the plant journal

The Plant Journal (2017)



RESOURCE

Analysis of transcriptional and epigenetic changes in hybrid vigor of allopolyploid *Brassica napus* uncovers key roles for small RNAs

Yifei Shen¹, Shuo Sun¹, Shuijin Hua², Enhui Shen¹, Chu-Yu Ye¹, Daguang Cai³, Michael P. Timko⁴, Qian-Hao Zhu⁵ and Longjiang Fan^{1,*}

¹Institute of Crop Science and Institute of Bioinformatics, Zhejiang University, Hangzhou 310058, China,

²Institute of Crop and Utilization of Nuclear Technology, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China,

³Institute of Phytopathology, Christian Albrechts University of Kiel, Hermann Rodewald Str. 9, D-24118 Kiel, Germany,
⁴Department of Biology, University of Virginia, Charlottesville, VA 22903, USA, and
⁵CSIRO Agriculture and Food, Black Mountain Laboratories, Canberra, ACT 2601, Australia

Received 13 January 2017; revised 8 May 2017; accepted 17 May 2017. *For correspondence (e-mail fanlj@zju.edu.cn).

SUMMARY

Heterosis is a fundamental biological phenomenon characterized by the superior performance of a hybrid compared with its parents. The underlying molecular basis for heterosis, particularly for allopolyploids, remains elusive. In this study we analyzed the transcriptomes of *Brassica napus* parental lines and their F_1 hybrids at three stages of early flower development. Phenotypically, the F₁ hybrids show remarkable heterosis in silique number and grain yield. Transcriptome analysis revealed that various phytohormone (auxin and salicylic acid) response genes are significantly altered in the F_1 hybrids relative to the parental lines. We also found evidence for decreased expression divergence of the homoeologous gene pairs in the allopolyploid F₁ hybrids and suggest that high-parental expression-level dominance plays an important role in heterosis. Small RNA and methylation studies aimed at examining the epigenetic effect of the changes in gene expression level in the F₁ hybrids showed that the majority of the small interfering RNA (siRNA) clusters had a higher expression level in the F1 hybrids than in the parents, and that there was an increase in genome-wide DNA methylation in the F1 hybrid. Transposable elements associated with siRNA clusters had a higher level of methylation and a lower expression level in the F₁ hybrid, implying that the non-additively expressed siRNA clusters resulted in lower activity of the transposable elements through DNA methylation in the hybrid. Our data provide insights into the role that changes in gene expression pattern and epigenetic mechanisms contribute to heterosis during early flower development in allopolyploid B. napus.

Keywords: Brassica napus, heterosis, allopolyploid, transcriptome, miRNA, siRNA, DNA methylation.

INTRODUCTION

Heterosis is a complex biological phenomenon in which hybrids exhibit superior phenotypic characteristics, such as enhanced biomass production, developmental rates, grain yields and stress tolerance, relative to their parents (Birchler *et al.*, 2003, 2010; Hochholdinger and Hoecker, 2007; Chen, 2010). Although heterosis has been exploited extensively in crop production for many decades, the molecular mechanisms underlying the phenomenon remain largely unknown despite more than a century of study (Birchler

@ 2017 The Authors The Plant Journal @ 2017 John Wiley & Sons Ltd

et al., 2003, 2010; Hochholdinger and Hoecker, 2007; Chen, 2010). Dominance, overdominance and epistasis have each been proposed as genetic explanations for heterosis from a quantitative perspective (Crow, 1948; Yu *et al.*, 1997). However, these hypotheses are largely conceptual and not connected to molecular principles, and therefore they fall short in explaining the molecular basis of heterosis (Birchler *et al.*, 2003; Hochholdinger and Hoecker, 2007; He *et al.*, 2011).

Several studies have shed light on the potential molecular mechanisms of heterosis in plants. A single heterozygous gene, SINGLE FLOWER TRUSS, was shown to cause hybrid vigor in tomato (Solanum lycopersicum), providing the first molecular evidence of heterosis being driven by a single overdominant gene (Krieger et al., 2010). Recent work has also shown that hormone-mediated defense and stress response networks contribute to heterosis in Arabidopsis F₁ hybrids (Groszmann et al., 2015). There are also data showing that epigenetic variation may contribute to the molecular mechanisms of complex traits, including hybrid vigor (Cubas et al., 1999; Manning et al., 2006; Shindo et al., 2006; Ni et al., 2009; Alonso-Peral et al., 2010). For example, Ni et al. (2009) showed that epigenetic modification of a few regulatory genes [e.g. the circadian oscillators CIRCADIAN CLOCK ASSOCIATED1 (CCA1) and LATE ELONGATED HYPOCOTYL (LHY)], were associated with growth vigor in Arabidopsis thaliana hvbrids.

The genetic composition of hybrid offspring is the result of information inherited from both parents. Theoretically, no new genes are produced in the F_1 hybrid, so heterosis is likely caused by differences resulting from qualitative or quantitative modifications of gene expression. According to genomic and epigenetic principles, offspring exhibit advantages in growth, stress resistance and adaptability because of interactions between alleles of parental genomes that change the regulatory network of related genes. Microarray and high-throughput sequencing technologies have enabled detailed investigations of the molecular basis of heterosis at the whole genome level (Zhang *et al.*, 2008; Ni *et al.*, 2009; Wei *et al.*, 2009; Lai *et al.*, 2010; Song *et al.*, 2010; Groszmann *et al.*, 2011).

Small RNAs, including microRNA (miRNA) and small interfering RNA (siRNA), affect gene expression by posttranscriptional mechanisms and epigenetic modifications (Vaucheret, 2006; Lu et al., 2012). In intraspecific Arabidopsis hybrids, variation in miRNA expression causes nonadditive expression of target genes that may affect growth, vigor and adaptation (Ha et al., 2009). Both cis- and transregulation of miRNAs and other genes may underlie natural variation in biochemical and metabolic pathways that affect growth vigor and stress responses (Ng et al., 2011, 2012). Moreover, siRNAs, particularly those associated with transposable elements (TEs), may serve as a buffer against genomic shock and influence gene expression by directing DNA methylation. For example, recent studies suggest that epigenetic diversity between parental lines, such as differences in the levels of the 24-nucleotide (nt) siRNAs in two Arabidopsis ecotypes, create allelic variants that influenced the activity of a large number of metabolic and regulatory genes that may contribute to heterotic phenotypes (Groszmann et al., 2011; Shen et al., 2012). These observations indicate that siRNAs may have an effect on

genome-wide gene expression bias and could be important determinants in plant heterosis.

Brassica napus (oilseed rape) is a major edible oil-producing crop, and its F1 hybrids have been widely used in commercial production for many years. The flower development process in Brassica napus is similar to that of Arabidopsis. There have been a few studies on the potential mechanism of heterosis in B. napus (Grant and Beversdorf, 1985; Chen et al., 2008). Most of these investigations, as well as other studies with Arabidopsis, focused on the seedling stage. Investigations of the effects of gene expression on heterosis during flower development are much more limited. Previous studies from our group showed that the early stages of flower development are vital for determining the number of flowers and siliques produced per plant, and consequently the yield (Zhang et al., 2016b). However, stage-specific transcriptomic changes, which are crucial for understanding the gene regulatory networks underlying flower development and the phenomenon of heterosis in B. napus, have not been thoroughly investigated.

The genome sequence of *B. napus* is now available (Chalhoub et al., 2014), and together with the next-generation sequencing technology provides an unprecedented opportunity to address questions concerning changes in gene expression and epigenetic variations between the two parents and their F₁ hybrids that are associated with heterosis in allopolyploid *B. napus*. In this study we investigated the dynamic transcriptomes, small RNAomes and DNA methylomes of an F₁ hybrid and its two parents at three early flower developmental stages of B. napus. Our results reveal that similarities and differences in gene expression and epigenetic changes exist between the F₁ hybrid and its parents at the early flower developmental stage, and these data provide insights into the role of transcriptional and epigenetic changes in heterosis of B. napus.

RESULTS

Heterosis of F₁ hybrids in oilseed rape

The F_1 hybrid (F_1 hereafter) ZYZ108 derived from the cross of MB (female) and ZY50 (male) parents has significantly increased silique numbers and grain yield (Figure 1a,b). The total grain yield of F_1 was 60% higher than the midparent value (MPV). Increased yield was mainly the result of significantly increased silique numbers (about 50% higher than MPV), because no significant difference was observed for the weight per seed between F_1 and the two parents. The increased silique number of F_1 was not the result of an extended flowering period because the flowering time of F_1 was in between that of the two parents and all three genotypes were harvested at the same time. The effect of heterosis on silique number is directly correlated Figure 1. Heterosis of the F_1 hybrid in oilseed rape and in representative plants of the three developmental stages examined in this study.

(a) Comparison of silique numbers among the two parents and their F_1 hybrid.

(b) Comparison of grain yield among the two parents and their F_1 hybrid.

