The Plant Cell, Vol. 26: 1878–1900, May 2014, www.plantcell.org

LARGE-SCALE BIOLOGY ARTICLE

mRNA and Small RNA Transcriptomes Reveal Insights into Dynamic Homoeolog Regulation of Allopolyploid Heterosis in Nascent Hexaploid Wheat[W][OPEN]

Aili Li,[a,1] Dengcai Liu,[b,1] Jun Wu,[c,1] Xubo Zhao,[a] Ming Hao,[b] Shuaifeng Geng,[a] Jun Yan,[c] Xiaoxue Jiang,[c] Lianquan Zhang,[b] Junyan Wu,[a] Lingjie Yin,[a] Rongzhi Zhang,[a] Liang Wu,[a] Youliang Zheng,[b] and Long Mao[a,2]

[a] National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China

[b] Triticeae Research Institute, Sichuan Agricultural University, Chengdu, Sichuan 611130, China

[c] Novogene Bioinformatics Institute, Beijing 100083, China

ORCID IDs: 0000-0001-9004-192X (A.L.); 0000-0001-5508-1125 (D.L.); 0000-0001-9488-7682 (J.W.); 0000-0002-3377-4040 (L.M.)

Nascent allohexaploid wheat may represent the initial genetic state of common wheat (*Triticum aestivum*), which arose as a hybrid between *Triticum turgidum* (AABB) and *Aegilops tauschii* (DD) and by chromosome doubling and outcompeted its parents in growth vigor and adaptability. To better understand the molecular basis for this success, we performed mRNA and small RNA transcriptome analyses in nascent allohexaploid wheat and its following generations, their progenitors, and the natural allohexaploid cultivar Chinese Spring, with the assistance of recently published A and D genome sequences. We found that nonadditively expressed protein-coding genes were rare but relevant to growth vigor. Moreover, a high proportion of protein-coding genes exhibited parental expression level dominance, with genes for which the total homoeolog expression level in the progeny was similar to that in *T. turgidum* potentially participating in development and those with similar expression to that in *Ae. tauschii* involved in adaptation. In addition, a high proportion of microRNAs showed nonadditive expression upon polyploidization, potentially leading to differential expression of important target genes. Furthermore, increased small interfering RNA density was observed for transposable element–associated D homoeologs in the allohexaploid progeny, which may account for biased repression of D homoeologs. Together, our data provide insights into small RNA–mediated dynamic homoeolog regulation mechanisms that may contribute to heterosis in nascent hexaploid wheat.

INTRODUCTION

The prevalence of polyploids in plants provides evidence for their evolutionary advantages in development and adaptation (Chen, 2007; Leitch and Leitch, 2008; Soltis et al., 2009). Many important crops, such as cotton (*Gossypium hirsutum*), Rapeseed (*Brassica napus*), and wheat (*Triticum aestivum*), are allopolyploids as well. Having multiple sets of genetic material in one cell may cause important changes at the genetic, gene expression, and epigenetic levels (Chen, 2007; Jackson and Chen, 2010). Heterozygosity and intergenomic interactions in allopolyploids may result in phenotypic variation and increased growth vigor, a phenomenon referred to as heterosis (Chen, 2007; Ni et al., 2009). Interestingly, allopolyploids often display enhanced heterosis relative to autopolyploids, and tetraploids show enhanced heterosis relative to diploids, as demonstrated in alfalfa (*Medicago sativa*; Groose et al., 1989; Bingham et al., 1994) and maize (*Zea mays*; Riddle et al., 2010; Yao et al., 2013). These effects in polyploid hybrids could be caused by increased genome dosage, allelic heterozygosity, and/or epigenetic changes (Chen, 2007, 2010; Riddle et al., 2010; Yao et al., 2013).

Bread wheat, or common wheat (*T. aestivum*), is a widely cultivated allohexaploid (2n = 6x = 42, AABBDD). It formed ~8000 years ago as a hybrid between the early-cultivated allotetraploid *Triticum turgidum* (2n = 4x = 28, AABB) and the diploid wild goat grass *Aegilops tauschii* (2n = 2x = 14, DD), followed by spontaneous chromosome doubling (Kihara, 1944; McFadden and Sears, 1946; Matsuoka, 2011; Madlung, 2013). Once established, hexaploid wheat outcompeted the tetraploid wheat varieties and, upon further domestication, showed broader adaptability to different photoperiod and vernalization conditions and improved tolerance to salt, low pH, aluminum, and frost (Dubcovsky and Dvorak, 2007). Hexaploid wheat also shows increased resistance to several pathogens and greater versatility for use in different food products (Dubcovsky and Dvorak, 2007). Since allohexaploid wheat underwent two relatively recent allopolyploidization events, it provides one of the few models for studying genetic interactions between the three homoeologous A, B, and D genomes.

Studies on duplicated genomes or genomic regions in ancient polyploids showed that they have often experienced unequal gene losses (or genome fractionation), with one genome or genomic region retaining more genes (dominant) than the other (more

[1] These authors contributed equally to this work.

[2] Address correspondence to maolong@caas.cn.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Long Mao (maolong@caas.cn).

[W] Online version contains Web-only data.

[OPEN] Articles can be viewed online without a subscription.

www.plantcell.org/cgi/doi/10.1105/tpc.114.124388

fractionated). More interestingly, genes located on the dominant genome or genomic region tend to have higher expression levels (Schnable et al., 2011; Cheng et al., 2012; Garsmeur et al., 2014). Such genome dominance phenomena have been reported for a few plants such as *Arabidopsis thaliana* (Wang et al., 2006), *Brassica* (Cheng et al., 2012), and maize (Schnable et al., 2011). Modern bread wheat is characterized by asymmetrical distribution of genetic loci underlying important agronomic traits among the three homoeologous subgenomes (Feldman et al., 2012), suggesting differential subgenome evolution. Natural hexaploid wheat (cv Chinese Spring) is estimated to have lost 10,000 to 16,000 genes following polyploidization, compared with the three diploid progenitors (Brenchley et al., 2012). Furthermore, accelerated accumulation of gene structure changes, such as alternative splicing and nonsynonymous and premature termination codon mutations, have been reported (Akhunov et al., 2013). Using the wheat syntenome and conserved orthologous set genes, Pont et al. (2013) proposed that the first neotetraploidization event resulted in subgenome dominance wherein the A subgenome was dominant over the B subgenome. The second neohexaploidization event led to a supradominance, with the D subgenome dominant over the tetraploid (subgenomes A and B). Nevertheless, expression data are needed to show whether genes in the dominant subgenomes have higher expression levels, as reported in other species (Schnable et al., 2011; Cheng et al., 2012; Garsmeur et al., 2014). Nascent allohexaploids of wheat can be synthesized in the laboratory via interspecific hybridization (Mestiri et al., 2010), facilitating the study of genetic, functional, and epigenetic changes during this process. Newly synthesized wheat allohexaploids may recapitulate the original genetic status of ancient wheat hybrids. It has long been known that nascent hexaploid wheat exhibits hybrid vigor and adaptive traits, including robust seedling growth, larger spikes, and enhanced salt tolerance (He et al., 2003; Colmer et al., 2006). Thus, transcriptome studies of nascent hexaploid wheat may shed light on the molecular basis for heterosis in hexaploid wheat.

