities worldwide (Wood et al. 2009; Rice et al. 2019). Moreover, a large range of structural variants have been identified in plants, including polyploid species, as reviewed by Zhang et al. (2018) and Schiessl et al. (2019b). Mid-sized structural variation studies have already revealed subpopulation divergence in polyploid species like wheat, cotton and oilseed rape (Walkowiak et al. 2020; Chawla et al. 2021; He et al. 2021). Nonetheless, the detection of mid-sized SVs *de-novo* mutations not present in the parents in polyploids is still limited due to large, repetitive and complex genomes that limit the differentiation between extremely similar sister lines (Belyeu et al. 2021; Yuan et al. 2021). Overall, few studies have evaluated spontaneous

genomic variants accelerating genomic diversity in one single generation. The proof-of-concept delivered in this thesis demonstrates the potential of NGS technologies in finding frequent spontaneous mutations that might aid in understanding post-polyploidization and hybrid divergence in future studies using larger and more diverse germplasm collections.

The impact of the spontaneous rearrangements observed in Chapter 2 also had implications in gene CNV that included flowering time genes, disease susceptibility and developmental functions. Changes in the number of homologs have been found to have profound effects on diseases response and flowering timing in oilseed rape (Schiessl et al. 2017; Gabur et al. 2020), and this kind of event also have implications for genome dosage (Lloyd et al. 2018; Birchler and Veitia 2021). The putative effect of the spontaneous large-scale rearrangements found between the Express 617 x G3D001 F_1 sister plants were additionally examined through methylation analyses from long read sequencing. As expected, regions with large-scale deletions displayed less methylated cytosines. However, the methylation levels were maintained overall despite the rearrangements, suggesting a mechanism where methylation is conserved despite large genomic rearrangements. Further studies with comprehensive gene expression and methylation datasets using large populations might aid to determine the mechanisms and roles of genome-wide methylation in loci with structural variants.

It is worthy to mention that since the commercialization of the first long reads sequencing platforms by PacBio in 2011 and by ONT in 2014, their functionalities have greatly expanded and enabled long read sequencing-based studies in methylation prediction, single cell sequencing, genome assembling and transcriptomics (Amarasinghe et al. 2020; Lebrigand et al. 2020). As long read sequencing improves and new technologies emerge, it is hoped that the detection of spontaneous rearrangements in complex crop genomes will become more feasible. Furthermore, exploring heterotic aspects outside the genomic scope such as RNA and methylation can bring additional novel breakthroughs in plant research.

4.2 Gene regulation in seed development

Early stage heterosis has played critical roles by triggering seed yield, biomass and early development and bringing advantages in overall performance in hybrid plants. For instance, inherited parental dominance in Arabidopsis was found to enhance seedling growth through genetic enrichment of cell cycle and photosynthetic functions (Liu et al. 2021b). In maize, seedling hybrids from cultivars B73 and Mo17 displayed remarkable non-additive gene expression linked with epigenetic features, having potential heterotic impacts (Luo et al. 2021). Similar examples are recorded also in oilseed rape and other Brassica species where early heterosis is shown in paternal dominant or hybrid-specific transcriptomic patterns (Wei et al. 2021; Zhang et al. 2021; Xiong et al. 2022).

The advantages of RNA-Seq have facilitated analyses of gene expression and genetic networks in large and diverse genotype panels by evaluating genome-wide transcripts and providing higher resolution transcriptomic profiling (An et al. 2019; Tan et al. 2022). Differential gene expression has dissected ontologically enriched genes with photosynthetic and metabolism regulation functions linked to early stage heterosis (Song et al. 2010; Shahzad et al. 2020; Wu et al. 2021a). In Chapter 3, differentially expressed genes associated to embryo development, photosynthesis and reproduction displayed transient patterns where the hybrid expression did not follow paternal trends during early seed formation, suggesting implications for early heterosis. In addition, more differential genes were observed in selfed F_1 plants in comparison to hybrids pollinated by outcrossing Express 617 and G3D001 during the experiment. Thus, indicating crucial differences between hybrid formation after parental outcrossing and hybrid selfing, and enabling a more refined heterotic gene dissection.

The role of miRNAs in plant development has been shown to be quite broad and cover multiple functions with consequences in crop heterosis (Dong et al. 2022b; Tang and Chu 2017). microRNAs triggering early stage heterosis have been reported in maize, Arabidopsis and

oilseed rape where miRNAs regulated germination, size, storability and fatty acid composition in seeds (Ding et al. 2012; Wang et al. 2016; Castillo-Bravo et al. 2022; Song et al. 2022). mRNA and miRNA interactions in Chapter 3 included miRNA families that have also been associated with organelle and embryo development in plants. For instance, miR172 and miR169 were found to regulate expression of genes involved in embryo development during seed formation in our studies, similarly as reported by Meinke (2020) and Nowak et al. (2022). Furthermore, miR165/166 families which are known to target the class III homeodomain leucine zipper (HD-ZIP III) transcription factors like *PHAVOLUTA* (*PHV*) and *PHABULOSA* (*PHB*), which were differentially expressed in our studies during hybrid seed development. Both of those genes indirectly regulate *LEAFY COTYLEDON 2* (*LEC2*), which encodes a DNA-binding protein that regulates embryo formation and seed size (Braybrook et al. 2006; Miller et al. 2019).

Seed development is characterized by the tissue differentiation of the endosperm, embryo and seed coat. Interestingly, the endosperm is characterised by methylation patterns that lead to the formation of epialleles via methylation and demethylation of parental gametes. For example, the results in Chapter 3 suggest that epialleles are associated with various functions such as photosynthesis, DNA transcription and lipid and carbohydrate metabolisms. In rice, corn and Arabidopsis, epialleles have been linked to enhanced performances through various generations (Greaves et al. 2015; Cao et al. 2022; Wang and Wang 2022). The endosperm has an unbalanced genome dosage due to double fertilization in plants, which also makes transcriptomic and methylation patterns valuable assets for research. siRNAs in Chapter 3 were mostly found in gene flanking regions that showed no differential expression, suggesting a role in regulating key genetic functions during seed formation as also observed by Lu et al. (2012) and Kirkbride et al. (2019). siRNAs are characterized by regulating genes through transposons methylation. Hence, genome-wide methylome profiling can aid to identify potential regulatory regions.

4.3 Transcriptomic and epigenomic patterns in early plant developmental stages

In Chapter 3, global genome-wide methylation patterns were associated with maternal dominant patterns, suggesting diverging roles between parents during seed development as reviewed extensively by Batista and Köhler (2020). Interestingly, these patterns deviated from female hypomethylation that is generally observed in hybrids of other plant species (Gehring et al. 2009; Zemach et al. 2010; Zhang et al. 2014). Instead, they revealed maternal hypermethylation in both outcrossed parents hybrids and selfed hybrids plants. Similar maternal methylation trends have been observed in other Brassica species (Liu et al. 2018; Grover et al. 2020). The latter authors proposed that a significant expression of siRNAs in the seed coat during seed formation in the seed could promote DNA methylation. Most of the differentially expressed siRNAs reflected maternal dominance in hybrids and were in flanking regions of non-differentially expressed genes, suggesting roles in maintaining essential genes functions and, possibly, regulating gene expression bias in repetitive allopolyploid genomes. Furthermore, maternal and paternal epialleles found in Chapter 3 during seed formation were linked to growth, development and metabolism transport functions. This highlights the range of implications that can be altered without direct genomic rearrangements. Evaluating clues outside the genome can be beneficial to find key loci and introgress variety in populations facing bottlenecks through transcriptomic and epigenomic profiling (Crisp et al. 2020).

Transcriptomic paternal patterns have also contributed to deciphering key loci in early plant developmental stages (Le et al. 2010; Das et al. 2015). In our study, the majority of DEGs displayed maternal expression dominance across distinct stages and tissues based on expression level dominance (ELD) analysis. EDL studies in polyploids like oilseed rape and cotton have also revealed mostly parental specific dominance rather than prominent transgressive expression (Yoo et al. 2013; Wu et al. 2018; Li et al. 2020b; Wei et al. 2021; Xiong et al. 2022). Remarkably, ELD and GO enrichment analysis revealed that transgressively upregulated genes

during early seed formation in Chapter 3 were linked to reproductive and developmental stage, with potential implications in early heterosis. Heterosis, however, is not confined to a unique paternal-offspring EDL classification pattern, as observed in differentially expressed miRNA families targeting key embryo and developmental genes through maternal dominance patterns. In addition, it was observed that 12-18% of all differentially expressed or methylated features displayed maternal dominance patterns. Although confirmation in larger germplasm collections with diverse maternal and paternal lines would be needed to validate such findings, the results provide a first indication for maternal specific effects influencing the transcriptome and epigenome landscape during seed formation. Similar observations were previously recorded in Arabidopsis (Lu et al. 2012; Grover et al. 2020).

Despite polyploids representing a more challenging network due to subgenome homologs and transcriptomic shock (Ng et al. 2012; Shin et al. 2022), bioinformatic and sequencing breakthroughs have, to a large extent, refined gene expression research in polyploids. This has turned, for instance, into an accelerated elucidation of diverse RNA types regarding performance, for example via fine mapping of flowering genes inside QTL-like segments or identification of genes regulating fatty acid seed content, seed weight and harvest index in oilseed rape (Wang et al. 2016; Dong et al. 2022a; Han et al. 2022; Zhang et al. 2022). Moreover, sRNAs interactions with targets genes in *B. napus* have revealed specific gene networks influencing drought, flowering and leaf and embryo development which can ultimately affect performance (Zhao et al. 2012; Jain et al. 2018; Wang et al. 2019b; Schiessl et al. 2020). In Chapter 3, thoroughly-annotated miRNA families such as miR166, mir169 and miR172 were found to potentially regulate *APETALA2-like* (*AP2-like*), *EMB* (*EMBRYO DEFECTIVE*) and HD-zip III genes during seed formation. Thus, expanding the repertoire of functional transcriptomic networks observed in *B. napus*.

Methylation profiling is another relevant strategy for detecting regulatory regions. Based on short and long reads in Chapters 2 and 3, it was found that the number of methylated cytosines was higher in the CHH context, while the methylation levels were higher in the CpG and CHG context in oilseed rape genotypes. This reflects similar previous findings in the same species. The regulation of each methylation context not only diverges between species (Niederhuth et al. 2016; Bartels et al. 2018) but also depends, to an extent, on specific regulating genes. For instance, CpG methylation is maintained by *METHYLTRANSFERASE 1* (*MET1*), CHG methylation is mainly regulated by *CHROMOMETHYLASE 3* (*CMT3*) and CHH methylation is retained by *DOMAINS REARRANGED METHYLTRANSFERASES* (*DRMs*) (see Stroud et al. 2013) . Genomic variants causing functional changes in these genes can lead to contrasting findings, as evidenced by increased dysregulated genome expression and seedling lethality when altering *MET1* in rice (Hu et al. 2014) or improved genome editing mutagenesis when using *CMT3* mutant lines in Arabidopsis (Weiss et al. 2022). However, not all genomic variants lead to clear patterns, as demonstrated in green foxtail (*Setaria viridis)*, where predominant methylation loss of CHH context occurred when altering *DRM* genes but no major differential gene expression was detected when low CHH methylation was present in nearby flanking genes in comparison to genotypes with higher CHH methylation levels (Read et al. 2022).

Similar results were observed in Chapter 3, regardless of the methylation context, where gene flanking regions which were significantly methylated were mostly in non-differentially expressed genes. This suggests a conservatory role in maintaining key genetic functions across all genotypes. Furthermore, not all methylation inside gene bodies led to a direct downregulation as normally expected, similarly to other studies reporting proportional gene hypermethylation and upregulation in human and mice cells (Arechederra et al. 2018; Rauluseviciute et al. 2020). Moreover, a high percentage of differentially methylated regions (DMRs) were within repeat motifs (43%) with the majority of them (70%) located in 5kbp

flanking regions as in other reports (Rong et al. 2021; Cao et al. 2022). Repetitive regions, including transposable elements, are a source of sRNAs that regulate transcriptomic expression (Nuthikattu et al. 2013; Ni et al. 2021; Ramakrishnan et al. 2021). In our study, abundant differentially expressed siRNAs co-localized in methylated flanking regions of non-differentially expressed genes which suggest a conservative regulatory role. The role of siRNAs in seed development is critical due to its involvement with RdDM and transcriptomic networks, as demonstrated previously in Brassicaceae species (Mosher and Melnyk 2010; Grover et al. 2018). A prevailing dilemma is that TEs are silenced by siRNAs but at the same time can be a source of TEs (McCue et al. 2015). Thus, refined spaciotemporal omics research would help in dissecting siRNA-TE interactions in the context of gene expression in polyploids.

Lastly, 392 putative epialleles involved in photosynthesis, DNA transcription and lipid and carbohydrate metabolism during seed formation and seedling development were identified based on differential gene methylation and expression in Chapter 3. Potential applications related to epialleles and methylation have been extensively described in previous literature (Kalisz and Purugganan 2004; Crisp et al. 2020; Mercé et al. 2020; Shaikh et al. 2022). Furthermore, research on inheritance of epigenomic features in plants has revealed that methylation is mostly stable across generations with few spontaneous epialleles occurring (Hofmeister et al. 2017) and infrequent epigenomic memory recalling events during stress (Crisp et al. 2016). Despite their scarcity, epialleles have displayed phenotypical differences in flowering time, disease resistance and pigmentation as well as heterotic patterns in maize, rice and Arabidopsis (Greaves et al. 2015; Cao et al. 2022; Wang and Wang 2022). Epigenomic profiling provides an opportunity to mine epigenome diversity and compare results to transcriptomic studies (Xu et al. 2019), yet the relationship to genomic rearrangements still needs to be further explored.

4.4 Effects of structural variants in gene regulation and methylation

Structural variants are credited with promotion of gene neo- and subfunctionalization in oilseed rape and other crops, thus leading to diversification and population divergence with crucial roles in polyploidisation, genomic imbalance and transcriptomic networks (Panchy et al. 2016; Xia et al. 2016; He et al. 2017b; Schiessl 2020). Among structural rearrangements, genomic duplications and deletions deserve special attention since they promote chromosome-scale transcriptomic effects, as clearly presented by Birchler and Yang (2022) with examples on gene hypofunctionalization (downregulated homologs), subfunctionalization and neofunctionalization. The same authors noticed that ancient whole genome duplication (WGD) events, that have occurred in most angiosperms, have led to a retention of genes coding for transcription factors and signalling components, whereas short segmental duplications tend to show less enrichment of such genes. In Chapter 2, large-scale rearrangements resulted in CNV either though NRHE or segmental deletions with enrichment of auxin metabolism, oxidation processes, histone methylation, cell organelles and enzyme activity. The CNVs also included gene homologs involved in flowering time, stress response and plant growth such as *FLOWERING LOCUS C* (*FLC*), *TERMINAL FLOWER 1* (*TFL1*), *WRKY DNA-binding protein 4* (*WRKY-4*) as well as *EMB*, *Arabidopsis heat shock promoter* (ATHSP) and *DWARF protein* (*DWARF*) genes. This indicates potential transcriptomic effects caused by spontaneous mutations in a single generation.

Effects of large-scale rearrangements on the epigenome were assessed by predicting methylation from long read sequencing data in Chapter 2, as also employed in Kirov et al. (2021) and Naish et al. (2021). CHH methylation levels were higher than other methylation contexts and the number of methylated cytosines (mC) was higher in the CpG and CHG contexts, as reported previously in other oilseed rape studies (Shen et al. 2017; Bartels et al. 2018). Interestingly, and regardless of the methylation context and genomic rearrangements,

the methylation levels were comparable across all F_1 sister plants. Likewise, similar findings were observed in Chapter 3, where a segmental deletion in C01 in the paternal genotype did not translate into diverging methylation profiles between genotypes during seed formation. This partially suggests buffering mechanisms or networks that aim at balancing methylation levels despite rearrangements, although much more detailed research is needed for proper validation. Similar complex and unclear scenarios were assessed when investigating the correlation between genome size and methylation levels among plant species. WGD events increase genome size and repetitive content, which in theory should result in higher methylation level. However, that has been reported to not always be the case in large genomes like faba bean (*Vicia faba,* 13 Gbp, Ellwood et al. 2008) and European mistletoe (*Viscum album*, ca. 88 Gbp, Novák et al. 2020), which have much lower methylation levels than smaller genomes like sunflower (*Helianthus annuus*, 3.6 Gbp, Badouin et al. 2017) and pearl millet (*Pennisetum glaucum*, 1.8 Gbp, Varshney et al. 2017). Furthermore, the genome complexity and methylation relationship might depends on methylation contexts (Alonso et al. 2015; Niederhuth et al. 2016). In coming years, it is expected that advances in NGS technologies, molecular biology and bioinformatics will contribute to clarifying such questions in complex plant genomes, as already applied in human genetics (Shi et al. 2020; Hawe et al. 2022).