(c) Representative plants of the three developmental stages (SAM1, SAM2 and YFB) from which tissues were analyzed in this study.



with the increased number of flowers in F_1 (Zhang *et al.*, 2016b). Therefore, we used tissues from three key stages during early flower development to investigate the potential molecular mechanisms underlying heterosis in *B. napus*.

Differentially expressed genes in the two parents are enriched in the F_1 hybrid

Dynamic changes of the transcriptomes in F_1 and its two parents were investigated by analyzing transcripts in shoot apical meristems (SAMs) at the transition meristem (SAM1) and flower meristem initiation (SAM2) stages, and young flower buds (YFB) at the early flower developmental stage (Figure 1c).

We first analyzed the transcriptomes of each hybrid and compared them with their parents at the same developmental stage (in Figure 2). Genes in F₁ that had expression levels significantly different from the MPV [$P \le 0.05$; false discovery rate (FDR) \leq 0.05] were designated as hybrid-MPV differentially expressed genes (DEGs) (Data S1 in the Supporting Information). Such genes could potentially be responsible for generating the heterosis phenotype. We found that only a small portion of all expressed genes were hybrid-MPV DEGs at the three stages (494-1592 or 0.81-2.31%) (Figure 3a-c,g). This suggests that the majority of genes in F₁ show expression additivity, with only a small fraction of the expressed genes show non-additivity of expression ($P \le 0.05$; FDR ≤ 0.05). In F₁, more upregulated than downregulated genes were observed in SAM1 and SAM2 (2.9-fold and 3.8-fold, respectively) and these differences were both significant (χ^2 test, $P \leq 0.01$) (Figure 3a,b). In contrast, in YFB, the number of up- and downregulated genes were similar (Figure 3c).

Heterosis in the allopolyploid F_1 hybrids of B. napus 3

DEGs between the male (M) and female (F) lines (designated M–F DEGs) were further investigated. In SAM1, SAM2 and YFB, 3320, 2347 and 3838 genes (3.8–5.6% of the expressed genes), respectively, were differentially expressed between the two parents ($P \le 0.05$; FDR ≤ 0.05) (Figure 3d–f). Interestingly, we found that a high percentage of the M–F DEGs overlapped with the hybrid-MPV DEGs in all three developmental stages, suggesting that genes differing in the expression level between the parents of the hybrid were overrepresented (5–13 times) among the genes differentially expressed in the hybrids (Figure 3d–f).

Gene Ontology (GO) enrichment analyses of the M–F DEGs identified 484, 290 and 556 significantly enriched GO function terms (Fisher's exact test $P \le 0.05$; FDR ≤ 0.05) in SAM1, SAM2 and YFB, respectively (Figure S1, Data S2). We used REViGO to cluster the functional categories across these lists of significantly enriched GO terms. Based on functional clustered GO terms, the enriched pathways associated with the M–F DEGs were found to be related to several different functional categories, of which 'defense and stress response' and 'hormone response' were the two prominent GO terms altered between the parental lines.

Hormone response genes are significantly altered in the F_1 hybrid

The GO enrichment analyses of the hybrid-MPV DEGs identified 370, 324 and 304 significantly enriched GO



Figure 2. Work flow of bioinformatics analysis conducted to investigate the heterosis mechanism in *Brassica napus*.

Two main aspects (changes in gene expression level and epigenetic regulation) were targeted in this study and diverse sequence data (in red) were generated by high-throughput sequencing. [Colour figure can be viewed at wileyonlinelibrary.com]

function terms (Fisher's exact test $P \le 0.05$; FDR ≤ 0.05) in SAM1, SAM2 and YFB, respectively (Figure 3h). Using the same approach applied in the analysis of GO terms associated with the M–F DEGs, we found that 'defense and stress response', 'hormone biosynthetic process' and 'hormone response' were the three most prominent functional clusters altered in F₁ (Data S3), which accounted for >50% of the biological terms in SAM1 and SAM2. Interestingly, a large proportion of the GO terms overrepresented between the parental lines and between F₁ and MPV overlapped in each of the three stages (SAM1 56%, SAM2 38%, YFB 36%; Figure S2), consistent with the observation of enrichment of the M–F DEGs in the F₁ hybrid. These results suggest that the variance of gene expression levels between the two parental lines contributes to heterosis in *B. napus*.

Biological processes related to defense responses were among the most prevalent GO terms related to DEGs in the hybrid_ENREF_9. A large proportion of terms were involved in defense and stress response in SAM1 (100/253, 27%), SAM2 (128/246, 40%) and YFB (54/304, 17.7%) (Data S4), with all three stages having enrichments in many of the same basal processes. Among the downregulated defense response genes are several genes well known to be induced by pathogen attack, including *PATHOGENESIS RELATED GENE 1 (PR1, BnaC03g45470D)* and *LIPID TRANSFER PROTEIN 3 (LTP3, BnaA03g13140D)*. Many upregulated genes associated with defense response known to function as repressors of plant defense, including WRKY33 (BnaA04g22040D), WRKY40 (BnaA07g35260D), PENETRATION 3 (PEN3, BnaCnng64010D) and ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 4 (ERF4, BnaC03g39000D), were also recognized. The pattern of altered defense-related gene expression observed in F_1 suggests that the hybrid may have decreased capabilities of basal defense response. Given the antagonistic relationship between plant immunity and plant growth, the lower level of defense-related cellular activity could be a significant factor in allowing for greater growth of the hybrid.

The altered salicylic acid responses in the F_1 hybrid. Salicylic acid is derived mainly from chorismate through two key pathways, the isochorismate (IC) pathway and the phenylalanine (PHE) pathway. The IC pathway is the predominant route for both basal and pathogen-induced SA production (Dempsey et al., 2011). Both ISOCHORISMATE SYNTHASE 1 (ICS1) and ICS2 were downregulated in F1 (Figure 4a,c). In the later steps of the IC pathway, AVRPPHB SUSCEPTIBLE 3 (PBS3) is important for accumulation of SA (Nobuta et al., 2007) and was also downregulated in F₁ (Figure 4a,c). These results suggested that the IC pathway is downregulated in F₁. A small proportion of SA is produced through the alternative PHE pathway involving four PHENYLALANINE AMMONIA-LYASE (PAL) genes (Rohde et al., 2004). In contrast to the downregulation of the IC pathway, we found that PAL1 and PAL2 were significantly upregulated in F_1 (Figure 4b,c). Interestingly, Figure 3. Overview of the genes differentially expressed in the F_1 hybrid compared with MPV (mean of the two parental values).

(a-c) The number of genes differentially expressed in the F1 hybrid compared with the expected MPV (P \leq 0.05; false discovery rate \leq 0.05) in SAM1, SAM2 and YFB. (d-f) Genes that were differentially expressed in the two parents (represented by black bars) were more frequently associated with differentially expressed genes in the F₁ hybrid (χ^2 test, $P \leq 0.01$). (g) Breakdown of DEGs in the F₁ hybrid showing the number of differentially expressed genes that were unique in one sample or overlapping in two or three samples used in this study. The numbers in brackets refer to the number of genes with a UniProtKB number. (h) The number of significantly enriched Gene Ontology terms for the differentially expressed genes shown in (g). [Colour figure can be viewed at wileyonlinelibrary.com]

Heterosis in the allopolyploid F₁ hybrids of B. napus 5



C4H, a downstream gene of the PHE pathway, was also significantly upregulated in F_1 . These results suggested that more cinnamic acid was used for biosynthesis of 4-hydroxycinnamic acid instead of SA (Figure 4b,c). The increasing consumption of chorismate in the PHE pathway leading to 4-hydroxycinnamic acid limited the SA biosynthesis via the major IC pathway. Consistently, we found that the concentration of SA in F_1 was significantly lower than in either parent in SAM1, SAM2 and YFB (Figure 4f).

In addition to the core SA biosynthesis genes, upstream repressors of SA biosynthesis showed upregulation in F_1 (Figure 4d). To further investigate the role of SA in heterosis, we checked expression changes of the known SA-responsive genes (based on AtGeneExpress) in F_1 . A large spectrum of SA-inducible genes, ranging from pathogenesis-related genes to *WRKY* transcription factors, were differentially expressed and downregulated in F_1 (Figure 4e). In contrast, the SA-repressed *ETHYLENE RESPONSE FAC-TORS* (*ERF*) were upregulated in F_1 (Figure 4e). Increased *ERF4* represses jasmonic acid (JA) signaling and can promote vegetative growth (Lyons *et al.*, 2013), whereas

ERF11 downregulates pathogen defense-related genes (Dombrecht *et al.*, 2007).