Recent work has shown that most structural changes in hexaploid wheat may take place during allotetraploidization (Mestiri et al., 2010; Zhao et al., 2011a, 2011b). Experiments with newly formed allotetraploid wheat revealed localized genomic changes and rapidly arising copy-number variation of gene homologs (Zhang et al., 2013b). Allohexaploid wheat, however, is genetically stable when considering true euploid plants (Mestiri et al., 2010; Zhao et al., 2011b; Zhang et al., 2013a) and is unlikely to have experienced significant genome structure changes. In this context, nascent allohexaploid wheat is a good model to study homoeolog expression dynamics and the corresponding genetic and epigenetic regulation.

Indeed, several studies have compared resynthesized and/or natural wheat allopolyploids directly to their parents using microarrays and other techniques (He et al., 2003; Bottley et al., 2006; Pumphrey et al., 2009; Akhunova et al., 2010; Chagué et al., 2010; Qi et al., 2012; Chelaifa et al., 2013). In these studies, the average or the sum of parental gene expression levels (i.e., mid-parent value [MPV]) was employed to determine nonadditive expression of genes in the allohexaploid progeny (Wang et al., 2006). However, additivity, rather than nonadditivity, was found to be pervasive (Chagué et al., 2010; Chelaifa et al., 2013). The lack of nonadditive genetic interactions in resynthesized wheat hybrids calls for additional explanations of the molecular basis underlying the growth vigor and adaptability of allohexaploid wheat. Recent work in allotetraploids showed that a significant proportion of genes are expressed at a level equal to that of one progenitor and different from that of the other one, irrespective of the comparison to the MPV and the additivity of expression (Rapp et al., 2009). This phenomenon is termed "expression-level dominance" (hereafter, ELD) for the implicated genes. However, the roles of these genes in allopolyploids remain unclear (Rapp et al., 2009; Grover et al., 2012; Yoo et al., 2013). We were interested in determining whether similar expression patterns may be found in nascent allohexaploid wheat.

Small RNAs, including microRNAs (miRNAs) and small interfering RNAs (siRNAs), regulate gene expression by posttranscriptional mechanisms and through epigenetic modifications (Vaucheret, 2006; Lu et al., 2012). In interspecific hybrids and allopolyploids in the genus *Arabidopsis*, variation in miRNA expression causes nonadditive expression of target genes that may affect growth vigor and adaptation (Ha et al., 2009a). Further study has shown that *cis*- and *trans*-regulation 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 regulate gene expression by directing DNA methylation. For example, siRNAs may function as regulators for parental genome imbalance and gene expression in triploid *Arabidopsis* created from a diploid maternal parent and a tetraploid paternal parent (Lu et al., 2012), leading to enlarged seed size. In *A. thaliana* and *Arabidopsis arenosa* interspecific hybrids and allotetraploids, siRNAs preferentially repress *A. thaliana* homoeologous loci when using the former as the maternal parent (Comai et al., 2000; Wang et al., 2004, 2006), causing morphological dominance of *A. arenosa* over *A. thaliana* in the allotetraploids. These observations indicate that siRNAs may regulate genome-wide gene expression bias in allopolyploids and are essential for plant heterosis. In wheat, which contains abundant TE-derived repetitive sequences (>80%), the number of siRNAs corresponding to TEs strongly decreased upon polyploidization (Kenan-Eichler et al., 2011), suggesting that they may play significant roles during allohexaploidization.

Emerging data demonstrate that allohexaploidization likely involves small RNA–mediated genome modification and gene regulation (Chen, 2013). However, analysis of genome-wide individual homoeolog expression and small RNA abundance requires genome sequence information (Ha et al., 2009b; Grover et al., 2012; Arnaud et al., 2013). The draft genome sequences of common wheat A and D genome donors (diploid wild einkorn *Triticum urartu* and diploid wild goat grass *Ae. tauschii*, respectively) have recently become available (Jia et al., 2013; Ling et al., 2013), with reasonable high sequence quality at gene coding regions and their vicinities. Together with next-generation sequencing technology, these resources provide unprecedented opportunities to address such questions as how allohexaploidization affects the expression of homoeologs and alters molecular pathways that may contribute to growth vigor and adaptation in nascent allohexaploid wheat and whether small RNAs play roles in this process. Toward this end, we performed RNA-seq and

small RNA analyses in three tissues of four consecutive generations of nascent allohexaploid wheat, their progenitors, and the common wheat cultivar Chinese Spring (CS). Our gene expression and small RNA abundance data suggest a possible small RNA–mediated regulatory mechanism for homoeolog expression that may be important for heterosis in nascent allohexaploid wheat.

RESULTS

Nascent Allohexaploid Wheat Resembles the Tetraploid Progenitor *T. turgidum* and Shows Heterosis

Despite the difference in ploidy, allohexaploid wheat can be generated using tetraploid *T. turgidum* as the maternal parent. In this study, we used *T. turgidum* ssp *dicoccon* accession PI94655 and the diploid *Ae. tauschii* ssp *strangulata* accession AS2404 as parental lines (Zhang et al., 2010). Hexaploid progeny were produced in F1 plants by spontaneous chromosome doubling. Genomic in situ hybridization and fluorescence in situ hybridization were performed on root tip cells for chromosome identification, and only plants with complete sets of chromosomes (28 from *T. turgidum* and 14 from *Ae. tauschii*) were used for further experimentation (Figure 1A). Our newly synthesized hexaploid wheat predominantly resembled *T. turgidum* in grain length, plant height, plant architecture, endosperm starch composition, powdery mildew resistance, and spring flowering habit (Figure 1; Supplemental Figure 1). Plants also displayed growth vigor in some traits when compared with parental lines, including more robust seedling growth and larger spikes with elongated rachis internodes (Figure 1B; Supplemental Figure 1A).