The results in Chapter 3 also revealed contrasting segmental gene expression covering large genomic regions in genotypes without large-scale deletions (e.g. chromosomes A03, A07 and C05). Mechanisms similar to long range epigenetic silencing (LRES) and systemic acquired silencing (SAS), or eQTL might induce large-scale transcriptomic patterns (Clark 2007; Druka et al. 2010; Rajeevkumar et al. 2015). It is also probable that chromatin interactions and genome spatial-time organization might define large-scale segmental expression (Doğan and Liu 2018; Takei et al. 2021). Hence, integrated omics experiments can provide insights into the transcriptome and epigenome of oilseed rape (Hu et al. 2022; Li et al. 2022; Shin et al. 2022).

4.5 Conclusion

Hybrid systems represents an important breeding target in agriculture for many crops, including oilseed rape (Steeg et al. 2022). Therefore, the aim of the present work was to detect hybrid specific, putatively heterotic multi-omics patterns in oilseed rape hybrids. The allopolyploid nature of oilseed rape is linked to multiple genomic rearrangements that can lead to pronounced diversification. Long read sequencing was used to detect early and rapid variation among F_1 plants with profound effects on gene CNV in a single generation. This provided plausible insights into post-polyploidisation mechanisms for rapid adaptation and diversification in a young allopolyploid species. Widespread rearrangements were found to diverge between F_1-sister lines, expanding similar observations in hybrid populations by Higgins et al. (2018) and Quezada-Martinez et al. (2022).

Structural variants can result in agronomically important phenotypes and genomic diversification (Yuan et al. 2021). Although their occurrence in oilseed rape across generations has been evaluated previously (Bird et al. 2021a; Ferreira de Carvalho et al. 2021), further research might still shed light on polyploid-specific patterns and post-polyploidization. For instance, it was recently shown in cotton that allopolyploids can accumulate deleterious mutations of diverse lengths faster than diploids due to their homology between subgenomes (Conover and Wendel 2022). Interestingly, polyploidy is mostly linked to plants with longer life cycles, regions with low competition with diploids and zones located far from the equator (Rice et al. 2019). Novel pangenomic and ecogeographical phylogenetic research, associating how mutation rates differ between plants based on ploidy levels, would be beneficial to identify the role of de-novo mutations in post-polyploidization diversification.

Transcriptomic and epigenomic networks are also key drivers of heterosis during early developmental stages such as seed formation and seedling development. Parental and hybrid gene expression patterns in Chapter 3 revealed genes regulating embryo and reproductive

patterns during seed formation, along with miRNA associations that have putative implication in early heterosis. The seed is characterized to be a hotspot for epigenomic and transcriptomic activity, as demonstrated through more elaborated experiments using laser dissection in oilseed rape seeds (Ziegler et al. 2019; Khan et al. 2022). Although the use of refined dissection approaches was out of the scope in the present study, global transcription and methylation patterns were found which included putatively inherited epigenetic alleles and identification of methylated regions. Genes which were not differentially expressed between genotypes had mostly differential methylation levels on their flanking regions, thus indicating the role of methylation in gene regulation. The potential of methylation in breeding is broad, with phenotypic implications on flowering time, disease susceptibility and abiotic stress responses (Mercé et al. 2020; Shaikh et al. 2022). Unmethylated regions (UMRs) also represent relevant areas for detailed future studies, as they generally correlate with accessible transcription sites and cis-regulatory elements (Crisp et al. 2020). Recent experiments evaluating heterosis in genetically identical *A. thaliana* F_1 sister plants revealed significant differences in leaf area and rosette size (Mehraj et al. 2020), suggesting that epigenetic factors might also be involved in diverging variation.

Further advances in bioinformatics and biology will likely provide considerably more insight into polyploidy and heterosis in the next decades. After the sequencing of the first plant genome in 2004, more than 1040 assembled plant genomes were registered in the plaBi database of sequenced accessions by September 15, 2022 (FJZ and HHU 2014). However, most assembled genomes to date are diploids (Marks et al. 2021) and the correlations of TE and methylation levels with genome sizes are not fully comprehended (Niederhuth et al. 2016; Novák et al. 2020). During the present work, additional field trials and draft genome assembling of the hybrid paternal line with long reads were carried out. The latter being currently under further

improvement to provide more insights on genome rearrangements and multiomics features in oilseed rape.

The range of variability surrounding the central dogma of genetics makes the integration of multiomics techniques necessary to elucidate transcriptomic and epigenomic networks (Buccitelli and Selbach 2020). In the present work, all omics sequencing approaches were performed from the same biological replicates in all experiments, optimizing the identification of multiomic interactions. Recent bioinformatic approaches are allowing the prediction of omic features and real-time interaction multiomics integration (MOI) analyses (Argelaguet et al. 2018; Seifert et al. 2018b; Jamil et al. 2020; Knoch et al. 2021).

Technological advancements in NGS and informatics have driven biological breakthroughs in biology, yet the plant kingdom has relatively lagged due to more complex genomes with multiple ploidies and repetitive content that require more bioinformatics resources. This might soon change due to the extremely rapid technological advancements developed during the recent pandemic, where scientific work fostered large-scale international collaborations and breakthroughs in bioinformatics and big data management so that petabyte-scale and large nucleic acid datasets could be handled in a cost-time efficient manner that has not been observed before (Maher and van Noorden 2021; Khare et al. 2021; Edgar et al. 2022). Plant breeding can apply similar big data methodology, integrating phenomic, genomic and environmental data following FAIR (Findable, Accessible, Interoperable, Reusable) protocols in open-source platforms like the Breedbase (Wilkinson et al. 2016; Morales et al. 2022; Schmidt and Hildebrandt 2017). The outlook on bioinformatic supported breeding approaches, which include diverse sequencing technologies coupled with deep epigenomic and transcriptomic profiling, seem promising strategies to elucidate the role of non-coding genomic loci and mine for multiomics diversity with relevant applications in modern agriculture.

5. Summary

5. Summary

Oilseed rape (*Brassica napus*) is the third-most widely grown crop globally for vegetable oil production after soybean and oil palm. The relevance of oilseed rape relies not only in its economic value but also in its allopolyploidy and close linkage to *Arabidopsis thaliana*, which make it a useful model crop for other allopolyploids like wheat, cotton and sugarcane. Part of the increase of *B. napus* yield is due to enhanced performance in F_1 hybrids, through exploitation of heterosis during the last few decades in all major oilseed rape production areas in the world. Until now, heterosis and hybrid-specific patterns have been mostly examined under allelic interaction hypotheses. Therefore, this thesis focuses on assessing genomic, transcriptomic and epigenomic features in the context of heterosis in a hybrid cross between two highly homozygous, yet genetically divergent, inbred oilseed rape accessions. Spontaneous large-scale genomic rearrangements that differed between F_1 sister plants generated early and rapid gene copy number variation in a single generation. The implications of such preliminary findings suggest possible scenarios of post-polyploidization mechanisms to survive the evolutionary dead end of allopolyploidization. Furthermore, gene expression, small RNA and genome-wide methylation during seed and seedling development in the F_1-hybrid were evaluated. Parental dominance was observed in transcriptomic and epigenomic patterns and key differentially expressed genes involved in reproduction and development during seed formation were detected through gene expression and microRNA analyses. Candidate epigenetic alleles with diverse core biological functions were identified and linkages between epigenomic and transcriptomic features were assessed. The insight provided through this research aims at exploring molecular patterns with potential implications in hybrids and heterosis and establishing frameworks for future multi-omics and pangenomic research in polyploid crops.

6. Zusammenfassung

6. Zusammenfasung

Raps (*Brassica napus*) ist nach Sojabohne und Ölpalme die weltweit am häufigsten angebaute Ölpflanze. Die Bedeutung von Raps beruht nicht nur auf seinem wirtschaftlichen Wert, sondern auch auf seiner Allopolyploidie und engen Verwandtschaft mit *Arabidopsis thaliana*. Ein Teil des Anstiegs des *B. napus*-Ertrags ist auf die verbesserte Leistung von F_1-Hybriden durch Heterosis zurückzuführen, die in den letzten Jahrzehnten in der Rapsproduktion in Deutschland, China und Kanada eingesetzt wurden. Bisher wurden heterotische- und hybridspezifische Muster vor allem unter allelischen Interaktionshypothesen untersucht. Daher konzentriert sich die vorliegende Arbeit auf die Bewertung genomischer, transkriptomischer und epigenomischer Merkmale in einer Hybridkreuzung zwischen zwei hochhomozygoten und genetisch divergenten Rapslinien. Spontane große genetische Varianten, die sich zwischen F_1-Schwesterpflanzen unterschieden, erzeugten eine schnelle Variation der Genkopienzahl in einer einzigen Generation. Die Auswirkungen solcher vorläufigen Ergebnisse deuten auf mögliche Szenarien von Mechanismen nach der Polyploidisierung hin, um evolutionären Engpässen zu entkommen. Darüber hinaus wurden die Muster von Genexpression, kleine RNA und genomweite Methylierung während der Samen- und Sämlingsentwicklung in der F_1-Hybride evaluiert. Transkriptomische und epigenomische Muster der elterlichen Dominanz erwiesen sich als prägender Faktor. Differentiell exprimierte Gene, die an der Reproduktion und Entwicklung während der Samenbildung beteiligt sind, wurden durch Genexpressions- und microRNA-Analysen nachgewiesen. Epigenetische Allele mit diversen biologischen Hauptfunktionen wurden identifiziert und Verbindungen zwischen epigenomischen und transkriptomischen Merkmalen bewertet. Die somit gewonnenen Erkenntnisse zielen darauf ab, molekulare Muster mit potenziellen Auswirkungen auf Heterosis und Hybridleistung zu untersuchen und Rahmenbedingungen für zukünftige Multiomics- und Pangenomik-Forschung bei polyploiden Nutzpflanzen zu etablieren.

7. References

7. References

Adu-Yeboah, Patricia; Malone, Jenna M.; Gill, Gurjeet; Preston, Christopher (2021): Non-Mendelian inheritance of gene amplification-based resistance to glyphosate in Hordeum glaucum (barley grass) from South Australia. In *Pest management science* 77 (10), pp. 4298–4302. DOI: 10.1002/ps.6518.

Alahakoon, Aruni Y.; Tongson, Eden; Meng, Wei; Ye, Zi-Wei; Russell, Derek A.; Chye, Mee-Len et al. (2022): Overexpressing Arabidopsis thaliana ACBP6 in transgenic rapid-cycling Brassica napus confers cold tolerance. In *Plant methods* 18 (1), p. 62. DOI: 10.1186/s13007-022-00886-y.

Alonge, Michael; Wang, Xingang; Benoit, Matthias; Soyk, Sebastian; Pereira, Lara; Zhang, Lei et al. (2020): Major Impacts of Widespread Structural Variation on Gene Expression and Crop Improvement in Tomato. In *Cell* 182 (1), 145-161.e23. DOI: 10.1016/j.cell.2020.05.021.

Alonso, Conchita; Pérez, Ricardo; Bazaga, Pilar; Herrera, Carlos M. (2015): Global DNA cytosine methylation as an evolving trait: phylogenetic signal and correlated evolution with genome size in angiosperms. In *Frontiers in genetics* 6, p. 4. DOI: 10.3389/fgene.2015.00004.

Alonso-Blanco, C.; Andrade, J.; Becker, C.; Bemm, F.; Bergelson, J.; Borgwardt, K. M.; et al. (2016): 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. In *Cell* 166 (2), pp. 481–491. DOI: 10.1016/j.cell.2016.05.063.

Alonso-Peral, Maria M.; Trigueros, Marina; Sherman, Bjorg; Ying, Hua; Taylor, Jennifer M.; Peacock, William J.; Dennis, Elizabeth S. (2017): Patterns of gene expression in developing embryos of Arabidopsis hybrids. In *The Plant journal : for cell and molecular biology* 89 (5), pp. 927–939. DOI: 10.1111/tpj.13432.

Amarasinghe, Shanika L.; Su, Shian; Dong, Xueyi; Zappia, Luke; Ritchie, Matthew E.; Gouil, Quentin (2020): Opportunities and challenges in long-read sequencing data analysis. In *Genome biology* 21 (1), p. 30. DOI: 10.1186/s13059-020-1935-5.

Ambros, Victor; Bartel, Bonnie; Bartel, David P.; Burge, Christopher B.; Carrington, James C.; Chen, Xuemei et al. (2003): A uniform system for microRNA annotation. In *RNA (New York, N.Y.)* 9 (3), pp. 277–279. DOI: 10.1261/rna.2183803.

An, Hong; Qi, Xinshuai; Gaynor, Michelle L.; Hao, Yue; Gebken, Sarah C.; Mabry, Makenzie E. et al. (2019): Transcriptome and organellar sequencing highlights the complex origin and diversification of allotetraploid Brassica napus. In *Nature communications* 10 (1), p. 2878. DOI: 10.1038/s41467-019-10757-1.

Arechederra, Maria; Daian, Fabrice; Yim, Annie; Bazai, Sehrish K.; Richelme, Sylvie; Dono, Rosanna et al. (2018): Hypermethylation of gene body CpG islands predicts high dosage of functional oncogenes in liver cancer. In *Nature communications* 9 (1), p. 3164. DOI: 10.1038/s41467-018-05550-5.

Argelaguet, Ricard; Velten, Britta; Arnol, Damien; Dietrich, Sascha; Zenz, Thorsten; Marioni, John C. et al. (2018): Multi-Omics Factor Analysis-a framework for unsupervised integration of multi-omics data sets. In *Molecular systems biology* 14 (6), e8124. DOI: 10.15252/msb.20178124.

Attia, T.; Röbbelen, G. (1986): Cytogenetic relationship within cultivated Brassica analyzed in amphihaploids from the three diploid ancestors. In *Can. J. Genet. Cytol.* 28 (3), pp. 323–329. DOI: 10.1139/g86-048.

Badouin, Hélène; Gouzy, Jérôme; Grassa, Christopher J.; Murat, Florent; Staton, S. Evan; Cottret, Ludovic et al. (2017): The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. In *Nature* 546 (7656), pp. 148–152. DOI: 10.1038/nature22380.

Bartels, Arthur; Han, Qiang; Nair, Pooja; Stacey, Liam; Gaynier, Hannah; Mosley, Matthew et al. (2018): Dynamic DNA Methylation in Plant Growth and Development. In *International journal of molecular sciences* 19 (7). DOI: 10.3390/ijms19072144.

Basunanda, P.; Radoev, M.; Ecke, W.; Friedt, W.; Becker, H. C.; Snowdon, R. J. (2010): Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (Brassica napus L.). In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 120 (2), pp. 271–281. DOI: 10.1007/s00122-009-1133-z.

Batista, Rita A.; Köhler, Claudia (2020): Genomic imprinting in plants-revisiting existing models. In *Genes & development* 34 (1-2), pp. 24–36. DOI: 10.1101/gad.332924.119.

Belyeu, Jonathan R.; Brand, Harrison; Wang, Harold; Zhao, Xuefang; Pedersen, Brent S.; Feusier, Julie et al. (2021): De novo structural mutation rates and gamete-of-origin biases revealed through genome sequencing of 2,396 families. In *American journal of human genetics* 108 (4), pp. 597–607. DOI: 10.1016/j.ajhg.2021.02.012.