The altered auxin and glucosinolate biosynthesis pathways in the F1 hybrid. Indole-3-acetic acid (IAA) is derived mainly from tryptophan through several key intermediates. While genes involved in the main IAA biosynthetic pathway (through indole-3-pyruvic acid, or IPA) were not altered in the hybrid relative to their MPVs, most genes involved in the indol-3-acetaldoxime (IAOX) pathway, which is specific to the Brassicaceae family, were downregulated in SAM1 and YFB (Figure 5a-c,d). In contrast, in SAM2, genes of the IAOX pathway were significantly upregulated, but AOP3, which converts indol-glucosinolates into hydroxyl-glucosinolate (hydroxyl-GS), was also significantly upregulated in F1, suggesting that the enhanced IAOX pathway in SAM2 was mainly used in biosynthesis of hydroxyl-GS rather than IAA (Figure 5b,d). In addition, AMIDASE 1 (AMI1), which is involved in the tryptophan \rightarrow IAOX \rightarrow indole-3-acetamide (IAM) \rightarrow IAA route, was downregulated in all three stages (Figure 5a-c,d).



Figure 4. Altered salicylic acid (SA) response in the F_1 hybrid.

Differential expression patterns of genes in (a) isochorismate (IC) and (b) phenylalanine (PHE) SA biosynthesis pathway in SAM1, SAM2 and YFB. (c) The SA biosynthesis pathway showing routes for production of bioactive SA. Gene names are colorcoded to represent their different changes in expression levels in the F1 hybrid; the same color codes are used for the biosynthetic products to reflect their expected changes. (d) Differential expression of the repressors of SA biosynthesis. (e) Differential expression of downstream SA-regulated genes in F1 hybrid (first three columns); the next two columns show the change in expression levels of the given gene in response to altered SA concentrations. (f) Comparison of the SA concentration in SAM1, SAM2 and YFB of the two parents and their F₁ hybrid.

Downregulation of the auxin-inducible marker genes also reflected the state of the decreased auxin level in F_1 (Figure 5e). These results suggest that the genes involved in the IAA biosynthesis pathway leading to accumulation of IAA were downregulated in all three stages. This is consistent with the observation that F_1 had a lower level of IAA than either parent in SAM1, SAM2 and YFB (Figure 5f).

An important role of high-parental expression level dominance in heterosis

Allopolyploids have been found to exhibit parental expression level dominance (ELD), a phenomenon in which some genes that are differentially expressed between the parental lines exhibit expression levels in F_1 that are statistically similar to that in one of the parents (Rapp, 2009; Yoo *et al.*, 2013). To dissect the expression pattern of the expressed genes (in Figure 2), we classified genes into 12 male-hybrid-female (M-H-F) expression patterns (Figure 6a) as described in Yoo *et al.* (2013). We further designated genes with expression levels in F_1 that are statistically similar to those in the male and female parent as ELD-M and ELD-F genes, respectively. A large number of ELD genes were identified in the three different developmental stages. This is in contrast to the limited number of genes found in these samples that exhibit additivity and transgressive regulation. This finding indicates that ELD plays an important role in allopolyploid hybrids. Among the genes exhibiting ELD some were expressed as highly as in the parent showing

Figure 5. Altered auxin response in the F₁ hybrid. (a-c) Auxin biosynthesis pathway showing routes for production of bioactive auxin (IAA). Gene names are color-coded to represent their different changes in expression levels in the F₁ hybrid; the same color codes are used for the biosynthetic products to reflect their expected changes. (d) Differential expression patterns of genes in auxin biosynthesis. (e) Differential expression of auxin-regulated genes in the F₁ hybrid. (f) Comparison of the IAA concentration in SAM1, SAM2 and YFB of the two parents and their F₁ hybrid.



the highest expression level (high-parental ELDs; types III and V in Figure 6a), while others had expression as low as that of the parent showing the lowest expression level (lowparental ELDs; types IV and VI in Figure 6a). In the three stages, more than 87% (1593/1826), 89% (1068/1191) and 73% (1443/1964) of ELD genes were high-parental ELDs in SAM1, SAM2 and YFB, respectively (Figure 6b). The number of high-parental ELD genes was significantly higher than that of the low-parental ELD genes in F₁ ($P \le 0.01$, pairwise *t*-test). Interestingly, a significant number of the hybrid-MPV DEGs overlapped with high-parental ELD genes in each developmental stage (of the total DEGs, 30.5% in SAM1, 21.1% in SAM2 and 9.6% in YFB). We also found 321, 104 and 153 ELD genes overlapping with the non-additively expressed genes in SAM1, SAM2 and YFB, respectively (Figure S3). These observations suggest that the high-parental ELD genes potentially contribute to heterosis in the hybrid, partly acting together with the non-additive mechanism. To further support this hypothesis, we found that the majority of the genes for IAA biosynthesis pathway were non-additively expressed, but *PAL1* (*BnaA04g21240D*) was a high-parental ELD gene (Figure 4b).

We next examined the possible functions of the genes exhibiting high-parental ELD in F_1 . A large proportion of the significantly enriched GO terms were found to be overlapping among the three stages [69.5% (221/328) for

Heterosis in the allopolyploid F_1 hybrids of B. napus 7



Figure 6. Parental expression level dominance (ELD) genes in the hybrid oilseed rape.

Genes with an expression level in the F₁ hybrid similar to the female parent are designated as ELD-F; genes with an expression level in the F₁ hybrid similar to the male parent are designated ELD-M.

(a) Expression patterns of twelve types of differentially expressed genes (DEGs). M, male parent; H, hybrid; F, female parent.

(b) Number of genes in each of the 12 types of DEGs in SAM1, SAM2 and YFB.

(c) Venn diagram showing the number of the significantly enriched Gene Ontology (GO) terms based on the genes with a high parental ELD (III and V in A) in each tissue.

(d) Prominent GO terms enriched in the genes with a high parental ELD in each tissue. [Colour figure can be viewed at wileyonlinelibrary.com]

SAM1, 73.2% (221/302) for SAM2 and 46.6% (221/474) for YFB], indicating that the high-parental ELD genes may be involved in similar biological processes in the three different developmental stages (Figure 6c). We found that the high-parental ELD genes (types III + V) were significantly enriched for GO terms related to response to defense, stress and hormone stimulus (Figure 6d). These results are consistent with the enrichment results of non-additively expressed genes, implying that ELD genes contributed to

the heterosis together with the non-additively expressed genes, and that both non-additivity and parental ELD, especially high-parental ELD, play important roles in heterosis in *B. napus*.

Significant trans-regulation of allelic expression in F₁ hybrid

It is not clear how two alleles of the same gene from the parental lines are regulated in F_1 . To distinguish the

contribution of each parental allele to the total gene expression level in F₁ (in Figure 2), we used parental allelic single nucleotide polymorphisms (SNPs) to determine the allele origin and its expression level in F₁. To this end, we first sequenced the genomes of the two parents to a 15× coverage and then identified allelic SNPs located in exons of all annotated genes between the two parents by separately mapping their genome sequencing data and RNA sequencing (RNA-seq) data to the *B. napus* reference genome. To ensure the accuracy of parental allele discrimination, only transcripts with allelic SNPs detected in both the transcriptome and re-sequenced genome were included in the read assignment and expression comparison.

To define *cis*- and/or *trans*-regulation of a given allele in F1 we used the principles documented by Shi et al. (2012) (Figure 7a). We used P-value A (P_A , χ^2 test) to distinguish the differences between the two parental alleles expressed in the hybrid and their relative expression patterns in the two parents to determine trans-regulation. The alleles are believed to be regulated by a trans-effect if the differences in their relative expression between the two parental alleles in F₁ were significantly different from their differences in expression in the two parents $(P_{\rm A} \leq 0.05)$. For the *cis*-effect, we used *P*-value B $(P_{\rm B}, \chi^2)$ test) to test whether the two parental alleles are expressed equally in the hybrid. If their expression levels were significantly different in the hybrid ($P_{\rm B} \leq 0.05$), the two parental alleles were suggested to be influenced by the ciseffect. By combining the P_A and P_B test results, four types of regulatory effects could be defined: trans-effect only, cis-effect only, both cis- and trans-effects, and no cis- and trans-effects (Figure 7a). After excluding genes with low expression levels [i.e. reads per kilobase of transcript per million mapped reads (or RPKM) ≤ 2] or with high sequence similarity between the two parental alleles (i.e. low SNP coverage), we identified approximately 4400 to 5400 genes showing significant difference in allelic expression in F_1 in the three developmental stages. In YFB, 147 (3.1%), 319 (6.7%) and 357 (7.5%) genes were found to be regulated by the cis-effect only, trans-effect only, and both cis- and trans-effects (cis-trans), respectively (Figure 7b). The remaining 3959 (82.8%) genes in YFB showed no allelic expression difference between the parents or in F₁ (i.e. no *cis-trans*). A similar proportional distribution of the four types of regulatory effects was observed in the two SAM samples.