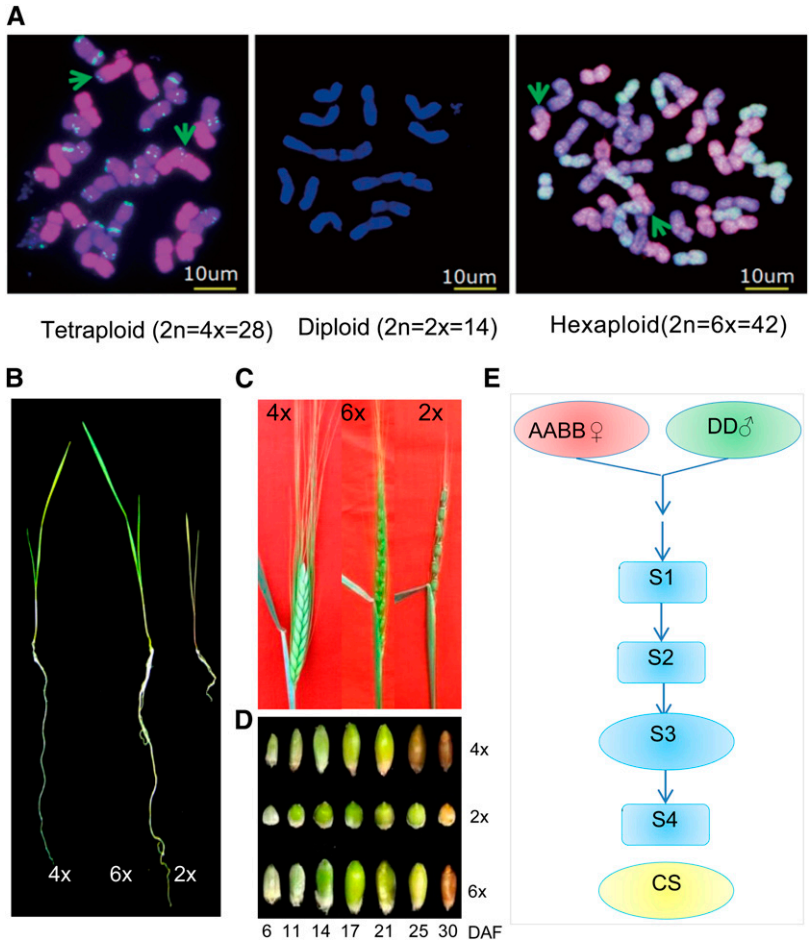


Figure 1. Characterization of Nascent Allohexaploid Wheat.

(A) Chromosomes in root tip cells. Green arrows indicate the 4A/7B chromosome translocations in synthetic hexaploid and its *T. turgidum* parent PI 94655 (4x).
(B) Seven-day-old seedlings of *T. turgidum* (4x), synthetic wheat (6x), and *Ae. tauschii* (2x).
(C) Heading stage spikes of the third generation of self-pollinated allohexaploid wheat (S3, 6x) and its parents.
(D) Developing seeds of S3 allohexaploid plants and their progenitors. 4x, *T. turgidum*; 2x, *Ae. tauschii*; 6x, newly synthesized allohexaploid wheat.
(E) Sampling schema. Samples in oval circles have biological replicates. AABB, *T. turgidum*; DD, *Ae. tauschii*; S1 to S4, consecutive generation of selfed allohexaploid wheat.

mRNA and Small RNA Transcriptome Sequencing in Nascent Hexaploid Wheat, Its Progenitors, and Common Wheat Chinese Spring

To study mRNA and small RNA expression dynamics during wheat allopolyploidization, we constructed and sequenced 25 RNA-seq libraries and 25 small RNA libraries from four consecutive generations of newly synthesized hexaploid wheat and their progenitors, with CS as a reference for natural hexaploid wheat (Supplemental Data Sets 1 and 2). Three types of tissues were analyzed: 7-d-old seedlings, heading-stage young spikes, and immature seed at 11 d after flowering (DAF) (Figure 1E). For gene expression quantification, we first evaluated the suitability of the three sets of gene models as references (Brenchley et al., 2012; Jia et al., 2013; Ling et al., 2013). We found that gene models from the 5x CS genome sequences gave rise to fragmented information and hence were not suitable for our purposes. Instead, we used the complete gene sets available from the D (*Ae. tauschii*) and A (*T. urartu*) genomes. For convenience, we hereafter use the species names and their genome representations interchangeably, i.e., DD for *Ae. tauschii* and AABB for *T. turgidum*. In most cases, we used the data from the third generation of self-pollinated plants (S3) as representative for newly synthesized allohexaploid progeny because biological duplicates were available for both RNA-seq and small RNA sequencing in S3, the two progenitors, and CS. Data from other generations (S1, S2, and S4) were used to identify the transgressive heritability of gene expression patterns across early generations.

Among all RNA-seq reads, 69% could be mapped to the D genome and 54% be mapped to the A genome (Supplemental Data Set 1 and Supplemental Figure 2). We then mapped them to the 43,150 annotated D gene models and found an average of 24,658 genes expressed (~57%) in each of the three tissues. A comparable number of A genome models (22,733 out of 34,879; ~65%) were detected in our RNA-seq data. We subsequently used BLAST to search the A and D gene models reciprocally and found that *Ae. tauschii* (D) appeared to confer more genome-specific genes (3316) than *T. urartu* (A; 1728) at an E value ≤ 1e-10 (BLASTN identity >90%; Supplemental Figures 3A to 3D). To avoid bias toward D-specific gene models, we blasted A gene models against the D genome and used BLAST-aligned homologous regions as references for gene expression quantification. We also used these D genome regions as references to identify single nucleotide polymorphisms (SNPs) between homoeologous reads of the two progenitors. SNP coordinates were compared with those of D gene models and total read numbers on one gene model were summed and used to represent the expression levels of individual homoeologs according to their parental origin (more details in following sections).

For interlibrary comparison, read numbers were normalized to relative abundance as reads per kilobase of exon model per million mapped base pairs (RPKM) (Mortazavi et al., 2008). We found that distributions of gene expression levels were comparable irrespective of tissue type or ploidy (Supplemental Figure 4). A-specific genes (P < 1e-32, ANOVA test) derived from A-D reciprocal BLAST searches (above) indeed showed no expression in *Ae. tauschii* tissues, while rarely could D-specific reads be mapped to the A genome (P < 1e-7; Supplemental Figure 5). Gene expression correlations between two biological replicates were high, with average coefficients (R^2) of 0.95, 0.97, and 0.89 for seedlings, young spikes, and immature seed, respectively (Figure 2A). A correlation dendrogram showed that synthetic wheat was clustered with the tetraploid progenitor in all three tissues studied, while natural hexaploid CS appeared to be the outlier (Figure 2B), indicating that the synthetic wheat retained similar gene expression patterns to the tetraploid progenitor and that there was marked divergence between the nascent and natural hexaploid wheat.

Tissue-specific expression was studied using an empirical cutoff value for positively expressed genes (RPKM ≥ 1) (Gan et al., 2010). The results showed no significant differences in the number of tissue-specifically expressed genes among *T. turgidum*, *Ae. tauschii*, and S3 progeny (Figures 2C to 2E). Next, we investigated Gene Ontology (GO) functional categories among these genes. In S3 plants, seedling-specific genes (1550 genes) were significantly enriched for secondary metabolic process and pathways for stress responses, and spike-specific genes (1534 genes) were significantly enriched for transport and cell growth pathways, whereas seed-specific genes (768 genes) were functionally enriched in the embryonic development pathway (Figures 2F to 2H). Similar results were obtained for tissue-specific genes in the two progenitors (Supplemental Figure 6). Therefore, we generated tissue-specific transcriptomes for newly synthesized hexaploid wheat and its progenitors for further comparative analysis.