Bente, Heinrich; Foerster, Andrea M.; Lettner, Nicole; Mittelsten Scheid, Ortrun (2021): Polyploidy-associated paramutation in Arabidopsis is determined by small RNAs, temperature, and allele structure. In *PLoS genetics* 17 (3), e1009444. DOI: 10.1371/journal.pgen.1009444.

Bhat, Javaid A.; Ali, Sajad; Salgotra, Romesh K.; Mir, Zahoor A.; Dutta, Sutapa; Jadon, Vasudha et al. (2016): Genomic Selection in the Era of Next Generation Sequencing for Complex Traits in Plant Breeding. In *Frontiers in genetics* 7, p. 221. DOI: 10.3389/fgene.2016.00221.

Birchler, James A.; Veitia, Reiner A. (2021): One Hundred Years of Gene Balance: How Stoichiometric Issues Affect Gene Expression, Genome Evolution, and Quantitative Traits. In *Cytogenetic and genome research* 161 (10-11), pp. 529–550. DOI: 10.1159/000519592.

Birchler, James A.; Yang, Hua (2022): The multiple fates of gene duplications: Deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. In *The Plant cell* 34 (7), pp. 2466–2474. DOI: 10.1093/plcell/koac076.

Bird, Kevin A.; Niederhuth, Chad E.; Ou, Shujun; Gehan, Malia; Pires, J. Chris; Xiong, Zhiyong et al. (2021a): Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid Brassica napus. In *The New phytologist* 230 (1), pp. 354–371. DOI: 10.1111/nph.17137.

Bird, Kevin A.; Pires, J. Chris; VanBuren, Robert; Xiong, Zhiyong; Edger, Patrick P. (2021b): Gene dosage constraints affect the transcriptional response to allopolyploidy and homoeologous exchange in resynthesized Brassica napus.

Bird, Kevin A.; VanBuren, Robert; Puzey, Joshua R.; Edger, Patrick P. (2018): The causes and consequences of subgenome dominance in hybrids and recent polyploids. In *The New phytologist* 220 (1), pp. 87–93. DOI: 10.1111/nph.15256.

Bohra, Abhishek; Jha, Uday C.; Adhimoolam, Premkumar; Bisht, Deepak; Singh, Narendra P. (2016): Cytoplasmic male sterility (CMS) in hybrid breeding in field crops. In *Plant cell reports* 35 (5), pp. 967–993. DOI: 10.1007/s00299-016-1949-3.

Bonjean, Alain. P.; Dequidt, Céline; Sang, Tina (2016): Rapeseed in China. In *OCL* 23 (6), D605. DOI: 10.1051/ocl/2016045.

Boveri, T. (1902): Über mehrpolige Mitosen als Mittel zur Analyse des Zellkerns. [Concerning multipolar mitoses as a means of analysing the cell nucleus.]. In *Würzburg and Verh. d. phys. med. Ges. zu Würzburg* 35.

Braybrook, Siobhan A.; Stone, Sandra L.; Park, Soomin; Bui, Anhthu Q.; Le, Brandon H.; Fischer, Robert L. et al. (2006): Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. In *Proceedings of the National Academy of Sciences of the United States of America* 103 (9), pp. 3468–3473. DOI: 10.1073/pnas.0511331103.

Bridges, C. B. (1936): The bar "gene" a duplication. In *Science (New York, N.Y.)* 83 (2148), pp. 210–211. DOI: 10.1126/science.83.2148.210.

Bruce, A. B. (1910): The Mendelian theory of heredity and the augmentation of vigor. In *Science (New York, N.Y.)* 32 (827), pp. 627–628. DOI: 10.1126/science.32.827.627-a.

BSA (2022): Federal Plant Variety Office (BSA) descriptive variety 2022 list: cereals, maize, oil and fibre plants, legumes, beets, catch crops. Available online at https://www.bundessortenamt.de/bsa/media/Files/BSL/bsl_getreide_2022.pdf., checked on September, 2022.

Buccitelli, Christopher; Selbach, Matthias (2020): mRNAs, proteins and the emerging principles of gene expression control. In *Nature reviews. Genetics* 21 (10), pp. 630–644. DOI: 10.1038/s41576-020-0258-4.

Canola Council of Canada (2021): History of canola seed development. Available online at https://www.canolacouncil.org/canola-encyclopedia/history-of-canola-seed-development/, checked on November, 2022.

Cao, Shuai; Wang, Longfei; Han, Tongwen; Ye, Wenxue; Liu, Yang; Sun, Yi et al. (2022): Small RNAs mediate transgenerational inheritance of genome-wide trans-acting epialleles in maize. In *Genome biology* 23 (1), p. 53. DOI: 10.1186/s13059-022-02614-0.

Carthew, Richard W.; Sontheimer, Erik J. (2009): Origins and Mechanisms of miRNAs and siRNAs. In *Cell* 136 (4), pp. 642–655. DOI: 10.1016/j.cell.2009.01.035.

Castillo-Bravo, Rosa; Fort, Antoine; Cashell, Ronan; Brychkova, Galina; McKeown, Peter C.; Spillane, Charles (2022): Parent-of-Origin Effects on Seed Size Modify Heterosis Responses in Arabidopsis thaliana. In *Frontiers in plant science* 13, p. 835219. DOI: 10.3389/fpls.2022.835219.

Chalhoub, Boulos; Denoeud, France; Liu, Shengyi; Parkin, Isobel A. P.; Tang, Haibao; Wang, Xiyin et al. (2014): Plant genetics. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. In *Science (New York, N.Y.)* 345 (6199), pp. 950–953. DOI: 10.1126/science.1253435.

Chao, Haoyu; Li, Tian; Luo, Chaoyu; Huang, Hualei; Ruan, Yingfei; Li, Xiaodong et al. (2020): BrassicaEDB: A Gene Expression Database for Brassica Crops. In *International journal of molecular sciences* 21 (16). DOI: 10.3390/ijms21165831.

Chawla, Harmeet Singh; Lee, HueyTyng; Gabur, Iulian; Vollrath, Paul; Tamilselvan-Nattar-Amutha, Suriya; Obermeier, Christian et al. (2021): Long-read sequencing reveals widespread intragenic structural variants in a recent allopolyploid crop plant. In *Plant biotechnology journal* 19 (2), pp. 240–250. DOI: 10.1111/pbi.13456.

Chen, Liangyu; Zhu, Yanyu; Ren, Xiaobo; Yao, Dan; Song, Yang; Fan, Sujie et al. (2022): Heterosis and Differential DNA Methylation in Soybean Hybrids and Their Parental Lines. In *Plants (Basel, Switzerland)* 11 (9). DOI: 10.3390/plants11091136.

Chen, Z. Jeffrey; Ni, Zhongfu (2006): Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. In *BioEssays : news and reviews in molecular, cellular and developmental biology* 28 (3), pp. 240–252. DOI: 10.1002/bies.20374.

Chittem, Kishore; Yajima, William R.; Goswami, Rubella S.; Del Río Mendoza, Luis E. (2020): Transcriptome analysis of the plant pathogen Sclerotinia sclerotiorum interaction with resistant and susceptible canola (Brassica napus) lines. In *PloS one* 15 (3), e0229844. DOI: 10.1371/journal.pone.0229844.

Clark, Susan J. (2007): Action at a distance: epigenetic silencing of large chromosomal regions in carcinogenesis. In *Human molecular genetics* 16 Spec No 1, R88-95. DOI: 10.1093/hmg/ddm051.

Clarke, Wayne E.; Higgins, Erin E.; Plieske, Joerg; Wieseke, Ralf; Sidebottom, Christine; Khedikar, Yogendra et al. (2016): A high-density SNP genotyping array for Brassica napus and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. In *TAG.*

Theoretical and applied genetics. Theoretische und angewandte Genetik 129 (10), pp. 1887–1899. DOI: 10.1007/s00122-016-2746-7.

Comai, Luca (2005): The advantages and disadvantages of being polyploid. In *Nature reviews. Genetics* 6 (11), pp. 836–846. DOI: 10.1038/nrg1711.

Conover, Justin L.; Wendel, Jonathan F. (2022): Deleterious Mutations Accumulate Faster in Allopolyploid Than Diploid Cotton (Gossypium) and Unequally between Subgenomes. In *Molecular biology and evolution* 39 (2). DOI: 10.1093/molbev/msac024.

Crisp, Peter A.; Ganguly, Diep; Eichten, Steven R.; Borevitz, Justin O.; Pogson, Barry J. (2016): Reconsidering plant memory: Intersections between stress recovery, RNA turnover, and epigenetics. In *Science advances* 2 (2), e1501340. DOI: 10.1126/sciadv.1501340.

Crisp, Peter A.; Marand, Alexandre P.; Noshay, Jaclyn M.; Zhou, Peng; Lu, Zefu; Schmitz, Robert J.; Springer, Nathan M. (2020): Stable unmethylated DNA demarcates expressed genes and their cis-regulatory space in plant genomes. In *Proceedings of the National Academy of Sciences of the United States of America* 117 (38), pp. 23991–24000. DOI: 10.1073/pnas.2010250117.

Cui, Chen; Wang, Jing-Jing; Zhao, Jian-Hua; Fang, Yuan-Yuan; He, Xiang-Feng; Guo, Hui-Shan; Duan, Cheng-Guo (2020): A Brassica miRNA Regulates Plant Growth and Immunity through Distinct Modes of Action. In *Molecular plant* 13 (2), pp. 231–245. DOI: 10.1016/j.molp.2019.11.010.

Dalton-Morgan, Jessica; Hayward, Alice; Alamery, Salman; Tollenaere, Reece; Mason, Annaliese S.; Campbell, Emma et al. (2014): A high-throughput SNP array in the amphidiploid species Brassica napus shows diversity in resistance genes. In *Functional & integrative genomics* 14 (4), pp. 643–655. DOI: 10.1007/s10142-014-0391-2.

Darwin, C. (1876): The effects of cross and self fertilisation in the vegetable kingdom. London: John Murry.

Das, Shabari Sarkar; Karmakar, Prakash; Nandi, Asis Kumar; Sanan-Mishra, Neeti (2015): Small RNA mediated regulation of seed germination. In *Frontiers in plant science* 6, p. 828. DOI: 10.3389/fpls.2015.00828.

Dickert, T. E.; Tracy, W. F. (2002): Heterosis for Flowering Time and Agronomic Traits among Early Open-pollinated Sweet Corn Cultivars. In *jashs* 127 (5), pp. 793–797. DOI: 10.21273/JASHS.127.5.793.

Diers, B. W.; McVetty, P. B. E.; Osborn, T. C. (1996): Relationship between Heterosis and Genetic Distance Based on Restriction Fragment Length Polymorphism Markers in Oilseed Rape (Brassica napus L.). In *Crop Sci.* 36 (1), pp. 79–83. DOI: 10.2135/cropsci1996.0011183X003600010014x.

Digital Science (2018): Dimensions [Software]. Available online at https://app.dimensions.ai., checked on November, 2022 under licence agreement.

Ding, Dong; Wang, Yinju; Han, Mingshui; Fu, Zhiyuan; Li, Weihua; Liu, Zonghua et al. (2012): MicroRNA transcriptomic analysis of heterosis during maize seed germination. In *PloS one* 7 (6), e39578. DOI: 10.1371/journal.pone.0039578.

Doğan, Ezgi Süheyla; Liu, Chang (2018): Three-dimensional chromatin packing and positioning of plant genomes. In *Nature plants* 4 (8), pp. 521–529. DOI: 10.1038/s41477-018-0199-5.

Dolatabadian, Aria; Yuan, Yuxuan; Bayer, Philipp Emanuel; Petereit, Jakob; Severn-Ellis, Anita; Tirnaz, Soodeh et al. (2022): Copy Number Variation among Resistance Genes Analogues in Brassica napus. In *Genes* 13 (11). DOI: 10.3390/genes13112037.

Dong, Hongli; Yang, Lei; Liu, Yilin; Tian, Guifu; Tang, Huan; Xin, Shuangshuang et al. (2022a): Detection of new candidate genes controlling seed weight by integrating gene coexpression analysis and QTL mapping in Brassica napus L. In *The Crop Journal*. DOI: 10.1016/j.cj.2022.09.009.

Dong, Qingkun; Hu, Binbin; Zhang, Cui (2022b): microRNAs and Their Roles in Plant Development. In *Frontiers in plant science* 13, p. 824240. DOI: 10.3389/fpls.2022.824240.

Druka, Arnis; Potokina, Elena; Luo, Zewei; Jiang, Ning; Chen, Xinwei; Kearsey, Mike; Waugh, Robbie (2010): Expression quantitative trait loci analysis in plants. In *Plant biotechnology journal* 8 (1), pp. 10–27. DOI: 10.1111/j.1467-7652.2009.00460.x.

Edgar, Robert C.; Taylor, Jeff; Lin, Victor; Altman, Tomer; Barbera, Pierre; Meleshko, Dmitry et al. (2022): Petabase-scale sequence alignment catalyses viral discovery. In *Nature* 602 (7895), pp. 142–147. DOI: 10.1038/s41586-021-04332-2.

Ellwood, Simon R.; Phan, Huyen T. T.; Jordan, Megan; Hane, James; Torres, Anna M.; Avila, Carmen M. et al. (2008): Construction of a comparative genetic map in faba bean (Vicia faba L.); conservation of genome structure with Lens culinaris. In *BMC genomics* 9, p. 380. DOI: 10.1186/1471-2164-9-380.

Escaramís, Geòrgia; Docampo, Elisa; Rabionet, Raquel (2015): A decade of structural variants: description, history and methods to detect structural variation. In *Briefings in functional genomics* 14 (5), pp. 305–314. DOI: 10.1093/bfgp/elv014.

Fang, Caochuang; Wang, Zhaoyang; Wang, Pengfei; Song, Yixian; Ahmad, Ali; Dong, Faming et al. (2021): Heterosis Derived From Nonadditive Effects of the BnFLC Homologs Coordinates Early Flowering and High Yield in Rapeseed (Brassica napus L.). In *Frontiers in plant science* 12, p. 798371. DOI: 10.3389/fpls.2021.798371.

FAO (2022): FAOSTAT. Available online at https://www.fao.org/faostat/en, checked on September, 2022.

Ferreira, M. E.; Satagopan, J.; Yandell, B. S.; Williams, P. H.; Osborn, T. C. (1995): Mapping loci controlling vernalization requirement and flowering time in Brassica napus. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 90 (5), pp. 727–732. DOI: 10.1007/BF00222140.

Ferreira de Carvalho, Julie; Stoeckel, Solenn; Eber, Frédérique; Lodé-Taburel, Maryse; Gilet, Marie-Madeleine; Trotoux, Gwenn et al. (2021): Untangling structural factors driving genome stabilization in nascent Brassica napus allopolyploids. In *The New phytologist* 230 (5), pp. 2072–2084. DOI: 10.1111/nph.17308.

FJZ; HHU (2014): Jülich Research Center (FZJ) and Heinrich Heine University Düsseldorf (HHU) plaBi database. Available online at https://www.plabipd.de, checked on 10/11/2022.

Fort, Antoine; Ryder, Peter; McKeown, Peter C.; Wijnen, Cris; Aarts, Mark G.; Sulpice, Ronan; Spillane, Charles (2016): Disaggregating polyploidy, parental genome dosage and hybridity contributions to heterosis in Arabidopsis thaliana. In *The New phytologist* 209 (2), pp. 590–599. DOI: 10.1111/nph.13650.

Friedt, Wolfgang; Tu, Jingxing; Fu, Tingdong (2018): Academic and Economic Importance of Brassica napus Rapeseed. In Shengyi Liu, Rod Snowdon, Boulos Chalhoub (Eds.): The Brassica napus Genome. Cham: Springer International Publishing (Compendium of Plant Genomes), pp. 1–20.

Fuentes, Roven Rommel; Chebotarov, Dmytro; Duitama, Jorge; Smith, Sean; La Hoz, Juan Fernando de; Mohiyuddin, Marghoob et al. (2019): Structural variants in 3000 rice genomes. In *Genome research* 29 (5), pp. 870–880. DOI: 10.1101/gr.241240.118.