Among the ELD genes and the hybrid-MPV DEGs, we identified 558 and 299 genes, respectively, showing significant allelic expression difference in F_1 in the three samples. Among the ELD genes, the percentage of *trans*-effects-only genes in SAM1 (46.4%), SAM2 (40.0%) and YFB (35.0%) was much higher than that observed among the total genes (6.9% for SAM1, 6.4% for SAM2 and 6.7% for YFB)



Figure 7. Distribution of genes with *cis*- and/or *trans*-effects in SAM1, SAM2 and YFB.

(b) Number and percentage of the genes without a *cis*- and *trans*-effects, with a *cis*- or *trans*-effects only, or with both *cis*- and *trans*-effects amongst the expression level dominance (ELD) genes, differentially expressed genes (DEGs) between the F₁ hybrid and mid-parent value, and all expressed genes in each tissue. [Colour figure can be viewed at wileyonlinelibrary.com]

(Figure 7b). These results underscore the importance of *trans*-regulation for the ELD genes in the *B. napus* F_1 hybrid. For the hybrid-MPV DEGs, although *trans*-effectsonly genes played an important role, *cis*-*trans*-effects seemed to make a similar contribution as *trans*-effects-only to the behavior of these genes (Figure 7b). For both the ELD genes and hybrid-MPV DEGs, *cis*-effects-only was the least detected regulatory effect (Figure 7b). Taken together, these results suggest that *trans*-effects-only play a more important role than *cis*-effects-only in shaping the expression profiles of the genes that are differentially expressed in F_1 .

⁽a) Flow chart showing quantification of *cis*- and *trans*-effects and classification of the transcripts with *cis*- and/or *trans*-effects. F, gene expression level in female parent; M, gene expression level in male parent; HF, female parental allelic expression level in hybrid; HM, male parental allelic expression level in hybrid.

Decreased divergence of expression of the homoeologous gene pairs in the F_1 hybrid

To analyze the expression levels of homoeologous genes in the A_n and C_n subgenomes of F_1 (in Figure 2), we first identified homoeologs expressed in the two parents and their F_1 hybrid in the three developmental stages. A total of 17 711, 16 969 and 18 308 such genes were identified in SAM1, SAM2 and YFB, respectively. In general, more C_n homoeologs were expressed at a higher level than their corresponding A_n homoeologs in all samples and genotypes (i.e. $C_n > A_n$; Figure 8a). We then calculated the fold changes of the expression level between homoeologous gene pairs in F_1 and the two parents. Compared with its parents, F_1 had fewer differentially expressed gene pairs. For example, the number of homoeologous gene pairs with a 16-fold difference in expression [i.e. $log_2(A_n/C_n) \geq 4.0$ or ≤ -4.0] in SAM1 (321), SAM2 (256) and YFB (324) of F_1 was significantly less than that of either parent ($P \leq 0.01$, Student's t-test) (Figure 8b).

We also considered the magnitude of the expression changes in homoeologous gene pairs between the two parents (MPV) and F₁. Compared with the MPV, F₁ showed less difference in expression [i.e. log₂(F₁/MPV) \leq 0)] for most homoeologous gene pairs (Figure 8c), suggesting a



Figure 8. The expression patterns of homoeologous genes of the A_n and C_n subgenomes in *Brassica napus*.

(a) Percentage of the homoeologous genes with a higher or lower expression level in the C_n subgenome than in the A_n subgenome.

(b) Number of homoeologous gene pairs with a 16-fold (i.e. $log_2(A_n/C_n) \ge 4.0$ or ≤ -4.0) expression difference in the two subgenomes of *B. napus*. MPV, mid-parent value.

(c) Distribution of the homoeologous gene pairs based on their expression changes between the F_1 hybrid and MPV [log₂(F_1 /MPV)] and their expression difference between the A_n and C_n homoeologs [log₂(A_n/C_n)] in each tissue. [Colour figure can be viewed at wileyonlinelibrary.com]

reduction of divergence of homoeolog expression as a result of hybridization. In addition, the homoeologous gene pairs which were more highly expressed in the C_n subgenome than in the A_n subgenome [i.e. $log_2(A_n/C_n \leq 0)$] had larger $log_2(F_1/MPV)$ values (i.e. they deviated more from the MPV) than those that expressed more highly in the A_n subgenome than in the C_n subgenome [i.e. $log_2(C_n/A_n \leq 0)$] (Figure 8c). This was because the C_n subgenome had more highly expressed homoeologs (Figure 8a). Taken together, these data showed a clear reduction in the expression divergence of homoeologs in F_1 following hybridization.

The F₁ hybrid has more non-additively repressed miRNAs

To characterize the potential role of miRNA and siRNA in heterosis of *B. napus*, we investigated the small RNA (sRNA) expression profiles of the two parents and their hybrid at three developmental stages (SAM1, SAM2 and YFB). Consistent with previous reports (Shen *et al.*, 2014, 2015), our data showed that the 24-nt class of sRNAs was the most abundant (>40% of the total reads) (Figure S4). Regarding miRNAs, in addition to the known *B. napus* miRNAs deposited in the miRbase database (release 21) and those reported in recent studies (Shen *et al.*, 2014, 2015), we identified 851 novel miRNAs from 393 families.

A total of 68 conserved miRNAs belonging to 48 families identified in the three stages were found to be non-additively expressed ($P \le 0.05$, FDR ≤ 0.05). Of the 68 miRNAs, 25, 12 and 12 were non-additively repressed in SAM1, SAM2 and YFB, respectively, while 16, 4 and 20 were nonadditively activated in SAM1, SAM2 and YFB, respectively. This set of miRNAs included some of the most functionally well-characterized miRNAs (e.g. miR156, miR159, miR165/ miR166, miR319, miR393 and miR2111) conserved between monocots and eudicots (Figure 9a). The target genes of these non-additively expressed miRNAs were predicted from the gene models in the *B. napus* genome annotation (Chalhoub et al., 2014) via psRobot analysis (Wu et al., 2012). A total of 117 target transcripts were predicted for 37 of the 68 conserved miRNAs (Data S5). We further investigated the correlations between miRNAs and targets by the Pearson correlation coefficient method (Rodgers and Nicewander, 1988). There were significant negative correlations between the expression levels of the conserved miRNAs and their corresponding target genes (r = -0.83). According to the annotation of the plant transcription factor database PlantTFDB (Jin et al., 2014), 103 targets of the non-additively expressed miRNA encoded 86 transcription factors from 11 families. Compared with all annotated oilseed rape genes, these target genes were significantly (P < 0.05) enriched for 17 biological process and 7 molecular function GO terms, including the IAA-mediated signaling pathway, response to stress and sequence-specific DNAbinding transcription factor activity.

In addition to conserved miRNAs, a total of 73 (36 families), 89 (32 families) and 85 (32 families) novel miRNAs were identified in SAM1, SAM2 and YFB, respectively (Data S6). Based on their expression levels in F_1 and parental lines, we found that F₁ had more downregulated novel miRNAs than upregulated ones in SAM1 (26/36, 72.22%), SAM2 (28/41, 68.29%) and YFB (23/32, 71.88%). To better understand the possible roles of the repressed miRNAs in heterosis, we performed GO enrichment analysis for their targets. We found that the biological process GO term 'protein amino acid phosphorylation' was enriched in both SAM1 (P = 0.0036) and SAM2 (P = 0.0176) (Figure 9b). We then used our previously generated degradome sequencing data (Shen et al., 2015) to validate the interaction between the novel miRNAs and their targets. From this analysis, we were able to confirm miRNA-mediated cleavage for 22 targets, including several genes encoding phosphorvlation-related protein kinases, such as BnaA09g39530D (Figure 9c). In addition, 128 (12%), 64 (13%) and 107 (7%) non-additively expressed genes, and 189 (12%), 85 (8%) and 100 (7%) ELD genes in SAM1, SAM2 and YFB, respectively, were found to be involved in the phosphorylation pathways (Data S7). These results suggest that the protein amino acid phosphorylation pathways and miRNA-mediated regulation of genes of the pathways might be a potential regulatory process contributing to heterosis.

Non-additive expression of the 24-nt siRNA clusters in the F_1 hybrid

To characterize the effects of 24-nt siRNAs on genome stability and gene expression, we surveyed 24-nt siRNA density and expression level in F_1 . In SAM1, SAM2 and YFB, about 200 000–250 000 24-nt siRNA clusters were identified. About 200 000 24-nt siRNA clusters were expressed in both parents and F_1 in each tissue, covering 26.50–39.86% of the *B. napus* genome. The length of the majority of 24-nt siRNA clusters ranged between 100 and 1000 nt with 250–300-nt siRNA clusters dominating all three stages (Figure S5).