Transcriptome Differentiation among Tissues and Species during Wheat Allohexaploidization

To study the effect of allohexaploidization on gene expression in synthetic wheat, we first compared gene expression levels between the two progenitors. Among ~20,000 genes expressed in each of the three tissues, young spikes showed the highest number of differentially expressed genes (P < 0.05, Benjamini-Hochberg [BH] multiple test correction) between *T. turgidum* and *Ae. tauschii* (AB-D_{dif}, 4319, 21.4%), followed by seedlings (1672, 8.8%) and then immature seed (215, 1.1%) (Figure 3A; Supplemental Figures 7A and 7B and Supplemental Table 1). Such differences in the transcriptomes were correlated with the morphological differences among the three tissues studied, as spikes differed most markedly between the tetraploid and diploid progenitors (Figures 1B to 1D; Supplemental Figure 1A).

Among 4319 genes in young spikes that were differentially expressed between the parental lines, 1866 (43.2%) were expressed at a higher level in the tetraploid (AB>D), with the remaining more highly expressed in the diploid (AB<D, 56.8%, 2453). Similarly, in both seedlings and immature seed, there were fewer AB>D genes (41.7 and 38.6%, respectively) than AB<D genes (58.3 and 61.4% in 1672 and 215 AB-D_{dif} genes, respectively) (Figure 3A; Supplemental Figures 7A and 7B and Supplemental Table 1). This observation conflicts with a previous report wherein the tetraploid progenitor conferred more high-expression genes (60%) than the diploid (40%) (Chelaifa et al., 2013).

We then studied gene expression levels between the progeny and each parental line. We found more differentially expressed genes (P < 0.05, BH multiple test correction) between progeny and the diploid progenitor (S3-D_{dif}) than between progeny and the tetraploid progenitor (S3-AB_{dif}). In young spikes, for instance,

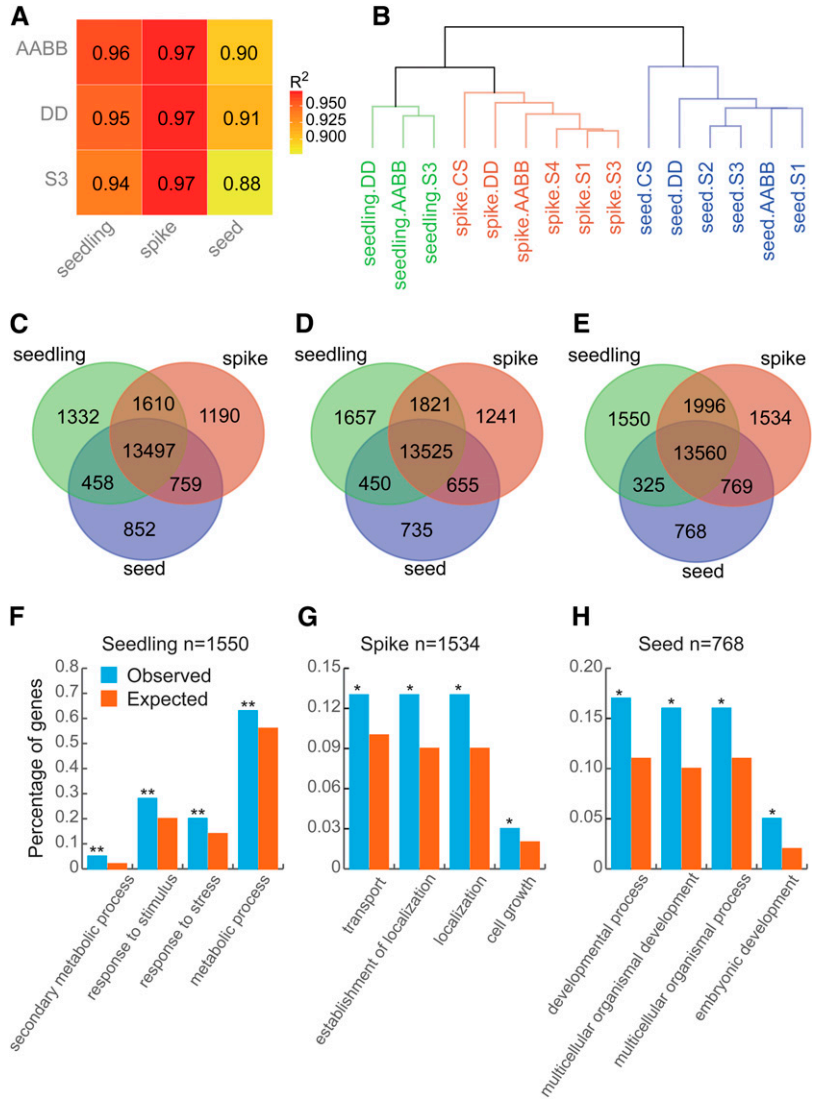
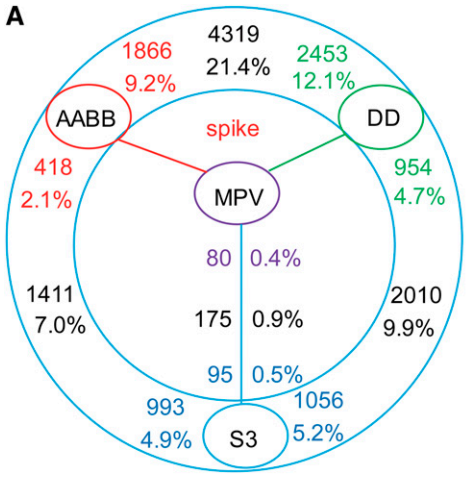


Figure 2. Global Characterization of Gene Expression Patterns among Three Tissues in Nascent Hexaploid Wheat and Its Progenitors.

AABB, *T. turgidum*; DD, *Ae. tauschii*; S3, the third generation of self-pollinated allohexaploid wheat.
(A) Correlation coefficients between gene expression data sets from two biological duplicates.
(B) Cluster dendrogram showing global relationships of gene expression in different tissues and species. The branch length indicates the degree of variance.
(C) to **(E)** Venn diagram analyses of tissue specific genes in *T. turgidum* **(C)**, *Ae. tauschii* **(D)**, and S3 allohexaploid wheat **(E)**.
(F) to **(H)** Functional categories of genes showing tissue-specific expression in seedlings **(F)**, young spikes **(G)**, and immature seed **(H)** of S3 plants. For each tissue, only genes whose transcripts were detected in both allohexaploid progeny and its progenitor were considered. FDR-adjusted P values, *P < 0.05 and **P < 0.01, respectively. Observed, numbers of genes observed in this study; Expected, numbers of genes in this same category in the GO enrichment analysis program.