Fujimoto, Ryo; Uezono, Kosuke; Ishikura, Sonoko; Osabe, Kenji; Peacock, W. James; Dennis, Elizabeth S. (2018): Recent research on the mechanism of heterosis is important for crop and vegetable breeding systems. In *Breeding science* 68 (2), pp. 145–158. DOI: 10.1270/jsbbs.17155.

Gabur, Iulian; Chawla, Harmeet Singh; Lopisso, Daniel Teshome; Tiedemann, Andreas von; Snowdon, Rod J.; Obermeier, Christian (2020): Gene presence-absence variation associates with quantitative

Verticillium longisporum disease resistance in Brassica napus. In *Scientific reports* 10 (1), p. 4131. DOI: 10.1038/s41598-020-61228-3.

Gabur, Iulian; Chawla, Harmeet Singh; Snowdon, Rod J.; Parkin, Isobel A. P. (2019): Connecting genome structural variation with complex traits in crop plants. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 132 (3), pp. 733–750. DOI: 10.1007/s00122-018-3233-0.

Gaebelein, Roman; Alnajar, Dima; Koopmann, Birger; Mason, Annaliese S. (2019): Hybrids between Brassica napus and B. nigra show frequent pairing between the B and A/C genomes and resistance to blackleg. In *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 27 (3), pp. 221–236. DOI: 10.1007/s10577-019-09612-2.

Gallego-Bartolomé, Javier (2020): DNA methylation in plants: mechanisms and tools for targeted manipulation. In *The New phytologist* 227 (1), pp. 38–44. DOI: 10.1111/nph.16529.

Gehring, Mary; Bubb, Kerry L.; Henikoff, Steven (2009): Extensive demethylation of repetitive elements during seed development underlies gene imprinting. In *Science (New York, N.Y.)* 324 (5933), pp. 1447–1451. DOI: 10.1126/science.1171609.

Girke, Andreas; Schierholt, Antje; Becker, Heiko C. (2012): Extending the rapeseed gene pool with resynthesized Brassica napus II: Heterosis. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 124 (6), pp. 1017–1026. DOI: 10.1007/s00122-011-1765-7.

Greaves, Ian K.; Gonzalez-Bayon, Rebeca; Wang, Li; Zhu, Anyu; Liu, Pei-Chuan; Groszmann, Michael et al. (2015): Epigenetic Changes in Hybrids. In *Plant physiology* 168 (4), pp. 1197–1205. DOI: 10.1104/pp.15.00231.

Griffiths-Jones, Sam; Grocock, Russell J.; van Dongen, Stijn; Bateman, Alex; Enright, Anton J. (2006): miRBase: microRNA sequences, targets and gene nomenclature. In *Nucleic acids research* 34 (Database issue), D140-4. DOI: 10.1093/nar/gkj112.

Groszmann, Michael; Greaves, Ian K.; Albertyn, Zayed I.; Scofield, Graham N.; Peacock, William J.; Dennis, Elizabeth S. (2011): Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor. In *Proceedings of the National Academy of Sciences of the United States of America* 108 (6), pp. 2617–2622. DOI: 10.1073/pnas.1019217108.

Grover, Jeffrey W.; Burgess, Diane; Kendall, Timmy; Baten, Abdul; Pokhrel, Suresh; King, Graham J. et al. (2020): Abundant expression of maternal siRNAs is a conserved feature of seed development. In *Proceedings of the National Academy of Sciences of the United States of America* 117 (26), pp. 15305–15315. DOI: 10.1073/pnas.2001332117.

Grover, Jeffrey W.; Kendall, Timmy; Baten, Abdul; Burgess, Diane; Freeling, Michael; King, Graham J.; Mosher, Rebecca A. (2018): Maternal components of RNA-directed DNA methylation are required for seed development in Brassica rapa. In *The Plant journal : for cell and molecular biology* 94 (4), pp. 575–582. DOI: 10.1111/tpj.13910.

Guan, Jiantao; Xu, Yaoguang; Yu, Yang; Fu, Jun; Ren, Fei; Guo, Jiying et al. (2021): Genome structure variation analyses of peach reveal population dynamics and a 1.67 Mb causal inversion for fruit shape. In *Genome biology* 22 (1), p. 13. DOI: 10.1186/s13059-020-02239-1.

Gullotta, Giorgio; Korte, Arthur; Marquardt, Sebastian (2022): Functional Variation in the Non-Coding Genome: Molecular Implications for Food Security. In *Journal of experimental botany*. DOI: 10.1093/jxb/erac395.

Hallauer, Arnel R. (2007): History, Contribution, and Future of Quantitative Genetics in Plant Breeding: Lessons From Maize. In *Crop Sci.* 47, S-4-S-19. DOI: 10.2135/cropsci2007.04.0002IPBS.

Han, Xu; Tang, Qingqing; Xu, Liping; Guan, Zhilin; Tu, Jinxing; Yi, Bin et al. (2022): Genome-wide detection of genotype environment interactions for flowering time in Brassica napus. In *Frontiers in plant science* 13, Article 1065766. DOI: 10.3389/fpls.2022.1065766.

Hasan, M.; Seyis, F.; Badani, A. G.; Pons-Kühnemann, J.; Friedt, W.; Lühs, W.; Snowdon, R. J. (2006): Analysis of Genetic Diversity in the Brassica napus L. Gene Pool Using SSR Markers. In *Genet Resour Crop Evol* 53 (4), pp. 793–802. DOI: 10.1007/s10722-004-5541-2.

Hatzig, Sarah; Breuer, Frank; Nesi, Nathalie; Ducournau, Sylvie; Wagner, Marie-Helene; Leckband, Gunhild et al. (2018): Hidden Effects of Seed Quality Breeding on Germination in Oilseed Rape (Brassica napus L.). In *Frontiers in plant science* 9, p. 419. DOI: 10.3389/fpls.2018.00419.

Hatzig, Sarah V.; Frisch, Matthias; Breuer, Frank; Nesi, Nathalie; Ducournau, Sylvie; Wagner, Marie-Helene et al. (2015): Genome-wide association mapping unravels the genetic control of seed germination and vigor in Brassica napus. In *Frontiers in plant science* 6, p. 221. DOI: 10.3389/fpls.2015.00221.

Hawe, Johann S.; Wilson, Rory; Schmid, Katharina T.; Zhou, Li; Lakshmanan, Lakshmi Narayanan; Lehne, Benjamin C. et al. (2022): Genetic variation influencing DNA methylation provides insights into molecular mechanisms regulating genomic function. In *Nature genetics* 54 (1), pp. 18–29. DOI: 10.1038/s41588-021-00969-x.

Hayward, Alice; Mason, Annaliese S.; Dalton-Morgan, Jessica; Zander, Manuel; Edwards, David; Batley, Jacqueline (2012): SNP discovery and applications in Brassica napus. In *Journal of Plant Biotechnology* 39 (1), pp. 49–61. DOI: 10.5010/JPB.2012.39.1.049.

He, Fei; Wang, Wei; Rutter, William B.; Jordan, Katherine W.; Ren, Jie; Taagen, Ellie et al. (2022): Genomic variants affecting homoeologous gene expression dosage contribute to agronomic trait variation in allopolyploid wheat. In *Nature communications* 13 (1), p. 826. DOI: 10.1038/s41467-022-28453-y.

He, Shoupu; Sun, Gaofei; Geng, Xiaoli; Gong, Wenfang; Dai, Panhong; Jia, Yinhua et al. (2021): The genomic basis of geographic differentiation and fiber improvement in cultivated cotton. In *Nature genetics* 53 (6), pp. 916–924. DOI: 10.1038/s41588-021-00844-9.

He, Yajun; Wu, Daoming; Wei, Dayong; Fu, Ying; Cui, Yixin; Dong, Hongli et al. (2017a): GWAS, QTL mapping and gene expression analyses in Brassica napus reveal genetic control of branching morphogenesis. In *Scientific reports* 7 (1), p. 15971. DOI: 10.1038/s41598-017-15976-4.

He, Zhesi; Wang, Lihong; Harper, Andrea L.; Havlickova, Lenka; Pradhan, Akshay K.; Parkin, Isobel A. P.; Bancroft, Ian (2017b): Extensive homoeologous genome exchanges in allopolyploid crops revealed by mRNAseq-based visualization. In *Plant biotechnology journal* 15 (5), pp. 594–604. DOI: 10.1111/pbi.12657.

Heard, Edith; Martienssen, Robert A. (2014): Transgenerational epigenetic inheritance: myths and mechanisms. In *Cell* 157 (1), pp. 95–109. DOI: 10.1016/j.cell.2014.02.045.

Hiebert, Colin W.; Moscou, Matthew J.; Hewitt, Tim; Steuernagel, Burkhard; Hernández-Pinzón, Inma; Green, Phon et al. (2020): Stem rust resistance in wheat is suppressed by a subunit of the mediator complex. In *Nature communications* 11 (1), p. 1123. DOI: 10.1038/s41467-020-14937-2.

Higgins, Erin E.; Clarke, Wayne E.; Howell, Elaine C.; Armstrong, Susan J.; Parkin, Isobel A. P. (2018): Detecting de Novo Homoeologous Recombination Events in Cultivated Brassica napus Using a Genome-Wide SNP Array. In *G3 (Bethesda, Md.)* 8 (8), pp. 2673–2683. DOI: 10.1534/g3.118.200118.

Hochholdinger, Frank; Hoecker, Nadine (2007): Towards the molecular basis of heterosis. In *Trends in plant science* 12 (9), pp. 427–432. DOI: 10.1016/j.tplants.2007.08.005.

Hofmeister, Brigitte T.; Lee, Kevin; Rohr, Nicholas A.; Hall, David W.; Schmitz, Robert J. (2017): Stable inheritance of DNA methylation allows creation of epigenotype maps and the study of epiallele inheritance patterns in the absence of genetic variation. In *Genome biology* 18 (1), p. 155. DOI: 10.1186/s13059-017-1288-x.

Hörandl, Elvira (2022): Novel Approaches for Species Concepts and Delimitation in Polyploids and Hybrids. In *Plants (Basel, Switzerland)* 11 (2). DOI: 10.3390/plants11020204.

Hotchkiss, R. D. (1948): The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. In *The Journal of biological chemistry* 175 (1), pp. 315–332.

Hu, Dandan; Jing, Jinjie; Snowdon, Rod J.; Mason, Annaliese S.; Shen, Jinxiong; Meng, Jinling; Zou, Jun (2021): Exploring the gene pool of Brassica napus by genomics-based approaches. In *Plant biotechnology journal* 19 (9), pp. 1693–1712. DOI: 10.1111/pbi.13636.

Hu, Lanjuan; Li, Ning; Xu, Chunming; Zhong, Silin; Lin, Xiuyun; Yang, Jingjing et al. (2014): Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. In *Proceedings of the National Academy of Sciences of the United States of America* 111 (29), pp. 10642–10647. DOI: 10.1073/pnas.1410761111.

Hu, Yue; Xiong, Jie; Shalby, Nesma; Zhuo, Chenjian; Jia, Yupeng; Yang, Qing-Yong; Tu, Jinxing (2022): Comparison of dynamic 3D chromatin architecture uncovers heterosis for leaf size in Brassica napus. In *Journal of Advanced Research*. DOI: 10.1016/j.jare.2022.01.001.

Huang, Daiqing; Koh, Chushin; Feurtado, J. Allan; Tsang, Edward W. T.; Cutler, Adrian J. (2013): MicroRNAs and their putative targets in Brassica napus seed maturation. In *BMC genomics* 14, p. 140. DOI: 10.1186/1471-2164-14-140.

Huang, Feiyi; Wu, Xiaoting; Hou, Xilin; Shao, Shuaixu; Liu, Tongkun (2018): Vernalization can regulate flowering time through microRNA mechanism in Brassica rapa. In *Physiologia plantarum* 164 (2), pp. 204–215. DOI: 10.1111/ppl.12692.

Huang, Hsi-Yuan; Lin, Yang-Chi-Dung; Cui, Shidong; Huang, Yixian; Tang, Yun; Xu, Jiatong et al. (2022): miRTarBase update 2022: an informative resource for experimentally validated miRNA-target interactions. In *Nucleic acids research* 50 (D1), D222-D230. DOI: 10.1093/nar/gkab1079.

Hurgobin, Bhavna; Golicz, Agnieszka A.; Bayer, Philipp E.; Chan, Chon-Kit Kenneth; Tirnaz, Soodeh; Dolatabadian, Aria et al. (2018): Homoeologous exchange is a major cause of gene presence/absence variation in the amphidiploid Brassica napus. In *Plant biotechnology journal* 16 (7), pp. 1265–1274. DOI: 10.1111/pbi.12867.

Ihien Katche, Elizabeth; Schierholt, Antje; Becker, Heiko C.; Batley, Jacqueline; Mason, Annaliese S. (2022): Fertility, genome stability, and homozygosity in a diverse set of resynthesized rapeseed lines. In *The Crop Journal*. DOI: 10.1016/j.cj.2022.07.022.

Jackson, R. C. (1982): Polyploidy and diploidy: new perspectives on chromosome pairing and its evolutionary implications. In *American Journal of Botany* 69 (9), pp. 1512–1523. DOI: 10.1002/j.1537-2197.1982.tb13400.x.

Jahnke, Stephanie; Sarholz, Barbara; Thiemann, Alexander; Kühr, Vera; Gutiérrez-Marcos, José F.; Geiger, Hartwig H. et al. (2010): Heterosis in early seed development: a comparative study of F1 embryo and endosperm tissues 6 days after fertilization. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 120 (2), pp. 389–400. DOI: 10.1007/s00122-009-1207-y.

Jain, Aditi; Anand, Saurabh; Singh, Neer K.; Das, Sandip (2018): Sequence and functional characterization of MIRNA164 promoters from Brassica shows copy number dependent regulatory diversification among homeologs. In *Functional & integrative genomics* 18 (4), pp. 369–383. DOI: 10.1007/s10142-018-0598-8.

Jamil, Ili Nadhirah; Remali, Juwairiah; Azizan, Kamalrul Azlan; Nor Muhammad, Nor Azlan; Arita, Masanori; Goh, Hoe-Han; Aizat, Wan Mohd (2020): Systematic Multi-Omics Integration (MOI) Approach in Plant Systems Biology. In *Frontiers in plant science* 11, p. 944. DOI: 10.3389/fpls.2020.00944.

Jian, Hongju; Zhang, Aoxiang; Ma, Jinqi; Wang, Tengyue; Yang, Bo; Shuang, Lan Shuan et al. (2019): Joint QTL mapping and transcriptome sequencing analysis reveal candidate flowering time genes in Brassica napus L. In *BMC genomics* 20 (1), p. 21. DOI: 10.1186/s12864-018-5356-8.

Johannsen, W. (1909): Elemente der exakten Erblichkeitslehre: Gustav Fischer.

Johnston, T. D. (1971): A comparison of inbred lines and their F1 hybrids in forage rape (Brassica napus L.). In *Euphytica* 20 (1), pp. 81–85. DOI: 10.1007/BF00146777.

Kalisz, Susan; Purugganan, Michael D. (2004): Epialleles via DNA methylation: consequences for plant evolution. In *Trends in ecology & evolution* 19 (6), pp. 309–314. DOI: 10.1016/j.tree.2004.03.034.

Kamthan, Ayushi; Chaudhuri, Abira; Kamthan, Mohan; Datta, Asis (2015): Small RNAs in plants: recent development and application for crop improvement. In *Frontiers in plant science* 6, p. 208. DOI: 10.3389/fpls.2015.00208.

Katche, Elvis; Gaebelein, Roman; Idris, Zurianti; Vasquez-Teuber, Paula; Lo, Yu-Tzu; Nugent, David et al. (2021): Stable, fertile lines produced by hybridization between allotetraploids Brassica juncea (AABB) and Brassica carinata (BBCC) have merged the A and C genomes. In *The New phytologist* 230 (3), pp. 1242–1257. DOI: 10.1111/nph.17225.