Based on comparison of expression levels, we found that 41–49% of the 24-nt siRNA clusters were differentially expressed between F₁ and its two parents in three stages. Interestingly, in F₁ the number of 24-nt siRNA clusters with an expression level higher than the MPV (64.7% in SAM1, 68.8% in SAM2 and 63.5% in YFB) was significantly higher than the number of 24-nt siRNA clusters with an expression level lower than the MPV (35.3% in SAM1, 31.2% in SAM2 and 36.5% in YFB) in all three developmental stages ($P \le 0.001$, pairwise Student's *t*-test) (Figure 10a). Consistent with this observation, the majority of 24-nt siRNA clusters had a $log_2(F_1/MPV)$ value of 0.2 (Figure 10b), significantly deviating from the null expectation [i.e. $log_2(F_1/MPV) = 0$]. We conducted the same analysis for the

© 2017 The Authors The Plant Journal © 2017 John Wiley & Sons Ltd, *The Plant Journal*, (2017), doi: 10.1111/tpj.13605





Figure 9. Changes of miRNAs and their targets in the F_1 hybrid.

(a) Expression changes of the differentially expressed conserved miRNAs in the F_1 hybrid compared with the mid-parent value (MPV) in the three tissues. (b) Significantly enriched Gene Ontology terms (colored in yellow) of the differentially expressed novel miRNAs targets in the F_1 hybrid.

(c) Examples of candidate targets confirmed to be regulated by novel miRNAs based on degradome analysis.

20–22-nt siRNA clusters and did not observe a similar expression pattern to that found for the 24-nt siRNA clusters. This result indicates that the biased, or non-additive, siRNA expression pattern in F_1 is unique to the 24-nt siRNA clusters.

To further investigate the expression level of the 24-nt siRNA clusters between F_1 and the two parents (MPV), we plotted the number of the 24-nt siRNA clusters identified in F_1 based on their expression level (Figure 10c). Although the peak appeared around the MPV, the distribution of the



Figure 10. Expression patterns of the 24-nucleotide (nt) siRNA clusters in the F₁ hybrid.

(a) Number of the 24-nt siRNA clusters that have a higher or lower expression level in the F_1 hybrid compared with the mid-parent value (MPV). (b) Distribution of the number of the 24-nt siRNA clusters based on their expression level changes between the F_1 hybrid and MPV in each tissue. (c) Distribution of the number of the 24-nt siRNA clusters based on their expression levels in the F_1 hybrid compared with those in the two parents and MPV.

(d) Number of transposable elements (Tes) that were up- or downregulated in the F_1 hybrid compared with MPV (fold change > 1.5). (e) Distribution of the number of TEs based on their expression level changes in the F_1 hybrid and MPV in each tissue. (f) Number of cytosine residues that have a higher or lower methylation level in the F_1 hybrid compared with MPV and the distribution of the methylated cytosine numbers based on the methylation level fold change in the F_1 hybrid and MPV in YFB. [Colour figure can be viewed at wileyonlinelibrary.com]

24-nt siRNA cluster number was negatively skewed in all three stages (skewness > 0), suggesting more 24-nt siRNA clusters were expressed more highly than the MPV in F₁.

The 24-nt siRNA clusters were found in a wide range of locations on all chromosomes, and their numbers in the following three genomic locations were compared: (i) TEs, (ii) gene-coding regions, and (iii) the \pm 1-kb flanking

regions of a gene. In both parents and the F₁ hybrid, more than 80% of the 24-nt siRNA clusters co-localized with TEs, although more 24-nt siRNA clusters were found in the \pm 1-kb flanking regions than in the gene-coding regions in all three developmental stages (Figure S6).

To investigate whether the TE-derived non-additively expressed 24-nt siRNA clusters influence TE transcription,

we compared the expression patterns of the TEs, from which 24-nt siRNA clusters were generated, in both parents and F₁. Of the approximately 1100–1600 TEs identified as being differentially expressed between MPV and F₁ in SAM1, SAM2 and YFB, 60–72% of them had a lower expression level in F₁ compared with the MPV, which was significantly higher than the number of the TEs having an expression level higher than the MPV (Student's *t*-test, $P \le 0.01$; Figure 10d). Consistent with this observation, the majority of TEs that generate 24-nt siRNAs had a log₂(F₁/MPV) value of -0.4 to -0.2 in SAM1 and SAM2, respectively, which deviated from the expected log₂(F₁/MPV) value of 0 (Figure 10e). Taken together, these results suggest that the activity of the TEs associated with 24-nt siRNA clusters was repressed in F₁.

Distribution analysis of the 24-nt siRNA clusters in the non-additively expressed genes found that most of them were located at the 1-kb upstream or downstream regions. In SAM1, SAM2 and YFB we found 745, 251 and 1086 non-additively expressed genes, respectively, containing siRNA clusters in their coding regions or 1-kb flanking regions. Among them, 438, 195 and 614 genes had siRNA clusters in the 1-kb upstream regions, 426, 192 and 615 genes had siRNA clusters in the 1-kb downstream regions, and 239, 130 and 412 genes had siRNA clusters in coding regions (Figure S7, Data S8). In addition 398 (38%), 208 (42%) and 701 (44%) non-additively expressed genes in SAM1, SAM2 and YFB, respectively, showed differential expression of the 24-nt siRNA clusters.

The F1 hybrid has an increased level of CG methylation

Genome wide, the average CG (CpG), CHG and CHH methylation levels were 58.2–61.1%, 25.7–27.4% and 10.4–12.6% in the two parents and their F₁ hybrid, respectively, in YFB. For all three methylation sequence contexts, F₁ always had a slightly higher level than the two parents (Figure 11a). Among the methylated cytosines, the majority were CG sites, followed by CHG and CHH sites in all three genomes. Compared with the two parents, F₁ had a relatively higher proportion of CG sites (77.3%) and a similar proportion of CHG and CHH sites as the female parent (Figure 11b).

According to the frequency distribution of the single base methylation level, in F₁ the number of unmethylated cytosines was decreased and the number of highly methylated cytosines was not increased (Figure 11c), implying that the increased methylation level in F₁ is mainly caused by the reduction in the number of unmethylated cytosines. In F₁, the majority (69%) of the methylated CG sites were inherited from both parents, but a significant number of the methylated CG sites were derived from only one of the parents (6%) or *de novo* methylation (5%) after hybridization (Figure 11d). The latter could explain the increased level of CG methylation in F₁.

Analysis of the location of the CG sites methylated in F1 but unmethylated in both parents revealed that 78% of these CG sites were associated with genes. Of these sites 23% was found in either the 1-kb upstream or 1-kb downstream regions of the genes, and 32% was found in the gene-coding regions. Only 3% were associated with TEs (Figure 11e). Thus, most of the CG sites that were methylated in F₁ but unmethylated in both parents were generather than TE-related. In addition, we found that the 24-nt siRNA clusters had a much higher frequency of methylated cytosines compared with the whole genome. For example, about 98% of cytosines in the 24-nt siRNA clusters identified in the two parents and F₁ were methylated. In addition, we found that the number of methylated CG sites with a higher methylation level in F_1 was ~23% more than the number of methylated CG sites with a lower methylation level in F_1 (insert in Figure 10f), consistent with the observation for the number of the 24-nt siRNA clusters (Figure 11a). This result indicates that the number of 24-nt siRNA clusters is directly related to the level of CG methylation and the transcriptional activity of TEs (Figure 10d).

Furthermore, among the total of 1592 non-additively expressed genes in F_1 , 1168 (73%), 1169 (73%) and 1043 (66%) showed a different level of DNA methylation in 1-kb upstream, 1-kb downstream and coding regions, respectively, implying an impact of the changed DNA methylation level on gene expression in the hybrid. One would assume that downregulation of the genes should be directly associated with increase in their DNA methylation levels or the expression levels of 24-nt siRNA clusters. However, our results did not support this assumption, which is consistent with what has been discovered in the Arabidopsis study (Shen *et al.*, 2012).

Taken together, CG methylation in both TEs and genecoding regions contributed to the increased methylation level in F_1 . In other words, genome-wide remodeling of DNA methylation may play a key role in heterosis by both influencing the activity of TEs and changing the expression level of genes.

DISCUSSION

Changes in gene expression level related to phytohormone response in the F₁ hybrid

Based on comparison of the transcriptome between the *B. napus* F_1 hybrid and its two parents in SAM and YFB, we found substantial transcriptional reprogramming following hybridization, with 0.8–2.3% of genes showing altered expression levels in the hybrid. The majority of these genes were upregulated in F_1 . Analysis of the biological process GO terms associated with DEGs in F_1 revealed a preponderance of defense- and hormone-related processes in all three samples. Interestingly, a recent study in Arabidopsis (Groszmann *et al.*, 2015) also reported that a



Figure 11. Methylation patterns in the two parents and their F_1 hybrid.

(a) The average methylation level of CpG, CHG and CHH.

(b) Percentage of the methylated cytosines in CpG, CHG and CHH.

(c) Histogram showing the distribution frequency of the cytosine methylation level.

(d) Distribution (by per cent of total) of the eight methylation patterns.