S3-D_{dif} genes accounted for 9.9% (2010) of the total expressed genes (20,204), more than the S3-AB_{dif} genes (7.0%, 1411; Figure 3A). Similar patterns were found for seedlings (Supplemental Figure 7A and Supplemental Table 1), although the numbers were smaller (248 S3-D_{dif} genes versus 132 S3-AB_{dif} genes out of a total of 18,925). In the natural allohexaploid wheat CS, we also found more CS-D_{dif} genes than CS-AB_{dif} genes in young spikes (4347 versus 2490; Supplemental Table 1), indicating that hexaploid wheat is indeed more highly diverged from *Ae. tauschii* than from *T. turgidum* in terms of transcriptome profiles.

To determine nonadditively expressed genes, we compared gene expression levels of nascent allohexaploid progeny with MPVs derived from the two parental lines (Wang et al., 2006). We found, consistent with the previous report (Chelaifa et al., 2013), expression additivity for the majority of transcripts in nascent synthetic wheat. Nonadditively expressed genes (P < 0.05, BH



B

Spike-nonadditively activated (95)

Ontology	GO term	FDR
BP	cell growth	0.00012
BP	regulation of anatomical structure size	0.00012
BP	regulation of cell size	0.00012
BP	regulation of cellular component size	0.00012
BP	growth	0.00073
MF	enzyme regulator activity	4.30E-08
MF	hydrolase activity	0.00029
MF	catalytic activity	0.00054

Spike-nonadditively repressed (80)

Ontology	GO term	FDR
BP	carbohydrate metabolic process	0.0022
MF	catalytic activity	1.50E-05
MF	hydrolase activity	0.00041
CC	external encapsulating structure	2.00E-05
CC	cell wall	2.00E-05

Figure 3. Nonadditively Expressed Genes in Young Spikes of Nascent Allohexaploid Wheat.

(A) Genes differentially expressed in S3 progeny and their tetraploid (AABB) and diploid (DD) progenitors. Numbers close to the species (colored) represent upregulated genes compared with the neighboring species. Percentages indicate those among all expressed genes in young spikes. The total number of genes differentially expressed between two species is given (black).

(B) GO enrichment analysis of nonadditively expressed genes. Shown are significantly enriched GO terms (Fisher test FDR < 0.05). BP, biological process; MF, molecular function; CC, cellular component.

multiple test correction) represented only a small portion of the expressed genes. In young spikes, for example, 175 (0.9% of 20,204 expressed genes) genes displayed nonadditive expression (Figure 3A; Supplemental Data Set 3). The percentage in seedlings was even lower (22, 0.1%) and none were detected in immature seed (Supplemental Figures 7A and 7B). GO enrichment analysis of genes nonadditively activated in S3 young spikes revealed enrichment for cell growth function (Figure 3B; false discovery rate [FDR] < 0.001, Yekudieli FDR dependency), including nine cell size regulation genes (such as expansins) and one similar to *OsGH3.1*, which affects auxin homeostasis in rice (*Oryza sativa*; Supplemental Table 2) (Jain et al., 2006; Du et al., 2012). By contrast, genes that were nonadditively repressed were enriched for carbohydrate metabolic process (Figure 3B). In addition, more than 21.4% of the nonadditively expressed genes (54 out of 252) were transgressively heritable across S1, S3, and S4 generations (Supplemental Figure 8 and Supplemental Table 3). From S1 to S4, the number of nonadditively expressed genes was reduced from 252 to 197 (Supplemental Figure 8). Analysis of more generations is needed to confirm whether or not such a change is related to inbreeding depression. Overall, the scarcity of nonadditively expressed genes in nascent allohexaploid wheat, as shown here and elsewhere (Chagué et al., 2010; Qi et al., 2012; Chelaifa et al., 2013), invites further exploration for additional mechanisms that may be relevant to heterosis in newly synthesized wheat.

Parental Expression Level Dominance in Nascent Allohexaploid Wheat

Genes in allopolyploids have been found to display parental ELD (Rapp et al., 2009; Flagel and Wendel, 2010; Grover et al., 2012), in which some genes are differentially expressed between the parental lines and exhibit expression level in the progeny that is statistically similar to that of one parent (Yoo et al., 2013). Here, we classified genes into 12 bins (as shown in Figure 4A) according to their differential expression patterns among the progenitors and/or the progeny (AB-S3-D_{dif}) following the method of Yoo et al. (2013). We designated genes with expression levels in S3 that are statistically similar to those in *T. turgidum* as ELD-ab genes and those similar to *Ae. tauschii* as ELD-d genes.

In young spikes, for example, more than 47% (2331) of AB-S3-D_{dif} genes showed parental ELD among genes differentially expressed between the progenitors and/or between the progeny and the two progenitors (4899) (Figure 4A; Supplemental Table 4, categories II, IV, IX, and XI). There were fewer genes showing parental ELD in seedlings and immature seed, representing ~20% of AB-S3-D_{dif} genes (346/1699 in seedlings; 46/222 in immature seed; Supplemental Data Set 4). S3 seedlings had significantly more ELD-ab genes than ELD-d genes (2-fold, 231/115) ($P < 2.185e-10$, binomial test) (Figure 4B), and the same was true of S3 young spikes (1.69-fold, 1465/866) ($P < 2.2e-16$) (Figure 4C). In other words, genes in nascent allohexaploid wheat displayed ELD bias toward the tetraploid progenitor.

We further pursued possible functions of genes showing parental ELD using S3 seedlings and young spikes as examples. In seedlings, GO analysis of ELD-ab genes (IV+IX, 231) revealed enrichment of genes for the molecular function "catalytic activity" ($P < 0.001$, BH multiple test correction), while among the ELD-d genes (II+XI, 115), the biological process term "response to stimulus" was enriched ($P < 0.001$) (Figure 4C). In S3 young spikes, the ELD-ab genes (IV+IX, 1465) were significantly enriched for "cytoplasm" ($P < 0.001$), while those of ELD-d (II+XI, 866) were enriched for "cell wall and membrane" ($P < 0.001$)