Kebede, Aida Z.; Johnston, Anne; Schneiderman, Danielle; Bosnich, Whynn; Harris, Linda J. (2018): Transcriptome profiling of two maize inbreds with distinct responses to Gibberella ear rot disease to identify candidate resistance genes. In *BMC genomics* 19 (1), p. 131. DOI: 10.1186/s12864-018-4513-4.

Khan, Deirdre; Ziegler, Dylan J.; Kalichuk, Jenna L.; Hoi, Vanessa; Huynh, Nina; Hajihassani, Abolfazl et al. (2022): Gene expression profiling reveals transcription factor networks and subgenome bias during Brassica napus seed development. In *The Plant journal : for cell and molecular biology* 109 (3), pp. 477–489. DOI: 10.1111/tpj.15587.

Khare, Shruti; Gurry, Céline; Freitas, Lucas; Schultz, Mark B.; Bach, Gunter; Diallo, Amadou et al. (2021): GISAID's Role in Pandemic Response. In *China CDC weekly* 3 (49), pp. 1049–1051. DOI: 10.46234/ccdcw2021.255.

Kippes, Nestor; Debernardi, Juan M.; Vasquez-Gross, Hans A.; Akpinar, Bala A.; Budak, Hikment; Kato, Kenji et al. (2015): Identification of the VERNALIZATION 4 gene reveals the origin of spring growth habit in ancient wheats from South Asia. In *Proceedings of the National Academy of Sciences of the United States of America* 112 (39), E5401-10. DOI: 10.1073/pnas.1514883112.

Kirkbride, Ryan C.; Lu, Jie; Zhang, Changqing; Mosher, Rebecca A.; Baulcombe, David C.; Chen, Z. Jeffrey (2019): Maternal small RNAs mediate spatial-temporal regulation of gene expression, imprinting, and seed development in Arabidopsis. In *Proceedings of the National Academy of Sciences of the United States of America* 116 (7), pp. 2761–2766. DOI: 10.1073/pnas.1807621116.

Kirov, Ilya; Polkhovskaya, Ekaterina; Dudnikov, Maxim; Merkulov, Pavel; Vlasova, Anastasia; Karlov, Gennady; Soloviev, Alexander (2021): Searching for a Needle in a Haystack: Cas9-Targeted Nanopore Sequencing and DNA Methylation Profiling of Full-Length Glutenin Genes in a Big Cereal Genome. In *Plants (Basel, Switzerland)* 11 (1). DOI: 10.3390/plants11010005.

Knoch, Dominic; Werner, Christian R.; Meyer, Rhonda C.; Riewe, David; Abbadi, Amine; Lücke, Sophie et al. (2021): Multi-omics-based prediction of hybrid performance in canola. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 134 (4), pp. 1147–1165. DOI: 10.1007/s00122-020-03759-x.

Körber, Niklas; Bus, Anja; Li, Jinquan; Higgins, Janet; Bancroft, Ian; Higgins, Erin Eileen et al. (2015): Seedling development traits in Brassica napus examined by gene expression analysis and association mapping. In *BMC plant biology* 15, p. 136. DOI: 10.1186/s12870-015-0496-3.

Körber, Niklas; Wittkop, Benjamin; Bus, Anja; Friedt, Wolfgang; Snowdon, Rod J.; Stich, Benjamin (2012): Seedling development in a Brassica napus diversity set and its relationship to agronomic performance. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 125 (6), pp. 1275–1287. DOI: 10.1007/s00122-012-1912-9.

Kozomara, Ana; Birgaoanu, Maria; Griffiths-Jones, Sam (2019): miRBase: from microRNA sequences to function. In *Nucleic acids research* 47 (D1), D155-D162. DOI: 10.1093/nar/gky1141.

Kumar, Suresh; Mohapatra, Trilochan (2021): Dynamics of DNA Methylation and Its Functions in Plant Growth and Development. In *Frontiers in plant science* 12, p. 596236. DOI: 10.3389/fpls.2021.596236.

Lauss, Kathrin; Wardenaar, René; Oka, Rurika; van Hulten, Marieke H. A.; Guryev, Victor; Keurentjes, Joost J. B. et al. (2018): Parental DNA Methylation States Are Associated with Heterosis in Epigenetic Hybrids. In *Plant physiology* 176 (2), pp. 1627–1645. DOI: 10.1104/pp.17.01054.

Le, Brandon H.; Cheng, Chen; Bui, Anhthu Q.; Wagmaister, Javier A.; Henry, Kelli F.; Pelletier, Julie et al. (2010): Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. In *Proceedings of the National Academy of Sciences of the United States of America* 107 (18), pp. 8063–8070. DOI: 10.1073/pnas.1003530107.

Lebrigand, Kevin; Magnone, Virginie; Barbry, Pascal; Waldmann, Rainer (2020): High throughput error corrected Nanopore single cell transcriptome sequencing. In *Nature communications* 11 (1), p. 4025. DOI: 10.1038/s41467-020-17800-6.

Lee, HueyTyng; Chawla, Harmeet Singh; Obermeier, Christian; Dreyer, Felix; Abbadi, Amine; Snowdon, Rod (2020): Chromosome-Scale Assembly of Winter Oilseed Rape Brassica napus. In *Frontiers in plant science* 11, p. 496. DOI: 10.3389/fpls.2020.00496.

Leijten, Willeke; Koes, Ronald; Roobeek, Ilja; Frugis, Giovanna (2018): Translating Flowering Time From Arabidopsis thaliana to Brassicaceae and Asteraceae Crop Species. In *Plants (Basel, Switzerland)* 7 (4). DOI: 10.3390/plants7040111.

Li, Aili; Liu, Dengcai; Wu, Jun; Zhao, Xubo; Hao, Ming; Geng, Shuaifeng et al. (2014): mRNA and Small RNA Transcriptomes Reveal Insights into Dynamic Homoeolog Regulation of Allopolyploid Heterosis in Nascent Hexaploid Wheat. In *The Plant cell* 26 (5), pp. 1878–1900. DOI: 10.1105/tpc.114.124388.

Li, Changsheng; Xiang, Xiaoli; Huang, Yongcai; Zhou, Yong; An, Dong; Dong, Jiaqiang et al. (2020a): Long-read sequencing reveals genomic structural variations that underlie creation of quality protein maize. In *Nature communications* 11 (1), p. 17. DOI: 10.1038/s41467-019-14023-2.

Li, Mengdi; Wang, Ruihua; Wu, Xiaoming; Wang, Jianbo (2020b): Homoeolog expression bias and expression level dominance (ELD) in four tissues of natural allotetraploid Brassica napus. In *BMC genomics* 21 (1), p. 330. DOI: 10.1186/s12864-020-6747-1.

Li, Ruijuan; Jeong, Kwangju; Davis, John T.; Kim, Seungmo; Lee, Soonbong; Michelmore, Richard W. et al. (2018): Integrated QTL and eQTL Mapping Provides Insights and Candidate Genes for Fatty Acid Composition, Flowering Time, and Growth Traits in a F2 Population of a Novel Synthetic Allopolyploid Brassica napus. In *Frontiers in plant science* 9, p. 1632. DOI: 10.3389/fpls.2018.01632.

Li, Zeyu; Li, Mengdi; Wang, Jianbo (2022): Asymmetric subgenomic chromatin architecture impacts on gene expression in resynthesized and natural allopolyploid Brassica napus. In *Communications biology* 5 (1), p. 762. DOI: 10.1038/s42003-022-03729-7.

Liu, Chang; Zhu, Xiya; Zhang, Jin; Shen, Meng; Chen, Kai; Fu, Xiangkui et al. (2022a): eQTLs play critical roles in regulating gene expression and identifying key regulators in rice. In *Plant biotechnology journal* 20 (12), pp. 2357–2371. DOI: 10.1111/pbi.13912.

Liu, Dongxu; Yu, Liangqian; Wei, Lulu; Yu, Pugang; Wang, Jing; Zhao, Hu et al. (2021a): BnTIR: an online transcriptome platform for exploring RNA-seq libraries for oil crop Brassica napus. In *Plant biotechnology journal* 19 (10), pp. 1895–1897. DOI: 10.1111/pbi.13665.

Liu, Jie; Wu, Yupo; Cui, Xiaobo; Zhang, Xiong; Xie, Meili; Liu, Lijiang et al. (2022b): Genome-wide characterization of ovate family protein gene family associated with number of seeds per silique in Brassica napus. In *Frontiers in plant science* 13, p. 962592. DOI: 10.3389/fpls.2022.962592.

Liu, Jing; Li, Jun; Liu, Hong-Fang; Fan, Shi-Hang; Singh, Surinder; Zhou, Xue-Rong et al. (2018): Genome-wide screening and analysis of imprinted genes in rapeseed (Brassica napus L.) endosperm. In *DNA research : an international journal for rapid publication of reports on genes and genomes* 25 (6), pp. 629–640. DOI: 10.1093/dnares/dsy030.

Liu, Liezhao; Qu, Cunmin; Wittkop, Benjamin; Yi, Bin; Xiao, Yang; He, Yajun et al. (2013): A high-density SNP map for accurate mapping of seed fibre QTL in Brassica napus L. In *PloS one* 8 (12), e83052. DOI: 10.1371/journal.pone.0083052.

Liu, Pei-Chuan; Peacock, W. James; Wang, Li; Furbank, Robert; Larkum, Anthony; Dennis, Elizabeth S. (2020): Leaf growth in early development is key to biomass heterosis in Arabidopsis. In *Journal of experimental botany* 71 (8), pp. 2439–2450. DOI: 10.1093/jxb/eraa006.

Liu, Wenwen; He, Guangming; Deng, Xing Wang (2021b): Biological pathway expression complementation contributes to biomass heterosis in Arabidopsis. In *Proceedings of the National Academy of Sciences of the United States of America* 118 (16). DOI: 10.1073/pnas.2023278118.

Lloyd, Andrew; Blary, Aurélien; Charif, Delphine; Charpentier, Catherine; Tran, Joseph; Balzergue, Sandrine et al. (2018): Homoeologous exchanges cause extensive dosage-dependent gene expression changes in an allopolyploid crop. In *The New phytologist* 217 (1), pp. 367–377. DOI: 10.1111/nph.14836.

Lolle, Susan J.; Victor, Jennifer L.; Young, Jessica M.; Pruitt, Robert E. (2005): Genome-wide non-mendelian inheritance of extra-genomic information in Arabidopsis. In *Nature* 434 (7032), pp. 505–509. DOI: 10.1038/nature03380.

Louwaars, Niels P. (2018): Plant breeding and diversity: A troubled relationship? In *Euphytica: Netherlands journal of plant breeding* 214 (7), p. 114. DOI: 10.1007/s10681-018-2192-5.

Lu, Jie; Zhang, Changqing; Baulcombe, David C.; Chen, Z. Jeffrey (2012): Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of Arabidopsis seeds. In *Proceedings of the National Academy of Sciences of the United States of America* 109 (14), pp. 5529–5534. DOI: 10.1073/pnas.1203094109.

Lu, Kun; Peng, Liu; Zhang, Chao; Lu, Junhua; Yang, Bo; Xiao, Zhongchun et al. (2017): Genome-Wide Association and Transcriptome Analyses Reveal Candidate Genes Underlying Yield-determining Traits in Brassica napus. In *Frontiers in plant science* 8, p. 206. DOI: 10.3389/fpls.2017.00206.

Lu, Kun; Wei, Lijuan; Li, Xiaolong; Wang, Yuntong; Wu, Jian; Liu, Miao et al. (2019): Whole-genome resequencing reveals Brassica napus origin and genetic loci involved in its improvement. In *Nature communications* 10 (1), p. 1154. DOI: 10.1038/s41467-019-09134-9.

Lunardon, Alice; Johnson, Nathan R.; Hagerott, Emily; Phifer, Tamia; Polydore, Seth; Coruh, Ceyda; Axtell, Michael J. (2020): Integrated annotations and analyses of small RNA-producing loci from 47 diverse plants. In *Genome research* 30 (3), pp. 497–513. DOI: 10.1101/gr.256750.119.

Luo, Jin-Hong; Wang, Min; Jia, Gui-Fang; He, Yan (2021): Transcriptome-wide analysis of epitranscriptome and translational efficiency associated with heterosis in maize. In *Journal of experimental botany* 72 (8), pp. 2933–2946. DOI: 10.1093/jxb/erab074.

Luo, Xiang; Tan, Yongqiang; Ma, Chaozhi; Tu, Jinxing; Shen, Jinxiong; Yi, Bin; Fu, Tingdong (2019): High-throughput identification of SNPs reveals extensive heterosis with intra-group hybridization and genetic characteristics in a large rapeseed (Brassica napus L.) panel. In *Euphytica* 215 (10). DOI: 10.1007/s10681-019-2484-4.

Maher, Brendan; van Noorden, Richard (2021): How the COVID pandemic is changing global science collaborations. In *Nature* 594 (7863), pp. 316–319. DOI: 10.1038/d41586-021-01570-2.

Mahmood, Sammina; Li, Zhaohong; Yue, Xiaopeng; Wang, Bo; Chen, Jun; Liu, Kede (2016): Development of INDELs markers in oilseed rape (Brassica napus L.) using re-sequencing data. In *Mol Breeding* 36 (6). DOI: 10.1007/s11032-016-0501-z.

Mahmoud, Medhat; Gobet, Nastassia; Cruz-Dávalos, Diana Ivette; Mounier, Ninon; Dessimoz, Christophe; Sedlazeck, Fritz J. (2019): Structural variant calling: the long and the short of it. In *Genome biology* 20 (1), p. 246. DOI: 10.1186/s13059-019-1828-7.

Malla, Stavroula; Brewin, Derek G. (2019): Crop research, biotech canola, and innovation policy in Canada: Challenges, opportunities, and evolution. In *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 67 (2), pp. 135–150. DOI: 10.1111/cjag.12195.

Marks, Rose A.; Hotaling, Scott; Frandsen, Paul B.; VanBuren, Robert (2021): Representation and participation across 20 years of plant genome sequencing. In *Nature plants* 7 (12), pp. 1571–1578. DOI: 10.1038/s41477-021-01031-8.

Mather, K. (1949): Biometrical genetics. London: Methuen.

Mayrose, Itay; Zhan, Shing H.; Rothfels, Carl J.; Arrigo, Nils; Barker, Michael S.; Rieseberg, Loren H.; Otto, Sarah P. (2015): Methods for studying polyploid diversification and the dead end hypothesis: a reply to Soltis et al. (2014). In *The New phytologist* 206 (1), pp. 27–35. DOI: 10.1111/nph.13192.

McClintock, B. (1950): The origin and behavior of mutable loci in maize. In *Proceedings of the National Academy of Sciences of the United States of America* 36 (6), pp. 344–355. DOI: 10.1073/pnas.36.6.344.

McCormick, Sheila (2018): RNA-directed DNA methylation and seed development: an unexpected difference between Arabidopsis thaliana and Brassica rapa. In *The Plant journal : for cell and molecular biology* 94 (4), pp. 573–574. DOI: 10.1111/tpj.13935.

McCue, Andrea D.; Panda, Kaushik; Nuthikattu, Saivageethi; Choudury, Sarah G.; Thomas, Erica N.; Slotkin, R. Keith (2015): ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. In *The EMBO journal* 34 (1), pp. 20–35. DOI: 10.15252/embj.201489499.

McNaughton, I. H.; Munro, Isabel K. (1972): Heterosis and its possible exploitation in swedes (Brassica napus L. ssp. rapifera). In *Euphytica* 21 (3), pp. 518–522. DOI: 10.1007/BF00039349.

Mehraj, Hasan; Kawanabe, Takahiro; Shimizu, Motoki; Miyaji, Naomi; Akter, Ayasha; Dennis, Elizabeth S.; Fujimoto, Ryo (2020): In Arabidopsis thaliana Heterosis Level Varies among Individuals in an F1 Hybrid Population. In *Plants (Basel, Switzerland)* 9 (4). DOI: 10.3390/plants9040414.

Meinke, D. W.; Cherry, J. M.; Dean, C.; Rounsley, S. D.; Koornneef, M. (1998): Arabidopsis thaliana: a model plant for genome analysis. In *Science (New York, N.Y.)* 282 (5389), 662, 679-82. DOI: 10.1126/science.282.5389.662.