(e) Distribution of genomic features with methylated cytosines in the F1 hybrid but not in either of the two parents. [Colour figure can be viewed at wileyonlinelibrary.com]

large proportion of GO terms of DEGs between F_1 and its parents were involved in defense and stress response. Among the defense- and hormone-related DEGs uncovered in the hybrid were those encoding enzymes involved in biosynthesis of IAA and SA together with downstream genes regulated by the two hormones. These altered gene activities are of potential importance in generating the heterotic phenotype considering that IAA and SA are key regulators of plant defense responses and are also major controllers of plant structure and growth (Busov et al., 2008; Kazan and Manners, 2009). We found that the SA level was significantly downregulated in F₁. Given the antagonistic relationship between plant immunity and plant growth, the lower level of defense-related cellular activity could be a significant factor in allowing for greater growth of the hybrid. Altered IAA responses have also been observed in rice hybrids (Zhang et al., 2014). Although changes in IAA response are common in hybrids, the molecular triggers appear to be variable amongst different hybrid systems. miRNAs and other small RNAs were linked to changes to the IAA network in hybrid rice (Zhang et al., 2014), whereas in the maize hybrids, allele variants of the IAA-regulated transcription factor ZmAGROS1 seemed to be the main contributors (Guo et al., 2014). We found that in the B. napus hybrids, the decreased IAA response mainly involves changes of IAA biosynthesis through the IAOX pathway, which is specific to Brassicaceae. The decreased IAA concentration in SAM and young flower bud potentially contributed to the heterosis through enhancement of branching due to a reduced apical dominance (Cline, 1991).

We also compared global patterns of gene expression changes between F1 and its parents. All possible modes of gene action, including additivity, high- and low-parent dominance, underdominance and overdominance, were observed in the B. napus hybrids. The majority of the expressed genes (>95%) exhibited expression patterns that are not statistically distinguishable from additivity, and 70-80% of the non-additive genes exhibited high-parent dominance, which is significantly higher than the proportion (15-20%) of the genes that exhibited low-parent dominance ($P \le 0.01$ based on the χ^2 test). Interestingly, a number of genes with high-parent dominance were related to defense and hormone response in the *B. napus* hybrid. coincident with the enriched functional processes found in the DEG analysis. These findings are consistent with the hypothesis that while multiple molecular mechanisms contribute to heterosis, high-parent dominance may be particularly important.

miRNAs are important epigenetic regulators in the F_1 hybrid

In Arabidopsis hybrids, variation in miRNA expression causes non-additive expression of target genes that may affect growth vigor (Chen, 2007, 2013; Ha and Chen, 2009; Ng *et al.*, 2012). A large number of miRNAs, including many functionally well-characterized miRNAs, were non-additively expressed in hybrids of *B. napus*. The observation of a substantial number of miRNAs displaying non-additive expression during early flower developmental stages indicates that those miRNAs may contribute to heterosis of *B. napus*. Upregulation of miR156 and

miR159/miR319 in particular could be beneficial in the plant's adaptation to stress by modulating plant morphological characteristics via transcriptional regulation of the transcription factors SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) and MYB/TCP (Palatnik et al., 2007; Wang et al., 2008). SPL genes are involved in floral transition and regulation of flowering (Manning et al., 2006). miR156 targets the SPL gene family to control the transition from the vegetative phase to the floral phase in Arabidopsis, rice and maize (Gang and Poethig, 2006). In rice, OsSPL14 promotes panicle branching and higher grain productivity (Miura et al., 2010). Notably, three B. napus homologs of OsSPL14 (i.e. BnaA06q14810D, BnaC06q37430D and BnaC05g16270D) appeared to be upregulated in F_1 in both SAM1 and SAM2. Therefore, it is plausible that the nonadditively regulated miR156-SPL pathway is associated with the heterosis of flower number and grain productivity of the B. napus hybrids. In Arabidopsis, miR166 is involved in SAM and lateral organ formation by regulating its target gene AtHD-ZIP (Williams et al., 2005). A B. napus homolog of AtHD-ZIP (BnaC08g35630D) was found to be non-additively expressed in F₁, suggesting that the miR166-HD-ZIP pathway might be conserved and contribute to the heterosis phenomenon through its action in the formation of the SAM.

TCP transcription factors direct the developmental processes determining leaf size, leaf form and flower symmetry (Danisman et al., 2012). In a previous study, overexpression of miR319, which specifically downrequlates TCP mRNAs, was found to result in uneven leaf shape and delayed flowering (Palatnik et al., 2007). Overexpression of miR159 has been reported to reduce accumulation of MYB mRNA and result in male sterility, whereas plants that overexpress miR159-resistant MYB33 have upwardly curled leaves, reduced stature and shortened petioles (Alonso-Peral et al., 2010). In addition, miR164, miR166, miR167 and miR390 have been shown to target ARF genes and to be regulated by IAA, suggesting that these miRNAs are also involved in the IAA response pathways (Mallory and Bartel, 2005). These miRNAs are highly conserved in plant species and their targets in B. napus have been validated (Shen et al., 2014, 2015). Based on their functionalities in other plant species, the non-additively expressed miRNAs identified in F1 are also most likely key regulators that are critical for growth vigor and adaptation of the *B. napus* F_1 hybrid.

Genetic buffering of siRNAs represses TE activity and changes gene expression levels through an enhanced level of methylation in the F_1 hybrid

siRNA-mediated epigenetic mechanisms have been shown to be important for the maintenance of genome stability of interspecific hybrids and allopolyploids (Ha and Chen, 2009; Lu *et al.*, 2012). In this study, we investigated the relationship between the expression level of 24-nt siRNA clusters and DNA methylation levels, and their changes in F1 relative to its parents. Although it is still too early to link the differences found between F₁ and its parents to specific heterotic traits, some interesting alterations in F1 were strongly correlated with the observed heterosis. We observed upregulation of the 24-nt siRNA clusters and an increased DNA methylation level in all three sequence contexts, particularly at the CG sites, in the *B. napus* F₁ hybrid (Figure 10a,b). We provided evidence to support that the non-additively expressed 24-nt siRNA clusters were linked to the reduced activity of TEs in the hybrid (Figure 10d). The observation of upregulated 24-nt siRNAs in the B. napus hybrid is different from what has been documented in other plant species, including Arabidopsis, rice and maize, in which reductions of 24-nt siRNAs following hybridization were observed (Ha and Chen, 2009; He et al., 2010; Groszmann et al., 2011; Barber et al., 2012). More studies are required to determine whether this is unique to the tissues (SAMs of the flower initiation stages and YFB) used in this study or to *B. napus*. Furthermore, it is interesting to note that most of the increased methylation in the B. na*pus* hybrid came from cytosines that were not methylated in one or both parents (Figure 11d), indicating the activities of trans chromosomal methylation (TCM) (Greaves et al., 2014) and *de novo* DNA methylation. We found that up to three-quarters of the methylated cytosines generated by the TCM and de novo DNA methylation mechanisms were associated with protein-coding genes (including 1 kb up-/ down-stream of genes and gene-encoding regions). In fact, the expression level of a gene is usually not directly correlated with its methylation level. For example, in Arabidopsis, two-thirds of the genes (>20 000) were methylated within the immediate context, including 1 kb upstream and downstream of the annotated gene model (Shen et al., 2012); however, in the *met1* and *ddc* mutants, although nearly all CG and non-CG methylation, respectively, was eliminated (Cao and Jacobsen, 2002; Tarig et al., 2003; Zhang *et al.*, 2006; Lister *et al.*, 2008), only hundreds of genes were overexpressed in these two mutants (Zhang *et al.*, 2006). In addition, the RdDM pathway or 24-nt siR-NAs seems not to be required for hybrid vigor (Barber *et al.*, 2012; Kawanabe *et al.*, 2016; Zhang *et al.*, 2016a). Therefore, the relationship between gene-associated methylation and the heterosis phenomenon in the *B. napus* F_1 hybrid should be further investigated.

In summary, we examined the differences between an F_1 hybrid and its parents at various developmental stages during floral development using a variety of -omic level technologies, and provided detailed information on the role of critical alterations in gene expression and epigenetic mechanisms contributing to heterosis in allopolyploid *B. napus.* Based on the results, we propose that the ELD genes and the non-additively expressed genes together contribute to heterosis mainly through the 'defense and stress response' and 'hormone response' pathways, which are regulated by both miRNAs and DNA methylation caused by changes in 24-nt siRNAs. DNA methylation could also play a role in heterosis by reducing the activity of transportable elements (Figure 12).

EXPERIMENTAL PROCEDURES

Plant materials

Brassica napus F_1 hybrid ZYZ108 (a major commercial variety used in the lower Yangtze river regions of China) derived by crossing the varieties of ZY50 and MB, a male-sterile line used as the male parent (ZY50) and a restorer line used as the female parent (MB), were used in this study. SAMs at the transition meristem stage (the stage where morphological changes are not visible but induction is taking place; designated sample SAM1) and flower meristem initiation stage (the stage where the first flower primordia are arising on the flanks of the inflorescence meristem; designated sample SAM2) and young flower buds at the early flower developmental stage (designated sample YFB) were collected and used in all experiments, unless otherwise indicated. Each genomic DNA and total RNA sample was extracted from a pool of tissue collected from around 20 to 30 plants.