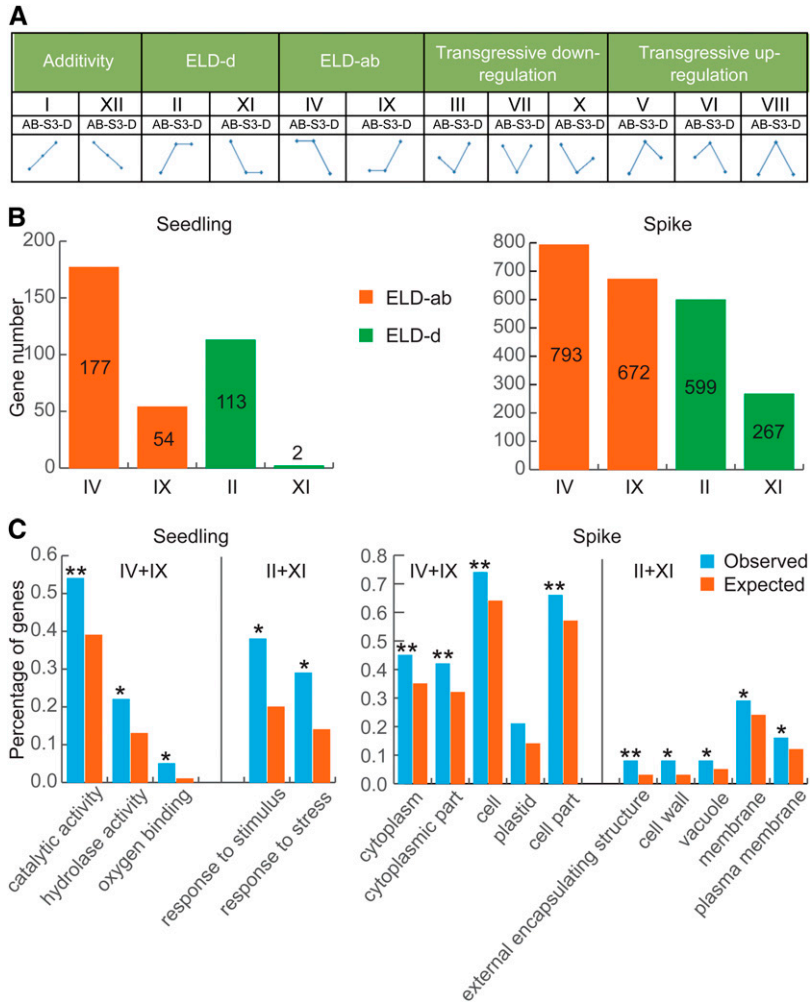
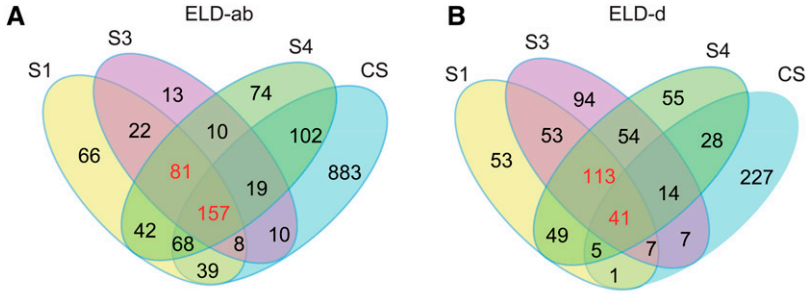


Figure 4. ELD of Genes in Seedlings and Young Spikes of Nascent Allohexaploid Wheat.

ELD-ab, genes with expression level similar to that in *T. turgidum*; ELD-d, genes with expression level similar to that in *Ae. tauschii*.
(A) Twelve bins of differentially expressed genes. AB, *T. turgidum*; D, *Ae. tauschii*; S3, the third generation of self-pollinated allohexaploid progeny.
(B) Numbers of genes showing four types of parental ELD in seedlings and young spikes. Bins II, IV, IX, and XI correspond to those in **(A)**.
(C) Enriched GO terms of genes showing parental ELD in seedlings and young spikes. Fisher test, *FDR < 0.05 and **FDR < 0.01.

(Figure 4C), indicating clear functional distinctions between the two groups of genes. In addition, nearly half (238 out of 483, 49.3%) of the ELD-ab genes were transgressively heritable across S1 to S4 (Figure 5A). A similar rate (154 out of 322, 47.8%) was found for ELD-d genes (Figure 5B). The ELD-ab heritable genes were enriched for "catalytic activity" and those of ELD-d were enriched for "transporter activity" (Figure 5C; Supplemental Data Set 4). In CS, more ELD-ab genes (157, 32.5%) were found than those of ELD-d (41, 12.7%; Figure 5A).

We then classified the genes showing parental ELD using the MapMan program and found that they fall into five major groups: secondary metabolism, lipid metabolism, RNA, stress, and transport (Figure 5D). The ELD-ab set contained more genes for secondary metabolism and lipid metabolism (17 and 11, respectively) than the ELD-d set (3 and 2, respectively), and the ELD-d set contained more stress and transport function genes (17 and 19, respectively) than the ELD-ab set (8 and 12, respectively). More interestingly, in the ELD-ab set, MADS box genes were found in "RNA process" bin, including key flower development genes *PISTILLATA* (*PI*, TRIUR3_15782), *APETALA3* (*AP3*, TRIUR3_34584), and *AGAMOUS* (*AG*, TRIUR3_29719) (Table 1), suggesting that ELD-ab genes may tend to participate in plant development. By contrast, among ELD-d genes for "stress" functions, eight were associated with disease resistance, including four *R* genes (TRIUR3_23925, TRIUR3_00931, TRIUR3_10098, and TRIUR3_06705), two chitinase genes (TRIUR3_14943 and TRIUR3_05515), and two HSP90-related genes (TRIUR3_16460 and TRIUR3_07638), which are absent among ELD-ab genes (Table 2). Two genes involved in the "transport" pathway, TRIUR3_06631 (a homolog of *Arabidopsis* potassium channel *AKT1*) and TRIUR3_24369 (a sodium-hydrogen antiporter), are interesting and may be implicated in salt



C

ELD-ab (81+157)

Ontology	GO term	FDR
BP	lipid metabolic process	2.40E-06
MF	catalytic activity	6.90E-08
CC	cytoplasmicpart	0.037

ELD-d (113+41)

Ontology	GO term	FDR
BP	transport	0.021
MF	transporter activity	0.00035
CC	vacuole	0.049

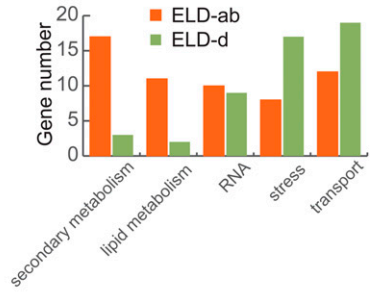


Figure 5. Transgressive Inheritance of Genes with Parental ELD in Young Spikes of Nascent Synthetic Wheat.

AABB, *T. turgidum*; DD, *Ae. tauschii*; S1 to S4, consecutive generations of self-pollinated allohexaploid wheat. ELD-ab, genes with expression level similar to that in *T. turgidum*; ELD-d, genes with expression level similar to that in *Ae. tauschii*.
(A) Venn diagram of ELD-ab genes among S1, S3, and S4 nascent allohexaploid wheat and their preservation in CS.
(B) Venn diagram of ELD-d genes among S1, S3, and S4 nascent allohexaploid wheat and their preservation in CS.
(C) Enriched GO terms among genes displaying parental ELD and shared by S1, S3, and S4 generations of nascent allohexaploid wheat.
(D) Top five MapMan bins of ELD-ab and ELD-d genes in **(C)**.

tolerance of nascent allohexaploid wheat (Supplemental Table 5), a trait that has recently been shown to be instantly acquired by *Arabidopsis* polyploids (Chao et al., 2013). ELD-d genes also included two key flowering genes, *LATE ELONGATED HYPOCOTYL* (*LHY*) for circadian rhythm (Schaffer et al., 1998) and the flowering promoter *CONSTANS* (*CO*) (Putterill et al., 1995) (Table 1).