Meinke, David W. (2020): Genome-wide identification of EMBRYO-DEFECTIVE (EMB) genes required for growth and development in Arabidopsis. In *The New phytologist* 226 (2), pp. 306–325. DOI: 10.1111/nph.16071.

Mendel, G. J. (1866): Versuche über Pflanzenhybriden. In *Verhandlungen des Naturforschenden Vereines in Brünn* 4, pp. 3–47.

Mercé, Clémentine; Bayer, Philipp E.; Tay Fernandez, Cassandria; Batley, Jacqueline; Edwards, David (2020): Induced Methylation in Plants as a Crop Improvement Tool: Progress and Perspectives. In *Agronomy* 10 (10), p. 1484. DOI: 10.3390/agronomy10101484.

Meyer, Rhonda C.; Witucka-Wall, Hanna; Becher, Martina; Blacha, Anna; Boudichevskaia, Anastassia; Dörmann, Peter et al. (2012): Heterosis manifestation during early Arabidopsis seedling development is characterized by intermediate gene expression and enhanced metabolic activity in the hybrids. In *The Plant journal : for cell and molecular biology* 71 (4), pp. 669–683. DOI: 10.1111/j.1365-313X.2012.05021.x.

Miller, Charlotte; Wells, Rachel; McKenzie, Neil; Trick, Martin; Ball, Joshua; Fatihi, Abdelhak et al. (2019): Variation in Expression of the HECT E3 Ligase UPL3 Modulates LEC2 Levels, Seed Size, and Crop Yields in Brassica napus. In *The Plant cell* 31 (10), pp. 2370–2385. DOI: 10.1105/tpc.18.00577.

Monroe, J. Grey; Srikant, Thanvi; Carbonell-Bejerano, Pablo; Becker, Claude; Lensink, Mariele; Exposito-Alonso, Moises et al. (2022): Mutation bias reflects natural selection in Arabidopsis thaliana. In *Nature* 602 (7895), pp. 101–105. DOI: 10.1038/s41586-021-04269-6.

Morales, Nicolas; Ogbonna, Alex C.; Ellerbrock, Bryan J.; Bauchet, Guillaume J.; Tantikanjana, Titima; Tecle, Isaak Y. et al. (2022): Breedbase: a digital ecosystem for modern plant breeding. In *G3 (Bethesda, Md.)* 12 (7). DOI: 10.1093/g3journal/jkac078.

Moreno, Pablo; Fexova, Silvie; George, Nancy; Manning, Jonathan R.; Miao, Zhichiao; Mohammed, Suhaib et al. (2022): Expression Atlas update: gene and protein expression in multiple species. In *Nucleic acids research* 50 (D1), D129-D140. DOI: 10.1093/nar/gkab1030.

Mosher, Rebecca A.; Melnyk, Charles W. (2010): siRNAs and DNA methylation: seedy epigenetics. In *Trends in plant science* 15 (4), pp. 204–210. DOI: 10.1016/j.tplants.2010.01.002.

Muir, Paul; Li, Shantao; Lou, Shaoke; Wang, Daifeng; Spakowicz, Daniel J.; Salichos, Leonidas et al. (2016): The real cost of sequencing: scaling computation to keep pace with data generation. In *Genome biology* 17, p. 53. DOI: 10.1186/s13059-016-0917-0.

Muyle, Aline M.; Seymour, Danelle K.; Lv, Yuanda; Huettel, Bruno; Gaut, Brandon S. (2022): Gene Body Methylation in Plants: Mechanisms, Functions, and Important Implications for Understanding Evolutionary Processes. In *Genome biology and evolution* 14 (4). DOI: 10.1093/gbe/evac038.

Naish, Matthew; Alonge, Michael; Wlodzimierz, Piotr; Tock, Andrew J.; Abramson, Bradley W.; Schmücker, Anna et al. (2021): The genetic and epigenetic landscape of the Arabidopsis centromeres. In *Science (New York, N.Y.)* 374 (6569), eabi7489. DOI: 10.1126/science.abi7489.

Ng, Danny W-K; Lu, Jie; Chen, Z. Jeffrey (2012): Big roles for small RNAs in polyploidy, hybrid vigor, and hybrid incompatibility. In *Current opinion in plant biology* 15 (2), pp. 154–161. DOI: 10.1016/j.pbi.2012.01.007.

Ni, Peng; Huang, Neng; Nie, Fan; Zhang, Jun; Zhang, Zhi; Wu, Bo et al. (2021): Genome-wide detection of cytosine methylations in plant from Nanopore data using deep learning. In *Nature communications* 12 (1), p. 5976. DOI: 10.1038/s41467-021-26278-9.

Niederhuth, Chad E.; Bewick, Adam J.; Ji, Lexiang; Alabady, Magdy S.; Kim, Kyung Do; Li, Qing et al. (2016): Widespread natural variation of DNA methylation within angiosperms. In *Genome biology* 17 (1), p. 194. DOI: 10.1186/s13059-016-1059-0.

Novák, Petr; Guignard, Maïté S.; Neumann, Pavel; Kelly, Laura J.; Mlinarec, Jelena; Koblížková, Andrea et al. (2020): Repeat-sequence turnover shifts fundamentally in species with large genomes. In *Nature plants* 6 (11), pp. 1325–1329. DOI: 10.1038/s41477-020-00785-x.

Nowak, Katarzyna; Morończyk, Joanna; Grzyb, Małgorzata; Szczygieł-Sommer, Aleksandra; Gaj, Małgorzata D. (2022): miR172 Regulates WUS during Somatic Embryogenesis in Arabidopsis via AP2. In *Cells* 11 (4). DOI: 10.3390/cells11040718.

Nowell, P. C.; Hungerford, D. A. (1960): Chromosome Studies on Normal and Leukemic Human Leukocytes. In *JNCI: Journal of the National Cancer Institute* 25, pp. 85–109. DOI: 10.1093/jnci/25.1.85.

Nuthikattu, Saivageethi; McCue, Andrea D.; Panda, Kaushik; Fultz, Dalen; DeFraia, Christopher; Thomas, Erica N.; Slotkin, R. Keith (2013): The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs. In *Plant physiology* 162 (1), pp. 116–131. DOI: 10.1104/pp.113.216481.

Obermeier, Christian; Mason, Annaliese S.; Meiners, Torsten; Petschenka, Georg; Rostás, Michael; Will, Torsten et al. (2022): Perspectives for integrated insect pest protection in oilseed rape breeding. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*. DOI: 10.1007/s00122-022-04074-3.

O'Brien, Jacob; Hayder, Heyam; Zayed, Yara; Peng, Chun (2018): Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. In *Frontiers in endocrinology* 9, p. 402. DOI: 10.3389/fendo.2018.00402.

One Thousand Plant Transcriptomes Initiative (2019): One thousand plant transcriptomes and the phylogenomics of green plants. In *Nature* 574 (7780), pp. 679–685. DOI: 10.1038/s41586-019-1693-2.

Osborn, Thomas C.; Butrulle, David V.; Sharpe, Andrew G.; Pickering, Kathryn J.; Parkin, Isobel A. P.; Parker, John S.; Lydiate, Derek J. (2003): Detection and effects of a homeologous reciprocal transposition in Brassica napus. In *Genetics* 165 (3), pp. 1569–1577. DOI: 10.1093/genetics/165.3.1569.

Otto, Sarah P. (2007): The evolutionary consequences of polyploidy. In *Cell* 131 (3), pp. 452–462. DOI: 10.1016/j.cell.2007.10.022.

Panchy, Nicholas; Lehti-Shiu, Melissa; Shiu, Shin-Han (2016): Evolution of Gene Duplication in Plants. In *Plant physiology* 171 (4), pp. 2294–2316. DOI: 10.1104/pp.16.00523.

Parkin, I. A.; Sharpe, A. G.; Keith, D. J.; Lydiate, D. J. (1995): Identification of the A and C genomes of amphidiploid Brassica napus (oilseed rape). In *Genome* 38 (6), pp. 1122–1131. DOI: 10.1139/g95-149.

Pelé, Alexandre; Rousseau-Gueutin, Mathieu; Chèvre, Anne-Marie (2018): Speciation Success of Polyploid Plants Closely Relates to the Regulation of Meiotic Recombination. In *Frontiers in plant science* 9, p. 907. DOI: 10.3389/fpls.2018.00907.

Pereira, Ricardo; Leitão, José M. (2021): A Non-Rogue Mutant Line Induced by ENU Mutagenesis in Paramutated Rogue Peas (Pisum sativum L.) Is Still Sensitive to the Rogue Paramutation. In *Genes* 12 (11). DOI: 10.3390/genes12111680.

Plieske, J.; Struss, D. (2001): Microsatellite markers for genome analysis in Brassica. I. development in Brassica napus and abundance in Brassicaceae species. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 102 (5), pp. 689–694. DOI: 10.1007/s001220051698.

Pokrovac, Ivan; Pezer, Željka (2022): Recent advances and current challenges in population genomics of structural variation in animals and plants. In *Frontiers in genetics* 13, Article 1060898. DOI: 10.3389/fgene.2022.1060898.

Powers, Leroy (1944): An Expansion of Jones's Theory for the Explanation of Heterosis. In *The American Naturalist* 78 (776), pp. 275–280. DOI: 10.1086/281199.

Provart, Nicholas J.; Alonso, Jose; Assmann, Sarah M.; Bergmann, Dominique; Brady, Siobhan M.; Brkljacic, Jelena et al. (2016): 50 years of Arabidopsis research: highlights and future directions. In *The New phytologist* 209 (3), pp. 921–944. DOI: 10.1111/nph.13687.

Qian, W.; Chen, X.; Fu, D.; Zou, J.; Meng, J. (2005): Intersubgenomic heterosis in seed yield potential observed in a new type of Brassica napus introgressed with partial Brassica rapa genome. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 110 (7), pp. 1187–1194. DOI: 10.1007/s00122-005-1932-9.

Quadrana, Leandro; Colot, Vincent (2016): Plant Transgenerational Epigenetics. In *Annual review of genetics* 50, pp. 467–491. DOI: 10.1146/annurev-genet-120215-035254.

Quezada-Martinez, Daniela; Addo Nyarko, Charles P.; Schiessl, Sarah V.; Mason, Annaliese S. (2021): Using wild relatives and related species to build climate resilience in Brassica crops. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 134 (6), pp. 1711–1728. DOI: 10.1007/s00122-021-03793-3.

Quezada-Martinez, Daniela; Zou, Jun; Zhang, Wenshan; Meng, Jinling; Batley, Jacqueline; Mason, Annaliese S. (2022): Allele segregation analysis of F1 hybrids between independent Brassica allohexaploid lineages. In *Chromosoma* 131 (3), pp. 147–161. DOI: 10.1007/s00412-022-00774-3.

Radoev, Mladen; Becker, Heiko C.; Ecke, Wolfgang (2008): Genetic analysis of heterosis for yield and yield components in rapeseed (Brassica napus L.) by quantitative trait locus mapping. In *Genetics* 179 (3), pp. 1547–1558. DOI: 10.1534/genetics.108.089680.

Rajeevkumar, Sarma; Anunanthini, Pushpanathan; Sathishkumar, Ramalingam (2015): Epigenetic silencing in transgenic plants. In *Frontiers in plant science* 6, p. 693. DOI: 10.3389/fpls.2015.00693.

Ramakrishnan, Muthusamy; Satish, Lakkakula; Kalendar, Ruslan; Narayanan, Mathiyazhagan; Kandasamy, Sabariswaran; Sharma, Anket et al. (2021): The Dynamism of Transposon Methylation for Plant Development and Stress Adaptation. In *International journal of molecular sciences* 22 (21). DOI: 10.3390/ijms222111387.

Raman, Harsh; Dalton-Morgan, Jessica; Diffey, Simon; Raman, Rosy; Alamery, Salman; Edwards, David; Batley, Jacqueline (2014): SNP markers-based map construction and genome-wide linkage analysis in Brassica napus. In *Plant biotechnology journal* 12 (7), pp. 851–860. DOI: 10.1111/pbi.12186.

Ran, Li-ping; Fang, Ting-ting; Rong, Hao; Jiang, Jin-jin; Fang, Yu-jie; Wang, You-ping (2016): Analysis of cytosine methylation in early generations of resynthesized Brassica napus. In *Journal of Integrative Agriculture* 15 (6), pp. 1228–1238. DOI: 10.1016/S2095-3119(15)61277-1.

Rauluseviciute, Ieva; Drabløs, Finn; Rye, Morten Beck (2020): DNA hypermethylation associated with upregulated gene expression in prostate cancer demonstrates the diversity of epigenetic regulation. In *BMC medical genomics* 13 (1), p. 6. DOI: 10.1186/s12920-020-0657-6.

Read, Andrew; Weiss, Trevor; Crisp, Peter A.; Liang, Zhikai; Noshay, Jaclyn; Menard, Claire C. et al. (2022): Genome-wide loss of CHH methylation with limited transcriptome changes in Setaria viridis DOMAINS REARRANGED METHYLTRANSFERASE (DRM) mutants. In *The Plant journal : for cell and molecular biology* 111 (1), pp. 103–116. DOI: 10.1111/tpj.15781.

Rice, Anna; Šmarda, Petr; Novosolov, Maria; Drori, Michal; Glick, Lior; Sabath, Niv et al. (2019): The global biogeography of polyploid plants. In *Nature ecology & evolution* 3 (2), pp. 265–273. DOI: 10.1038/s41559-018-0787-9.

Rong, Hao; Yang, Wenjing; Zhu, Haotian; Jiang, Bo; Jiang, Jinjin; Wang, Youping (2021): Genomic imprinted genes in reciprocal hybrid endosperm of Brassica napus. In *BMC plant biology* 21 (1), p. 140. DOI: 10.1186/s12870-021-02908-8.

Rousseau-Gueutin, Mathieu; Belser, Caroline; Da Silva, Corinne; Richard, Gautier; Istace, Benjamin; Cruaud, Corinne et al. (2020): Long-read assembly of the Brassica napus reference genome Darmor-bzh. In *GigaScience* 9 (12). DOI: 10.1093/gigascience/giaa137.

Roy, Jayanta; Shaikh, T. M.; Del Río Mendoza, Luis; Hosain, Shakil; Chapara, Venkat; Rahman, Mukhlesur (2021): Genome-wide association mapping and genomic prediction for adult stage sclerotinia stem rot resistance in Brassica napus (L) under field environments. In *Scientific reports* 11 (1), p. 21773. DOI: 10.1038/s41598-021-01272-9.

Samans, Birgit; Chalhoub, Boulos; Snowdon, Rod J. (2017): Surviving a Genome Collision: Genomic Signatures of Allopolyploidization in the Recent Crop Species Brassica napus. In *The plant genome* 10 (3). DOI: 10.3835/plantgenome2017.02.0013.

Saxena, Rachit K.; Edwards, David; Varshney, Rajeev K. (2014): Structural variations in plant genomes. In *Briefings in functional genomics* 13 (4), pp. 296–307. DOI: 10.1093/bfgp/elu016.

Saxena, Rachit K.; Jiang, Yong; Khan, Aamir W.; Zhao, Yusheng; Kumar Singh, Vikas; Bohra, Abhishek et al. (2021): Characterization of heterosis and genomic prediction-based establishment of heterotic patterns for developing better hybrids in pigeonpea. In *The plant genome* 14 (3), e20125. DOI: 10.1002/tpg2.20125.

Schiessl, Sarah (2020): Regulation and Subfunctionalization of Flowering Time Genes in the Allotetraploid Oil Crop Brassica napus. In *Frontiers in plant science* 11, p. 605155. DOI: 10.3389/fpls.2020.605155.

Schiessl, Sarah; Huettel, Bruno; Kuehn, Diana; Reinhardt, Richard; Snowdon, Rod (2017): Post-polyploidisation morphotype diversification associates with gene copy number variation. In *Scientific reports* 7, p. 41845. DOI: 10.1038/srep41845.