Figure 12. Hypothetic model of heterosis in *Brassica napus* (ELD, expression level dominance; nt, nucleotide). [Colour figure can be viewed at wileyonlinelibrary. com]

© 2017 The Authors

The Plant Journal © 2017 John Wiley & Sons Ltd, The Plant Journal, (2017), doi: 10.1111/tpj.13605

Library generation for high-throughput sequencing and bioinformatics analyses

Methods S1 gives details about whole-genome DNA re-sequencing, mRNA and sRNA sequencing, methylC sequencing and bioinformatics analyses (Figures S8 and S9, Data S9–S12).

ACCESSION NUMBERS

Sequence data from this article can be found in the Gen-Bank/EMBL databases with BioProject ID PRJNA338132.

ACKNOWLEDGEMENTS

This work was supported by the National Basic Research Program of China (2015CB150200) and Jiangsu Collaborative Innovation Center for Modern Crop Production.

AUTHOR CONTRIBUTIONS

LF designed the research and wrote the article. YS performed the experiment, conducted bioinformatics analyses and wrote the article. SH performed the field experiment. SS, ES, CYY and DC discussed the results. MPT and QHZ discussed the results and wrote the article.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Venn diagram showing the number of the significantly enriched Gene Ontology terms based on the genes differentially expressed between the parental lines.

Figure S2. Overlapping analysis of the Gene Ontology terms between hybrid-MPV and two parental lines.

Figure S3. Overlapping analysis between the ELD genes and the non-additively expressed genes in SAM1, SAM2, and YFB.

Figure S4. Length distribution of small RNAs in each sample used in this study.

Figure S5. Length distribution of the 24-nucleotide siRNA clusters in the three samples.

Figure S6. Distribution of the 24-nucleotide siRNA clusters in the 1 kb upstream, 1 kb downstream, transcribed region and transposable elements in the three tissues.

Figure S7. Venn diagram showing the number of the non-additively expressed genes that had siRNA clusters in coding regions or 1-kb flanking regions in SAM1, SAM2 and YFB.

Figure S8. Multi-dimensional scaling plot showing the clustering of the individual RNA sequencing samples.

Figure S9. Correlation analysis of gene expression between the two replicates of SAM1, SAM2 or YFB in each of the three genotypes.

Data S1. Overlapping differentially expressed genes between SAM1, SAM2 and young flower bud in hybrid.

Data S2. Gene Ontology enrichment lists for SAM1, SAM2 and young flower bud between the parents.

Data S3. Gene Ontology enrichment lists for SAM1, SAM2 and young flower bud in the hybrid.

Data S4. Hierarchical clustering in the significantly enriched Gene Ontology terms associated with the differentially expressed genes of the F_1 hybrids in SAM1 and SAM2.

Data S5. Conserved miRNAs and the recognized target genes.

Data S6. Novel miRNAs and miRNA families identified in each sample and each tissue.

Data S7. Lists of the non-additively expressed genes and ELD genes related to phosphorylation pathways.

Data S8. Number of the non-additively expressed genes that had siRNA clusters in coding regions or 1-kb flanking regions in SAM1, SAM2 and young flower bud.

Data S9. The experiment design and RNA sequencing information.

 $\ensuremath{\text{Data S10}}$. The experiment design and small RNA sequencing information.

Data S11. The genome sequencing information of both parents.

Data S12. The methylC sequencing information for the cross combination.

Methods S1. Library generation for high-throughput sequencing and bioinformatics analyses.

REFERENCES

- Alonso-Peral, M.M., Li, J., Li, Y., Allen, R.S., Schnippenkoetter, W., Ohms, S., White, R.G. and Millar, A.A. (2010) The microRNA159-regulated GAMYB-like genes inhibit growth and promote programmed cell death in Arabidopsis. *Plant Physiol.* **154**, 757–771.
- Barber, W.T., Zhang, W., Win, H., Varala, K.K., Dorweiler, J.E., Hudson, M.E. and Moose, S.P. (2012) Repeat associated small RNAs vary among parents and following hybridization in maize. *Proc. Natl Acad. Sci. USA*, 109, 10444–10449.
- Birchler, J.A., Auger, D.L. and Riddle, N.C. (2003) In search of the molecular basis of heterosis. *Plant Cell*, 15, 2236–2239.
- Birchler, J.A., Yao, H., Chudalayandi, S., Vaiman, D. and Veitia, R.A. (2010) Heterosis. *Plant Cell*, **22**, 2105–2112.
- Busov, V.B., Brunner, A.M. and Strauss, S.H. (2008) Genes for control of plant stature and form. New Phytol. 177, 589–607.
- Cao, X. and Jacobsen, S.E. (2002) Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc. Natl Acad. Sci. USA*, 99, 16491–16498.
- Chalhoub, B., Denoeud, F., Liu, S. et al. (2014) Plant genetics. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science, 345, 950–953.
- Chen, Z.J. (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu. Rev. Plant Biol.* 58, 377–406.
- Chen, Z.J. (2010) Molecular mechanisms of polyploidy and hybrid vigor. Trends Plant Sci. 15, 57–71.
- Chen, Z.J. (2013) Genomic and epigenetic insights into the molecular bases of heterosis. Nat. Rev. Genet. 14, 471–482.
- Chen, X., Li, M., Shi, J., Fu, D., Qian, W., Zou, J., Zhang, C. and Meng, J. (2008) Gene expression profiles associated with intersubgenomic heterosis in *Brassica napus. Theor. Appl. Genet.* **117**, 1031–1040.
- Cline, M.G. (1991) Apical dominance. Bot. Rev. 57, 318-358.
- Crow, J.F. (1948) Alternative hypotheses of hybrid vigor. Genetics, 33, 477– 487.
- Cubas, P., Vincent, C. and Coen, E. (1999) An epigenetic mutation responsible for natural variation in floral symmetry. *Nature*, 401, 157–161.
- Danisman, S., Van der Wal, F., Dhondt, S., Waites, R., de Folter, S., Bimbo, A., van Dijk, A.D., Muino, J.M., Cutri, L. and Dornelas, M.C. (2012) Arabidopsis class I and class II TCP transcription factors regulate jasmonic acid metabolism and leaf development antagonistically. *Plant Physiol.* 159, 1511–1523.
- Dempsey, D.M.A., Vlot, A.C., Wildermuth, M.C. and Klessig, D.F. (2011) Salicylic acid biosynthesis and metabolism. *Arabidopsis Book*, 9, e0156.
- Dombrecht, B., Xue, G.P., Sprague, S.J., Kirkegaard, J.A., Ross, J.J., Reid, J.B., Fitt, G.P., Sewelam, N., Schenk, P.M. and Manners, J.M. (2007) MYC2 differentially modulates diverse jasmonate-dependent functions in Arabidopsis. *Plant Cell*, **19**, 2225–2245.