Homoeolog Expression in Nascent Allohexaploid Wheat

ELD in polyploid progeny reflects the collective expression of all homoeologs. To distinguish homoeolog contributions to total gene expression level in newly synthesized allohexaploid wheat, we identified SNPs by separately mapping *T. turgidum* and *Ae. tauschii* reads to the D genome sequences and comparing each base position corresponding to exons of 33,151 gene models that were shared with the A gene set (Supplemental Figure 3D). Before mapping the reads, polymorphic loci between A and D genomes were masked to avoid mapping bias that may not be excluded using mismatch cutoff limitations. The method of Degner et al. (2009) was adapted to filter heterozygous loci at A and D homologous regions. For SNP calling, we also masked the heterozygous sites among *T. turgidum* reads, and thus we considered only sites that were monomorphic between A and B homoeologs (called 'AB' homoeologs), but polymorphic between AB and D (see Methods for more details).

SNPs that differed between two parental lines were used to quantify homoeolog expression in the allohexaploid wheat. We plotted the relative density of AB genome SNP depths against their ratios in total SNPs and found that the median value was close to 0.6, indicating that more AB homoeologs fell in higher expression categories than did D homoeologs in allohexaploid wheat, irrespective of whether the A or the D genome was used as a reference (Figures 6A and 6B). In addition, the CS genome was used as cross-validation for the effect of the bias-removal step. A similar pattern for SNP depth plot was found (Supplemental Figure 9). SNPs were then used as markers to discriminate D homoeologs from those of AB, with normalized read counts or SNP depths representing their expression levels. For convenience, we use AB_{pa} and D_{pa} to represent total expression levels of genes in *T. turgidum* and *Ae. tauschii*, respectively, and AB_{pr} and D_{pr} for corresponding homoeologs in the allohexaploid progeny.

To study the effect of polyploidization on homoeolog expression, we evaluated their expression levels in the hexaploid progeny relative to those in the progenitors. In young spikes, among 10,376 SNP-marked homoeologs, more than 59% (6153) showed similar expression between the two progenitors and in the progeny, i.e., when two homoeologs were expressed at similar levels between the two parental lines (AB_{pa}=D_{pa}), the expression of AB_{pr} was also equal to D_{pr} in the allohexaploid progeny (AB_{pr}=D_{pr}; Figure 6C; Supplemental Table 6). However, only 298 parental differential expression genes displayed such preconditioned expression

Table 1. Parental ELD Genes of MapMan "RNA process" Bin That Were Heritable among S1, S3, and S4 and Preserved in CS in Young Spikes

Gene Model	Category[a]	MapMan Bin Name	MapMan Annotation[b]
ELD-ab			
TRIUR3_20260	IV	RNA.processing.ribonucleases	RNS1 (RIBONUCLEASE1); endoribonuclease/ribonuclease
TRIUR3_32519	IV	RNA.transcription	DNA binding/DNA-directed RNA polymerase
TRIUR3_32766	IX	RNA.regulation of transcription.bHLH,Basic Helix-Loop-Helix family	Protein homodimerization/transcription activator
TRIUR3_08114	IX	RNA.regulation of transcription.C2C2(Zn) GATA transcription factor family	Zinc finger (GATA type) family protein
TRIUR3_15782	IV	RNA.regulation of transcription.MADS box transcription factor family	PI
TRIUR3_34584	IV	RNA.regulation of transcription.MADS box transcription factor family	AG
TRIUR3_29719	IV	RNA.regulation of transcription.MADS box transcription factor family	AP3
TRIUR3_16869	IX	RNA.regulation of transcription.NAC domain transcription factor family	REVERSE DUF547 protein
TRIUR3_23294	IX	RNA.regulation of transcription.Chromatin Remodeling Factors	DRD1 (DEFECTIVE IN RNA-DIRECTED DNA METHYLATION1)
TRIUR3_24217	IV	RNA.regulation of transcription.unclassified	CPPM_DIACA putative carboxyvinyl-carboxyphosphonate phosphorylmutase
ELD-d			
TRIUR3_22250	II	RNA.processing	PAB7 [POLY(A) BINDING PROTEIN 7]
TRIUR3_27020	XI	RNA.regulation of transcription.MYB domain transcription factor family	*Hordeum vulgare* (barley) REB1 protein
TRIUR3_13396	XI	RNA.regulation of transcription.MYB-related transcription factor family	LHY
TRIUR3_15470	II	RNA.regulation of transcription.bZIP transcription factor family	ATBZIP25, BZO2H4, BZIP25
TRIUR3_05382	II	RNA.regulation of transcription.zf-HD	ATHB22 (HOMEOBOX PROTEIN22)
TRIUR3_34989	II	RNA.regulation of transcription.unclassified	Chloroplast nucleoid DNA binding protein
TRIUR3_20618	II	RNA.regulation of transcription.unclassified	Zinc finger CONSTANS-related
TRIUR3_08677	II	RNA.RNA binding	RNA recognition motif–containing protein
TRIUR3_31910	II	RNA.RNA binding	Binding/zinc ion binding

[a]As shown in Figure 4A.
[b]Using rice proteins as template.

patterns, with 108 showing $AB_{pr}>D_{pr}$ and 190 $AB_{pr}<D_{pr}$ (Figure 6C; Supplemental Table 6). By contrast, 2503 genes showed different or even opposite homoeolog expression patterns in the progeny relative to their patterns in the two progenitors. For example, 86% (2158 out of 2503) of $AB_{pa}=D_{pa}$ genes displayed $AB_{pr}>D_{pr}$ patterns in allohexaploid progeny (FDR < 0.001; binomial test), while 24 (1%) genes showed $AB_{pa}<D_{pa}$ ($P < 0.001$; binomial test). Interestingly, 320 (13%) $AB_{pa}<D_{pa}$ genes reversed their expression patterns in the progeny ($AB_{pr}>D_{pr}$), while only 1 (0.04%) $AB_{pa}>D_{pa}$ gene was reversed. Among 10,376 SNP-marked homoeolog pairs, 1422 (14%) parental differential genes ($AB_{pa}>D_{pa}$ and $AB_{pa}<D_{pa}$) lost their homoeolog expression difference and became $AB_{pr}=D_{pr}$ (Supplemental Table 6), while 24% (2478 out of 10,376) were switched to $AB_{pr}>D_{pr}$ and only 0.2% (25 out of 10,376) to $AB_{pr}<D_{pr}$ during polyploidization. In total, 2586 (novel, 2478; preconditioned, 108) genes showed $AB_{pr}>D_{pr}$ patterns and only 215 (novel, 25; preconditioned, 190) displayed $AB_{pr}<D_{pr}$ ($P < 2.2e\text{-}16$, binomial test) in hexaploid progeny, indicating strong bias toward higher AB homoeolog expression in young spikes. Thus, for a large number of genes, biased homoeolog expression was not preconditioned in the progenitors but was dynamically regulated during allohexaploidization.