Schiessl, Sarah V.; Quezada-Martinez, Daniela; Orantes-Bonilla, Mauricio; Snowdon, Rod J. (2020): Transcriptomics reveal high regulatory diversity of drought tolerance strategies in a biennial oil crop. In *Plant science : an international journal of experimental plant biology* 297, p. 110515. DOI: 10.1016/j.plantsci.2020.110515.

Schiessl, Sarah V.; Quezada-Martinez, Daniela; Tebartz, Ellen; Snowdon, Rod J.; Qian, Lunwen (2019a): The vernalisation regulator FLOWERING LOCUS C is differentially expressed in biennial and annual Brassica napus. In *Scientific reports* 9 (1), p. 14911. DOI: 10.1038/s41598-019-51212-x.

Schiessl, Sarah-Veronica; Katche, Elvis; Ihien, Elizabeth; Chawla, Harmeet Singh; Mason, Annaliese S. (2019b): The role of genomic structural variation in the genetic improvement of polyploid crops. In *The Crop Journal* 7 (2), pp. 127–140. DOI: 10.1016/j.cj.2018.07.006.

Schmidt, Bertil; Hildebrandt, Andreas (2017): Next-generation sequencing: big data meets high performance computing. In *Drug discovery today* 22 (4), pp. 712–717. DOI: 10.1016/j.drudis.2017.01.014.

Seifert, Felix; Thiemann, Alexander; Grant-Downton, Robert; Edelmann, Susanne; Rybka, Dominika; Schrag, Tobias A. et al. (2018a): Parental Expression Variation of Small RNAs Is Negatively Correlated with Grain Yield Heterosis in a Maize Breeding Population. In *Frontiers in plant science* 9, p. 13. DOI: 10.3389/fpls.2018.00013.

Seifert, Felix; Thiemann, Alexander; Schrag, Tobias A.; Rybka, Dominika; Melchinger, Albrecht E.; Frisch, Matthias; Scholten, Stefan (2018b): Small RNA-based prediction of hybrid performance in maize. In *BMC genomics* 19 (1), p. 371. DOI: 10.1186/s12864-018-4708-8.

Shahzad, Kashif; Zhang, Xuexian; Guo, Liping; Qi, Tingxiang; Tang, Huini; Zhang, Meng et al. (2020): Comparative transcriptome analysis of inbred lines and contrasting hybrids reveals overdominance mediate early biomass vigor in hybrid cotton. In *BMC genomics* 21 (1), p. 140. DOI: 10.1186/s12864-020-6561-9.

Shaikh, Aamir Ali; Chachar, Sadaruddin; Chachar, Muzafaruddin; Ahmed, Nazir; Guan, Changfei; Zhang, Pingxian (2022): Recent Advances in DNA Methylation and Their Potential Breeding Applications in Plants. In *Horticulturae* 8 (7), p. 562. DOI: 10.3390/horticulturae8070562.

Sharma, Eshan; Jain, Mukesh; Khurana, Jitendra P. (2019): Differential quantitative regulation of specific gene groups and pathways under drought stress in rice. In *Genomics* 111 (6), pp. 1699–1712. DOI: 10.1016/j.ygeno.2018.11.024.

Sharpe, A. G.; Parkin, I. A.; Keith, D. J.; Lydiate, D. J. (1995): Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (Brassica napus). In *Genome* 38 (6), pp. 1112–1121. DOI: 10.1139/g95-148.

Shen, Yifei; Sun, Shuo; Hua, Shuijin; Shen, Enhui; Ye, Chu-Yu; Cai, Daguang et al. (2017): Analysis of transcriptional and epigenetic changes in hybrid vigor of allopolyploid Brassica napus uncovers key roles for small RNAs. In *The Plant journal : for cell and molecular biology* 91 (5), pp. 874–893. DOI: 10.1111/tpj.13605.

Shi, Jiaqin; Li, Ruiyuan; Zou, Jun; Long, Yan; Meng, Jinling (2011): A dynamic and complex network regulates the heterosis of yield-correlated traits in rapeseed (Brassica napus L.). In *PloS one* 6 (7), e21645. DOI: 10.1371/journal.pone.0021645.

Shi, Xinghua; Radhakrishnan, Saranya; Wen, Jia; Chen, Jin Yun; Chen, Junjie; Lam, Brianna Ashlyn et al. (2020): Association of CNVs with methylation variation. In *NPJ genomic medicine* 5, p. 41. DOI: 10.1038/s41525-020-00145-w.

Shin, Hosub; Park, Jeong Eun; Park, Hye Rang; Choi, Woo Lee; Yu, Seung Hwa; Koh, Wonjun et al. (2022): Admixture of divergent genomes facilitates hybridization across species in the family Brassicaceae. In *The New phytologist* 235 (2), pp. 743–758. DOI: 10.1111/nph.18155.

Shull, G. H. (1908): The Composition of a Field of Maize. In *Journal of Heredity* os-4 (1), pp. 296–301. DOI: 10.1093/jhered/os-4.1.296.

Shull, G. H. (1914): Duplicate genes for capsule-form inBursa bursa-pastoris. In *Z.Ver-erbungslehre* 12 (1), pp. 97–149. DOI: 10.1007/BF01837282.

Shull, G. H. (1948): What Is "Heterosis"? In *Genetics* 33 (5), pp. 439–446. DOI: 10.1093/genetics/33.5.439.

Snowdon, Rod J.; Abbadi, Amine; Kox, Tobias; Schmutzer, Thomas; Leckband, Gunhild (2015): Heterotic Haplotype Capture: precision breeding for hybrid performance. In *Trends in plant science* 20 (7), pp. 410–413. DOI: 10.1016/j.tplants.2015.04.013.

Soltis, Douglas E.; Segovia-Salcedo, María Claudia; Jordon-Thaden, Ingrid; Majure, Lucas; Miles, Nicolas M.; Mavrodiev, Evgeny V. et al. (2014): Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose et al.. In *The New phytologist* 202 (4), pp. 1105–1117. DOI: 10.1111/nph.12756.

Song, Gui-Sheng; Zhai, Hong-Li; Peng, Yong-Gang; Zhang, Lei; Wei, Gang; Chen, Xiao-Ying et al. (2010): Comparative transcriptional profiling and preliminary study on heterosis mechanism of super-hybrid rice. In *Molecular plant* 3 (6), pp. 1012–1025. DOI: 10.1093/mp/ssq046.

Song, Jia-Ming; Guan, Zhilin; Hu, Jianlin; Guo, Chaocheng; Yang, Zhiquan; Wang, Shuo et al. (2020): Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of Brassica napus. In *Nature plants* 6 (1), pp. 34–45. DOI: 10.1038/s41477-019-0577-7.

Song, K.; Lu, P.; Tang, K.; Osborn, T. C. (1995): Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution. In *Proceedings of the National Academy of Sciences of the United States of America* 92 (17), pp. 7719–7723. DOI: 10.1073/pnas.92.17.7719.

Song, Yongfeng; Lv, Zhichao; Wang, Yue; Li, Chunxiang; Jia, Yue; Zhu, Yong et al. (2022): Identification of miRNAs Mediating Seed Storability of Maize during Germination Stage by High-Throughput Sequencing, Transcriptome and Degradome Sequencing. In *International journal of molecular sciences* 23 (20). DOI: 10.3390/ijms232012339.

Springer, Nathan M.; Ying, Kai; Fu, Yan; Ji, Tieming; Yeh, Cheng-Ting; Jia, Yi et al. (2009): Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. In *PLoS genetics* 5 (11), e1000734. DOI: 10.1371/journal.pgen.1000734.

Stange, Madlen; Barrett, Rowan D. H.; Hendry, Andrew P. (2021): The importance of genomic variation for biodiversity, ecosystems and people. In *Nature reviews. Genetics* 22 (2), pp. 89–105. DOI: 10.1038/s41576-020-00288-7.

Stearns, Frank W. (2010): One hundred years of pleiotropy: a retrospective. In *Genetics* 186 (3), pp. 767–773. DOI: 10.1534/genetics.110.122549.

Stebbins, G. Ledyard (1945): The cytological analysis of species hybrids. II. In *Bot. Rev* 11 (9), pp. 463–486. DOI: 10.1007/BF02861140.

Steeg, Emily M. S. ter; Struik, Paul C.; Visser, Richard G. F.; Lindhout, Pim (2022): Crucial factors for the feasibility of commercial hybrid breeding in food crops. In *Nature plants* 8 (5), pp. 463–473. DOI: 10.1038/s41477-022-01142-w.

Stein, Anna; Coriton, Olivier; Rousseau-Gueutin, Mathieu; Samans, Birgit; Schiessl, Sarah V.; Obermeier, Christian et al. (2017): Mapping of homoeologous chromosome exchanges influencing quantitative trait variation in Brassica napus. In *Plant biotechnology journal* 15 (11), pp. 1478–1489. DOI: 10.1111/pbi.12732.

Stroud, Hume; Greenberg, Maxim V. C.; Feng, Suhua; Bernatavichute, Yana V.; Jacobsen, Steven E. (2013): Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. In *Cell* 152 (1-2), pp. 352–364. DOI: 10.1016/j.cell.2012.10.054.

Sun, Silong; Zhou, Yingsi; Chen, Jian; Shi, Junpeng; Zhao, Haiming; Zhao, Hainan et al. (2018): Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes. In *Nature genetics* 50 (9), pp. 1289–1295. DOI: 10.1038/s41588-018-0182-0.

Sun, Zhizhong; Peng, Jianxiang; Lv, Qiming; Ding, Jia; Chen, Siyang; Duan, Meijuan et al. (2022): Dissecting the genetic basis of the heterosis of Y900, an elite super-hybrid rice.

Sutton, W. S. (1902): On the morphology of the chromoso group in Brachystola magna. In *The Biological Bulletin* 4 (24-39).

Szadkowski, E.; Eber, F.; Huteau, V.; Lodé, M.; Huneau, C.; Belcram, H. et al. (2010): The first meiosis of resynthesized Brassica napus, a genome blender. In *The New phytologist* 186 (1), pp. 102–112. DOI: 10.1111/j.1469-8137.2010.03182.x.

Takei, Yodai; Yun, Jina; Zheng, Shiwei; Ollikainen, Noah; Pierson, Nico; White, Jonathan et al. (2021): Integrated spatial genomics reveals global architecture of single nuclei. In *Nature* 590 (7845), pp. 344–350. DOI: 10.1038/s41586-020-03126-2.

Tan, Zengdong; Peng, Yan; Xiong, Yao; Xiong, Feng; Zhang, Yuting; Guo, Ning et al. (2022): Comprehensive transcriptional variability analysis reveals gene networks regulating seed oil content of Brassica napus. In *Genome biology* 23 (1), p. 233. DOI: 10.1186/s13059-022-02801-z.

Tang, Jiuyou; Chu, Chengcai (2017): MicroRNAs in crop improvement: fine-tuners for complex traits. In *Nature plants* 3, p. 17077. DOI: 10.1038/nplants.2017.77.

Tang, Shan; Zhao, Hu; Lu, Shaoping; Yu, Liangqian; Zhang, Guofang; Zhang, Yuting et al. (2021): Genome- and transcriptome-wide association studies provide insights into the genetic basis of natural

variation of seed oil content in Brassica napus. In *Molecular plant* 14 (3), pp. 470–487. DOI: 10.1016/j.molp.2020.12.003.

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

Udall, Joshua A.; Quijada, Pablo A.; Osborn, Thomas C. (2005): Detection of chromosomal rearrangements derived from homologous recombination in four mapping populations of Brassica napus L. In *Genetics* 169 (2), pp. 967–979. DOI: 10.1534/genetics.104.033209.

van de Peer, Yves; Mizrachi, Eshchar; Marchal, Kathleen (2017): The evolutionary significance of polyploidy. In *Nature reviews. Genetics* 18 (7), pp. 411–424. DOI: 10.1038/nrg.2017.26.

Varshney, Rajeev K.; Shi, Chengcheng; Thudi, Mahendar; Mariac, Cedric; Wallace, Jason; Qi, Peng et al. (2017): Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. In *Nature biotechnology* 35 (10), pp. 969–976. DOI: 10.1038/nbt.3943.

Vollrath, Paul; Chawla, Harmeet S.; Alnajar, Dima; Gabur, Iulian; Lee, HueyTyng; Weber, Sven et al. (2021a): Dissection of Quantitative Blackleg Resistance Reveals Novel Variants of Resistance Gene Rlm9 in Elite Brassica napus. In *Frontiers in plant science* 12, p. 749491. DOI: 10.3389/fpls.2021.749491.

Vollrath, Paul; Chawla, Harmeet S.; Schiessl, Sarah V.; Gabur, Iulian; Lee, HueyTyng; Snowdon, Rod J.; Obermeier, Christian (2021b): A novel deletion in FLOWERING LOCUS T modulates flowering time in winter oilseed rape. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 134 (4), pp. 1217–1231. DOI: 10.1007/s00122-021-03768-4.

Wagner, M. (1954): Inzucht und Heterosis bei Raps und Rübsen. In *Z. Pflanzenzüchtg* 33, pp. 237–266.

Walden, Nora; German, Dmitry A.; Wolf, Eva M.; Kiefer, Markus; Rigault, Philippe; Huang, Xiao-Chen et al. (2020): Nested whole-genome duplications coincide with diversification and high morphological disparity in Brassicaceae. In *Nature communications* 11 (1), p. 3795. DOI: 10.1038/s41467-020-17605-7.

Walkowiak, Sean; Gao, Liangliang; Monat, Cecile; Haberer, Georg; Kassa, Mulualem T.; Brinton, Jemima et al. (2020): Multiple wheat genomes reveal global variation in modern breeding. In *Nature* 588 (7837), pp. 277–283. DOI: 10.1038/s41586-020-2961-x.

Wang, Bo; Tseng, Elizabeth; Baybayan, Primo; Eng, Kevin; Regulski, Michael; Jiao, Yinping et al. (2020): Variant phasing and haplotypic expression from long-read sequencing in maize. In *Communications biology* 3 (1), p. 78. DOI: 10.1038/s42003-020-0805-8.

Wang, Bo; Wu, Zhikun; Li, Zhaohong; Zhang, Qinghua; Hu, Jianlin; Xiao, Yingjie et al. (2018): Dissection of the genetic architecture of three seed-quality traits and consequences for breeding in Brassica napus. In *Plant biotechnology journal* 16 (7), pp. 1336–1348. DOI: 10.1111/pbi.12873.

Wang, Jia; Jian, Hongju; Wang, Tengyue; Wei, Lijuan; Li, Jiana; Li, Chao; Liu, Liezhao (2016): Identification of microRNAs Actively Involved in Fatty Acid Biosynthesis in Developing Brassica napus Seeds Using High-Throughput Sequencing. In *Frontiers in plant science* 7, p. 1570. DOI: 10.3389/fpls.2016.01570.

Wang, Li; Wu, Li Min; Greaves, Ian K.; Zhu, Anyu; Dennis, Elizabeth S.; Peacock, W. James (2017): PIF4-controlled auxin pathway contributes to hybrid vigor in Arabidopsis thaliana. In *Proceedings of the National Academy of Sciences of the United States of America* 114 (17), E3555-E3562. DOI: 10.1073/pnas.1703179114.

Wang, Long; Ji, Yilun; Hu, Yingwen; Hu, Huaying; Jia, Xianqin; Jiang, Mengmeng et al. (2019a): The architecture of intra-organism mutation rate variation in plants. In *PLoS biology* 17 (4), e3000191. DOI: 10.1371/journal.pbio.3000191.

Wang, Mengyao; Wang, Jianbo (2022): Transcriptome and DNA methylome analyses provide insight into the heterosis in flag leaf of inter-subspecific hybrid rice. In *Plant molecular biology* 108 (1-2), pp. 105–125. DOI: 10.1007/s11103-021-01228-7.

Wang, Tengyue; Ping, Xiaoke; Cao, Yanru; Jian, Hongju; Gao, Yumin; Wang, Jia et al. (2019b): Genome-wide exploration and characterization of miR172/euAP2 genes in Brassica napus L. for likely role in flower organ development. In *BMC plant biology* 19 (1), p. 336. DOI: 10.1186/s12870-019-1936-2.