- Gang, W. and Poethig, R.S. (2006) Temporal regulation of shoot development in Arabidopsis thaliana by Mir156 and its target SPL3. Development, 133, 3539–3547.
- Grant, I. and Beversdorf, W. (1985) Heterosis and combining ability estimates in spring-planted oilseed rape (*Brassica napus* L.). Can. J. Genet. Cytol. 27, 472–478.
- Greaves, I.K., Groszmann, M., Wang, A., Peacock, W.J. and Dennis, E.S. (2014) Inheritance of trans chromosomal methylation patterns from arabidopsis F1 hybrids. *Proc. Natl Acad. Sci. USA*, **111**, 2017–2022.
- Groszmann, M., Greaves, I.K., Albertyn, Z.I., Scofield, G.N., Peacock, W.J. and Dennis, E.S. (2011) Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor. *Proc. Natl Acad. Sci. USA*, 108, 2617–2622.
- Groszmann, M., Gonzalez-Bayon, R., Lyons, R.L., Greaves, I.K., Kazan, K., Peacock, W.J. and Dennis, E.S. (2015) Hormone-regulated defense and stress response networks contribute to heterosis in Arabidopsis F1 hybrids. *Proc. Natl Acad. Sci. USA*, 112, E6397–E6406.
- Guo, M., Rupe, M.A., Wei, J. et al. (2014) Maize ARGOS1 (ZAR1) transgenic alleles increase hybrid maize yield. J. Exp. Bot. 65, 249–260.
- Ha, M. and Chen, Z.J. (2009) Small RNAs serve as a genetic buffer against genomic shock in arabidopsis interspecific hybrids and allopolyploids. *Proc. Natl Acad. Sci. USA*. **106**, 17835–17840.
- Ha, M., Kim, E.-D. and Chen, Z.J. (2009) Duplicate genes increase expression diversity in closely related species and allopolyploids. *Proc. Natl Acad. Sci. USA*, 106, 2295–2300.
- He, G.M., Zhu, X.P., Elling, A.A. et al. (2010) Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell*, 22, 17–33.
- He, G., Elling, A.A. and Deng, X.W. (2011) The epigenome and plant development. Annu. Rev. Plant Biol. 62, 411–435.
- Hochholdinger, F. and Hoecker, N. (2007) Towards the molecular basis of heterosis. *Trends Plant Sci.* 12, 427–432.
- Jin, J., Zhang, H., Kong, L., Gao, G. and Luo, J. (2014) PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.* 42, 1182–1187.
- Kawanabe, T., Ishikura, S., Miyaji, N. et al. (2016) Role of DNA methylation in hybrid vigor in Arabidopsis thaliana. Proc. Natl Acad. Sci. USA, 113, E6704–E6711.
- Kazan, K. and Manners, J.M. (2009) Linking development to defense: auxin in plant–pathogen interactions. *Trends Plant Sci.* 14, 373–382.
- Krieger, U., Lippman, Z.B. and Zamir, D. (2010) The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat. Genet.* 42, 459–463.
- Lai, J.S., Li, R.Q., Xu, X. et al. (2010) Genome-wide patterns of genetic variation among elite maize inbred lines. Nat. Genet. 42, 1027–1030.
- Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H. and Ecker, J.R. (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell*, **133**, 523–536.
- Lu, J., Zhang, C., Baulcombe, D.C. and Chen, Z.J. (2012) Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of Arabidopsis seeds. *Proc. Natl Acad. Sci. USA*, 109, 5529–5534.
- Lyons, R., Iwase, A., Gänsewig, T., Sherstnev, A., Duc, C., Barton, G.J., Hanada, K., Higuchi-Takeuchi, M., Matsui, M. and Sugimoto, K. (2013) The RNA-binding protein FPA regulates flg22-triggered defense responses and transcription factor activity by alternative polyadenylation. *Sci. Rep.* **3**, 2866.
- Mallory, A.C. and Bartel, B. (2005) MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. *Plant Cell*, 17, 1360–1375.
- Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A.J., King, G.J., Giovannoni, J.J. and Seymour, G.B. (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat. Genet.* 38, 948–952.
- Miura, K., Ikeda, M., Matsubara, A., Song, X.J., Ito, M., Asano, K., Matsuoka, M., Kitano, H. and Ashikari, M. (2010) OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nat. Genet.* 42, 545–549.
- Ng, D.W., Zhang, C., Miller, M., Palmer, G., Whiteley, M., Tholl, D. and Chen, Z.J. (2011) cis-and trans-Regulation of miR163 and target genes

confers natural variation of secondary metabolites in two Arabidopsis species and their allopolyploids. *Plant Cell*, **23**, 1729–1740.

- Ng, D.W., Lu, J. and Chen, Z.J. (2012) Big roles for small RNAs in polyploidy, hybrid vigor, and hybrid incompatibility. *Curr. Opin. Plant Biol.* 15, 154–161.
- Ni, Z., Kim, E.-D., Ha, M., Lackey, E., Liu, J., Zhang, Y., Sun, Q. and Chen, Z.J. (2009) Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature*, 457, 327–331.
- Nobuta, K., Okrent, R., Stoutemyer, M., Rodibaugh, N., Kempema, L., Wildermuth, M. and Innes, R. (2007) The GH3 acyl adenylase family member PBS3 regulates salicylic acid-dependent defense responses in Arabidopsis. *Plant Physiol.* 144, 1144–1156.
- Palatnik, J.F., Wollmann, H., Schommer, C. et al. (2007) Sequence and expression differences underlie functional specialization of Arabidopsis microRNAs miR159 and miR319. Dev. Cell, 13, 115–125.
- Rapp, R.A., Udall, J.A. and Wendel, J.F. (2009) Genomic expression dominance in allopolyploids. *BMC Biology*, 7, 1–10.
- Rodgers, J.L. and Nicewander, W.A. (1988) Thirteen ways to look at the correlation coefficient. Am. Stat. 42, 59–66.
- Rohde, A., Morreel, K., Ralph, J., Goeminne, G., Hostyn, V., De Rycke, R., Kushnir, S., Van Doorsselaere, J., Joseleau, J.-P. and Vuylsteke, M. (2004) Molecular phenotyping of the pal1 and pal2 mutants of *Arabidop-sis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. *Plant Cell*, 16, 2749–2771.
- Shen, H.S., He, H., Li, J.G. et al. (2012) Genome-wide analysis of DNA methylation and gene expression changes in two Arabidopsis ecotypes and their reciprocal hybrids. *Plant Cell*, 24, 875–892.
- Shen, D., Suhrkamp, I., Wang, Y., Liu, S., Menkhaus, J., Verreet, J.A., Fan, L. and Cai, D. (2014) Identification and characterization of micro-RNAs in oilseed rape (*Brassica napus*) responsive to infection with the pathogenic fungus *Verticillium longisporum* using Brassica AA (*Brassica rapa*) and CC (*Brassica oleracea*) as reference genomes. *New Phy*tol. 204, 577–594.
- Shen, E., Zou, J., Hubertus Behrens, F. et al. (2015) Identification, evolution, and expression partitioning of miRNAs in allopolyploid Brassica napus. J. Exp. Bot. 66, 7241–7253.
- Shi, X., Ng, D.W., Zhang, C., Comai, L., Ye, W. and Chen, Z.J. (2012) Cisand trans-regulatory divergence between progenitor species determines gene-expression novelty in Arabidopsis allopolyploids. *Nature Communications*, 3, 950.
- Shindo, C., Lister, C., Crevillen, P., Nordborg, M. and Dean, C. (2006) Variation in the epigenetic silencing of FLC contributes to natural variation in Arabidopsis vernalization response. *Genes Dev.* 20, 3079–3083.
- Song, G.S., Zhai, H.L., Peng, Y.G. et al. (2010) Comparative transcriptional profiling and preliminary study on heterosis mechanism of super-hybrid rice. Mol. Plant, 3, 1012–1025.
- Tariq, M., Saze, H., Probst, A.V., Lichota, J., Habu, Y. and Paszkowski, J. (2003) Erasure of CpG methylation in Arabidopsis alters patterns of histone H3 methylation in heterochromatin. *Proc. Natl Acad. Sci. USA*, 100, 8823–8827.
- Vaucheret, H. (2006) Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes Dev.* 20, 759–771.
- Wang, J.W., Schwab, R., Czech, B., Mica, E. and Weigel, D. (2008) Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in *Arabidopsis thaliana*. *Plant Cell*, **20**, 1231–1243.
- Wei, G., Tao, Y., Liu, G. et al. (2009) A transcriptomic analysis of superhybrid rice LYP9 and its parents. Proc. Natl Acad. Sci. USA, 106, 7695–7701.
- Williams, L., Grigg, S.P., Xie, M., Christensen, S. and Fletcher, J.C. (2005) Regulation of Arabidopsis shoot apical meristem and lateral organ formation by microRNA miR166 g and its AtHD-ZIP target genes. *Development*, 132, 3657–3668.
- Wu, H.J., Ma, Y.K., Chen, T., Wang, M. and Wang, X.J. (2012) PsRobot: a web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res.* 40, 22–28.
- Yoo, M.J., Szadkowski, E. and Wendel, J.F. (2013) Homoeolog expression bias and expression level dominance in allopolyploid cotton. *Heredity*, 110, 171–180.
- Yu, S., Li, J., Xu, C., Tan, Y., Gao, Y., Li, X., Zhang, Q. and Maroof, M.S. (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl Acad. Sci. USA*, **94**, 9226–9231.

- Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W.-L., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M. and Jacobsen, S.E. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell*, **126**, 1189–1201.
- Zhang, H.Y., He, H., Chen, L.B. et al. (2008) A genome-wide transcription analysis reveals a close correlation of promoter INDEL polymorphism and heterotic gene expression in rice hybrids. *Mol. Plant*, 1, 720–731.
- Zhang, L., Peng, Y., Wei, X., Dai, Y., Yuan, D., Lu, Y., Pan, Y. and Zhu, Z. (2014) Small RNAs as important regulators for the hybrid vigour of super-hybrid rice. J. Exp. Bot. 65, 5989–6002.
- Zhang, O.Z., Wang, D., Lang, Z.B. et al. (2016a) Methylation interactions in Arabidopsis hybrids require RNA-directed DNA methylation and are influenced by genetic variation. Proc. Natl Acad. Sci. USA, 113, E4248– E4256.
- Zhang, Y., Zhang, D., Yu, H., Lin, B., Fu, Y. and Hua, S. (2016b) Floral initiation in response to planting date reveals the key role of floral meristem differentiation prior to budding in canola (*Brassica napus* L.). Front. Plant Sci. 7, 1369.