Wasserman, M. (1968): Recombination-induced chromosomal heterosis. In *Genetics* 58 (1), pp. 125–139. DOI: 10.1093/genetics/58.1.125.

Wei, Yunxiao; Li, Guoliang; Zhang, Shujiang; Zhang, Shifan; Zhang, Hui; Sun, Rifei et al. (2021): Analysis of Transcriptional Changes in Different Brassica napus Synthetic Allopolyploids. In *Genes* 12 (1). DOI: 10.3390/genes12010082.

Weiss, Trevor; Crisp, Peter A.; Rai, Krishan M.; Song, Meredith; Springer, Nathan M.; Zhang, Feng (2022): Epigenetic features drastically impact CRISPR-Cas9 efficacy in plants. In *Plant physiology* 190 (2), pp. 1153–1164. DOI: 10.1093/plphys/kiac285.

Werner, Christian R.; Qian, Lunwen; Voss-Fels, Kai P.; Abbadi, Amine; Leckband, Gunhild; Frisch, Matthias; Snowdon, Rod J. (2018): Genome-wide regression models considering general and specific combining ability predict hybrid performance in oilseed rape with similar accuracy regardless of trait architecture. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 131 (2), pp. 299–317. DOI: 10.1007/s00122-017-3002-5.

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, I. Jsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie et al. (2016): The FAIR Guiding Principles for scientific data management and stewardship. In *Scientific data* 3, p. 160018. DOI: 10.1038/sdata.2016.18.

Wood, Troy E.; Takebayashi, Naoki; Barker, Michael S.; Mayrose, Itay; Greenspoon, Philip B.; Rieseberg, Loren H. (2009): The frequency of polyploid speciation in vascular plants. In *Proceedings of the National Academy of Sciences of the United States of America* 106 (33), pp. 13875–13879. DOI: 10.1073/pnas.0811575106.

Wu, Dezhi; Liang, Zhe; Yan, Tao; Xu, Ying; Xuan, Lijie; Tang, Juan et al. (2019): Whole-Genome Resequencing of a Worldwide Collection of Rapeseed Accessions Reveals the Genetic Basis of Ecotype Divergence. In *Molecular plant* 12 (1), pp. 30–43. DOI: 10.1016/j.molp.2018.11.007.

Wu, Jian; Lin, Li; Xu, Meiling; Chen, Peipei; Liu, Dongxiao; Sun, Qinfu et al. (2018): Homoeolog expression bias and expression level dominance in resynthesized allopolyploid Brassica napus. In *BMC genomics* 19 (1), p. 586. DOI: 10.1186/s12864-018-4966-5.

Wu, Jianzhong; Sun, Dequan; Zhao, Qian; Yong, Hongjun; Zhang, Degui; Hao, Zhuanfang et al. (2021a): Transcriptome Reveals Allele Contribution to Heterosis in Maize. In *Frontiers in plant science* 12, p. 739072. DOI: 10.3389/fpls.2021.739072.

Wu, Xilin; Liu, Yan; Zhang, Yaowei; Gu, Ran (2021b): Advances in Research on the Mechanism of Heterosis in Plants. In *Frontiers in plant science* 12, p. 745726. DOI: 10.3389/fpls.2021.745726.

Wu, Ying; Lin, Fan; Zhou, Yao; Wang, Jie; Sun, Shuai; Wang, Bin et al. (2021c): Genomic mosaicism due to homoeologous exchange generates extensive phenotypic diversity in nascent allopolyploids. In *National science review* 8 (5), nwaa277. DOI: 10.1093/nsr/nwaa277.

Xia, Shengqian; Wang, Zhixin; Zhang, Haiyan; Hu, Kaining; Zhang, Zhiqiang; Qin, Maomao et al. (2016): Altered Transcription and Neofunctionalization of Duplicated Genes Rescue the Harmful Effects of a Chimeric Gene in Brassica napus. In *The Plant cell* 28 (9), pp. 2060–2078. DOI: 10.1105/tpc.16.00281.

Xiong, Jie; Hu, Kaining; Shalby, Nesma; Zhuo, Chenjian; Wen, Jing; Yi, Bin et al. (2022): Comparative transcriptomic analysis reveals the molecular mechanism underlying seedling biomass heterosis in Brassica napus. In *BMC plant biology* 22 (1), p. 283. DOI: 10.1186/s12870-022-03671-0.

Xiong, Zhiyong; Gaeta, Robert T.; Pires, J. Chris (2011): Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid Brassica napus. In *Proceedings of the National Academy of Sciences of the United States of America* 108 (19), pp. 7908–7913. DOI: 10.1073/pnas.1014138108.

Xu, Jing; Chen, Guo; Hermanson, Peter J.; Xu, Qiang; Sun, Changshuo; Chen, Wenqing et al. (2019): Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize. In *Genome biology* 20 (1), p. 243. DOI: 10.1186/s13059-019-1859-0.

Yang, Sha; Zhang, Zhuqing; Chen, Wenchao; Li, Xuefeng; Zhou, Shudong; Liang, Chengliang et al. (2021): Integration of mRNA and miRNA profiling reveals the heterosis of three hybrid combinations of Capsicum annuum varieties. In *GM crops & food* 12 (1), pp. 224–241. DOI: 10.1080/21645698.2020.1852064.

Yeaman, Sam (2013): Genomic rearrangements and the evolution of clusters of locally adaptive loci. In *Proceedings of the National Academy of Sciences of the United States of America* 110 (19), E1743-51. DOI: 10.1073/pnas.1219381110.

Yoo, M-J; Szadkowski, E.; Wendel, J. F. (2013): Homoeolog expression bias and expression level dominance in allopolyploid cotton. In *Heredity* 110 (2), pp. 171–180. DOI: 10.1038/hdy.2012.94.

Yu, Ning; Gan, Guiyun; Li, Dandan; Li, Weiliu; Jiang, Yaqin; Yang, Qihong et al. (2022): Genetic changes in F1 hybrids of the genetically divergent Solanum L., Solanum melongena L. × Solanum aethiopicum L. In *Plant Growth Regul*. DOI: 10.1007/s10725-022-00926-6.

Yu, Xiaowen; Zhao, Zhigang; Zheng, Xiaoming; Zhou, Jiawu; Kong, Weiyi; Wang, Peiran et al. (2018): A selfish genetic element confers non-Mendelian inheritance in rice. In *Science (New York, N.Y.)* 360 (6393), pp. 1130–1132. DOI: 10.1126/science.aar4279.

Yuan, Yuxuan; Bayer, Philipp E.; Batley, Jacqueline; Edwards, David (2021): Current status of structural variation studies in plants. In *Plant biotechnology journal* 19 (11), pp. 2153–2163. DOI: 10.1111/pbi.13646.

Zanini, Silvia F.; Bayer, Philipp E.; Wells, Rachel; Snowdon, Rod J.; Batley, Jacqueline; Varshney, Rajeev K. et al. (2022): Pangenomics in crop improvement-from coding structural variations to finding regulatory variants with pangenome graphs. In *The plant genome* 15 (1), e20177. DOI: 10.1002/tpg2.20177.

Zemach, Assaf; Kim, M. Yvonne; Silva, Pedro; Rodrigues, Jessica A.; Dotson, Bradley; Brooks, Matthew D.; Zilberman, Daniel (2010): Local DNA hypomethylation activates genes in rice endosperm. In *Proceedings of the National Academy of Sciences of the United States of America* 107 (43), pp. 18729–18734. DOI: 10.1073/pnas.1009695107.

Zhang, Chao; Chang, Wei; Li, Xiaodong; Yang, Bo; Zhang, Liyuan; Xiao, Zhongchun et al. (2022): Transcriptome and Small RNA Sequencing Reveal the Mechanisms Regulating Harvest Index in Brassica napus. In *Frontiers in plant science* 13, p. 855486. DOI: 10.3389/fpls.2022.855486.

Zhang, Libin; He, Jianjie; He, Hongsheng; Wu, Jiangsheng; Li, Maoteng (2021): Genome-wide unbalanced expression bias and expression level dominance toward Brassica oleracea in artificially synthesized intergeneric hybrids of Raphanobrassica. In *Horticulture research* 8 (1), p. 246. DOI: 10.1038/s41438-021-00672-2.

Zhang, Mei; Xie, Shaojun; Dong, Xiaomei; Zhao, Xin; Zeng, Biao; Chen, Jian et al. (2014): Genome-wide high resolution parental-specific DNA and histone methylation maps uncover patterns of imprinting regulation in maize. In *Genome research* 24 (1), pp. 167–176. DOI: 10.1101/gr.155879.113.

Zhang, Xingtan; Chen, Xuequn; Liang, Pingping; Tang, Haibao (2018): Cataloging Plant Genome Structural Variations. In *Current issues in molecular biology* 27, pp. 181–194. DOI: 10.21775/cimb.027.181.

Zhao, Chuanji; Xie, Meili; Liang, Longbing; Yang, Li; Han, Hongshi; Qin, Xinrong et al. (2022): Genome-Wide Association Analysis Combined With Quantitative Trait Loci Mapping and Dynamic Transcriptome Unveil the Genetic Control of Seed Oil Content in Brassica napus L. In *Frontiers in plant science* 13, p. 929197. DOI: 10.3389/fpls.2022.929197.

Zhao, Xinwang; Li, Bao; Zhang, Ka; Hu, Kaining; Yi, Bin; Wen, Jing et al. (2016): Breeding signature of combining ability improvement revealed by a genomic variation map from recurrent selection population in Brassica napus. In *Scientific reports* 6, p. 29553. DOI: 10.1038/srep29553.

Zhao, Ying-Tao; Wang, Meng; Fu, San-Xiong; Yang, Wei-Cai; Qi, Cun-Kou; Wang, Xiu-Jie (2012): Small RNA profiling in two Brassica napus cultivars identifies microRNAs with oil production- and development-correlated expression and new small RNA classes. In *Plant physiology* 158 (2), pp. 813–823. DOI: 10.1104/pp.111.187666.

Zhou, Shirong; Xing, Meiqing; Zhao, Zhilong; Gu, Yincong; Xiao, Yunping; Liu, Qiaoquan; Xue, Hongwei (2021): DNA methylation modification in heterosis initiation through analyzing rice hybrid contemporary seeds. In *The Crop Journal* 9 (5), pp. 1179–1190. DOI: 10.1016/j.cj.2020.12.003.

Zhu, Anyu; Wang, Aihua; Zhang, You; Dennis, Elizabeth S.; Peacock, W. James; Greaves, And Ian K. (2020): Early Establishment of Photosynthesis and Auxin Biosynthesis Plays a Key Role in Early Biomass Heterosis in Brassica napus (Canola) Hybrids. In *Plant & cell physiology* 61 (6), pp. 1134–1143. DOI: 10.1093/pcp/pcaa038.

Ziegler, Dylan J.; Khan, Deirdre; Kalichuk, Jenna L.; Becker, Michael G.; Belmonte, Mark F. (2019): Transcriptome landscape of the early Brassica napus seed. In *Journal of integrative plant biology* 61 (5), pp. 639–650. DOI: 10.1111/jipb.12812.

Zou, Jun; Hu, Dandan; Mason, Annaliese S.; Shen, Xueqi; Wang, Xiaohua; Wang, Nian et al. (2018): Genetic changes in a novel breeding population of Brassica napus synthesized from hundreds of crosses between B. rapa and B. carinata. In *Plant biotechnology journal* 16 (2), pp. 507–519. DOI: 10.1111/pbi.12791.

Zou, Jun; Zhu, Jiali; Huang, Shunmou; Tian, Entang; Xiao, Yong; Fu, Donghui et al. (2010): Broadening the avenue of intersubgenomic heterosis in oilseed Brassica. In *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 120 (2), pp. 283–290. DOI: 10.1007/s00122-009-1201-4.

List of abbreviations

- AFLP: Amplified fragment length polymorphisms
- AP2-like: APETALA2-like
- ATHSP: Arabidopsis heat shock promoter
- C: Cytosine
- cDNA: Complementary DNA
- CMS: Cytoplasmic male sterility
- CMT3: Chromomethylase 3
- CNV: Copy number variation
- CO: Crossover
- CpGi: CpG islands
- DEG: Differentially expressed genes
- DE-miRNA: Differentially expressed miRNA
- DE-siRNA: Differentially expressed siRNA
- DMR: Differentially methylated region
- DNA: Deoxyribonucleic acid
- DRM: Domains rearranged methyltransferase
- DSB: Double-strand breaks
- DWARF: DWARF protein
- ELD: Expression level dominance
- EMB: Embryo defective
- eQTL: Expression quantitative trait loci
- FDR: False discovery rate
- FLC: FLOWERING LOCUS C
- GO: Gene ontology
- GS: Genomic selection
- GWAS: Genome-wide association studies
- HD-ZIP III: Class III homeodomain leucine zipper
- HE: Homoeologous exchange
- Indels: Insertion-deletion polymorphisms
- LEC2: LEAFY COTYLEDON 2
- LRES: Long range epigenetic silencing
- mC: Methylated cytosine
- MET1: Methyltransferase 1
- MOI: Multiomics integration
- miRNA: MicroRNA
- mRNA: Messenger RNA
- NGS: Next-generation sequencing
- NHEJ: Non-homologous end-joining
- NRHE: Non-reciprocal homoeologous exchanges
- ONT: Oxford Nanopore Technology
- PacBio: Pacific Biosciences
- PAV: Presence-absence variation
- PCR: Polymerase chain reaction
- PHB: PHABULOSA
- PHV: PHAVOLUTA

- QTL: Quantitative trait loci
- RdDM: RNA-directed DNA methylation
- RFLP: Restriction fragment length polymorphism
- RHE: Reciprocal homoeologous exchanges
- RNA: Ribonucleic acid
- RNA-Seq: RNA sequencing
- SAS: Systemic acquired silencing
- siRNA: Small interfering RNA
- SNP: Single nucleotide polymorphism
- sRNA: Small RNA
- SSR: Simple sequence repeats
- SV: Structural variation
- TE: Transposable element
- TFL 1: TERMINAL FLOWER 1
- UMR: Unmethylated regions
- WGBS: Whole genome bisulfite sequencing
- WGD: Whole genome duplication
- WRKY 4: WRKY DNA-binding protein 4

Declaration

I declare that the dissertation here submitted is entirely my own work, written without any illegitimate help by any third party and solely with materials as indicated in the dissertation. I have indicated in the text where I have used texts from already published sources, either word for word or in substance, and where I have made statements based on oral information given to me. At all times during the investigations carried out by me and described in the dissertation, I have followed the principles of good scientific practice as defined in the 'Statutes of Justus Liebig University Giessen for the Safeguarding of Good Scientific Practice'

Acknowledgements

I would like to express my sincere gratitude to Prof. Rod Snowdon for giving me the chance and support to pursue a PhD degree. I learned skills that I had only heard of in lectures and seemed so far-away for me to understand back then. I would also like to thank Rod for taking time for discussions and showing me to think outside the box for problem solving. Likewise, I would like to extend my appreciation to our project partners at Huazhong University of Agriculture in Wuhan, who despite the pandemic, did their best to cooperate; hence, special thanks go to Prof. Jun Zou and her team. In the same manner, the work described in this thesis was carried thanks to the financial support from DFG and NCSI, and to the bioinformatic resources from the denbi network and BCF from JLU-Giessen. Furthermore, I would like to kindly thank the Plant Breeding Institute colleagues in JLU-Giessen for their help with field and lab work.

A deep, most-sincere thanks go to my parents, Eduardo and Isis, who have showed me beyond words what love and sacrifice means. For their 24/7 reachability to answer my phone calls, even at 3 AM, to listen to me and cheer me up, and for their unconditional support and patience: I greatly thank them! Likewise, I thank my partner in surf and brother, Eduardo, for being here with and for me. Also, I thank my relatives Mary, Mami Estela, Mercy, Julio and Emilio for their support!

Finally, the great value of this PhD for me goes far beyond the academic side and rather roots itself in comprehending better who I am and realizing, among all distresses, the love from God and my family towards me; and for that, I am forever grateful.