We further investigated the contribution of homoeolog expression bias (defined as unequal relative expression of homoeologs in progeny) toward parental ELD (defined as total expression of homoeologs in progeny being the same as only one of the parents). We found evidence of interactive regulation of the AB and D homoeologs for ~44% (179 out of 405) of genes showing parental ELD (Supplemental Table 7). As shown in Figure 6D, for ELD-ab genes (IV+IX), the number of genes for which only D homoeologs were up- or downregulated relative to their parental expression levels was higher than that for which only AB homoeologs showed a changed (55 versus 29). Conversely, among ELD-d genes (II+XI), the number of genes for which only AB homoeologs were up- or downregulated relative to their parental expression levels was higher than that for which only D homoeologs showed a change (54 versus 27). Nevertheless, the total numbers of genes showing differential regulation of only one of the homoeologs in the progeny relative to their parents were similar for D and AB homoeologs (82 versus 83). Thus, there was no clearly biased AB or D homoeolog

Table 2. Parental ELD Genes of MapMan "Stress" Bin in Young Spikes That Were Transgressively Heritable among S1, S3, and S4 and Preserved in CS

Gene Model	Category[a]	MapMan Bin Name	Annotation[b]
ELD-ab			
TRIUR3_33888	IV	Stress.abiotic	γ-Glutamyltranspeptidase 1 precursor
TRIUR3_29968	IV	Stress.abiotic	α-DOX2, putative
TRIUR3_10635	IV	Stress.abiotic.heat	RNA polymerase–associated transcription specificity factor
TRIUR3_14289	IX	Stress.biotic	γ Herpesvirus capsid protein
TRIUR3_24271	IX	Stress.abiotic	BAG domain–containing protein
TRIUR3_07748	IX	Stress.biotic	Transferase family
TRIUR3_25446	IX	Stress.biotic.PR-proteins	Dirigent-like protein pDIR17
TRIUR3_07957	IX	Stress.abiotic.unspecified	Cupin domain containing protein
ELD-d			
TRIUR3_16460	XI	Stress.abiotic.drought/salt	Activator of Hsp90 ATPase homolog 1-like
TRIUR3_33197	II	Stress.abiotic.drought/salt	Methyltransferase
TRIUR3_21012	II	Stress.abiotic.drought/salt	Methyltransferase
TRIUR3_21476	XI	Stress.abiotic.heat	DnaK family protein
TRIUR3_07638	XI	Stress.abiotic.heat	Hsp90 protein
TRIUR3_18140	XI	Stress.abiotic.heat	MreB/Mbl protein
TRIUR3_32736	XI	Stress.abiotic.heat	NADH dehydrogenase subunit 5 C terminus
TRIUR3_07085	II	Stress.abiotic.unspecified	Pollen allergen and extensin family
TRIUR3_14943	II	Stress.biotic	CHIT3, Chitinase family protein
TRIUR3_05515	II	Stress.biotic	CHIT7, Chitinase family protein precursor
TRIUR3_05094	II	Stress.biotic	Thaumatin
TRIUR3_06705	II	Stress.biotic.PR-proteins	Disease resistance protein RGA2
TRIUR3_00931	II	Stress.biotic.PR-proteins	NB-ARC domain–containing protein
TRIUR3_23925	II	Stress.biotic.PR-proteins	NBS-LRR disease resistance protein
TRIUR3_10098	II	Stress.biotic.PR-proteins	NBS-LRR disease resistance protein
TRIUR3_06270	II	Stress.biotic.PR-proteins	Pollen signaling protein with adenylyl cyclase activity
TRIUR3_17466	II	Stress.biotic.respiratory burst	Respiratory burst NADPH oxidase

[a]As shown in Figure 4A.
[b]By BLASTP against rice protein data set with E value < 1e-20.

contribution to parental ELD. Only ~15% (61 out of 405) of genes showed no change in homoeolog expression levels in both the progeny and the progenitors (Supplemental Table 7). Therefore, the expression of a marked portion of genes with parental ELD in allohexaploid progeny is the result of dynamic regulation of AB and D homoeologs.

Nonadditive Expression of miRNAs and Its Relevance to Growth Vigor and Adaptation

Small RNAs are important players in interspecific hybridization and polyploidization (Ha et al., 2009b; Ng et al., 2012). To study the expression dynamics of small RNAs and their potential roles in gene expression regulation in wheat allopolyploids, we generated ~317 million small RNA sequencing reads by Illumina sequencing of 25 small RNA libraries from the same plant tissues used for RNA-seq analyses (Figures 1B to 1D). After removing adapter and structural RNA sequences, we identified ~186 million 20- to 25-nucleotide small RNA sequences. Among them, ~62 million (~33%) perfectly matched the A genome and ~82 million (43.9%) perfectly matched the D genome, representing 33 million nonredundant small RNAs (Supplemental Data Set 2). Here, the D genome sequence was used to identify small RNAs derived from *Ae. tauschii* and the D subgenome of nascent allohexaploid progeny, while the A genome sequence was used to identify A subgenome–derived small RNAs in *T. turgidum* and nascent allohexaploid progeny.

In young spikes and immature seed, we found that *Ae. tauschii* small RNAs contained a higher proportion of 21-nucleotide small RNAs than those of *T. turgidum* (Supplemental Figures 10A and 10B). Numerous miRNAs were identified that were held in common in the allohexaploids and their progenitors, demonstrating that they are conserved among species. Of 154 miRNAs identified in this work, nearly all were present in the hexaploid progeny and its progenitors (Supplemental Data Set 5). The percentage of miRNAs with parental differential expression was high in seedlings (42/139, 30.2%), young spikes (57/135, 42.2%), and immature seed (25/117, 21.4%) (Supplemental Table 8), indicating that different species and tissues possess characteristic miRNA expression patterns.

A total of 32 miRNAs (32/135, 23.7%) in S3 young spikes were found to be nonadditively expressed when compared with their MPVs ($P < 0.05$, BH multiple test correction), whereas there were five in immature seed and none in seedlings (Supplemental Table 9). Among these nonadditively expressed miRNAs from the young spikes, 20 were nonadditively repressed, 75% (15/20) of which were from miRNAs showing higher expression levels in *Ae. tauschii* than in *T. turgidum* (Supplemental Tables 8 and 9). These included functionally well characterized miRNAs, such as miR169, miR172, miR319, and miR827, that are conserved between monocots and

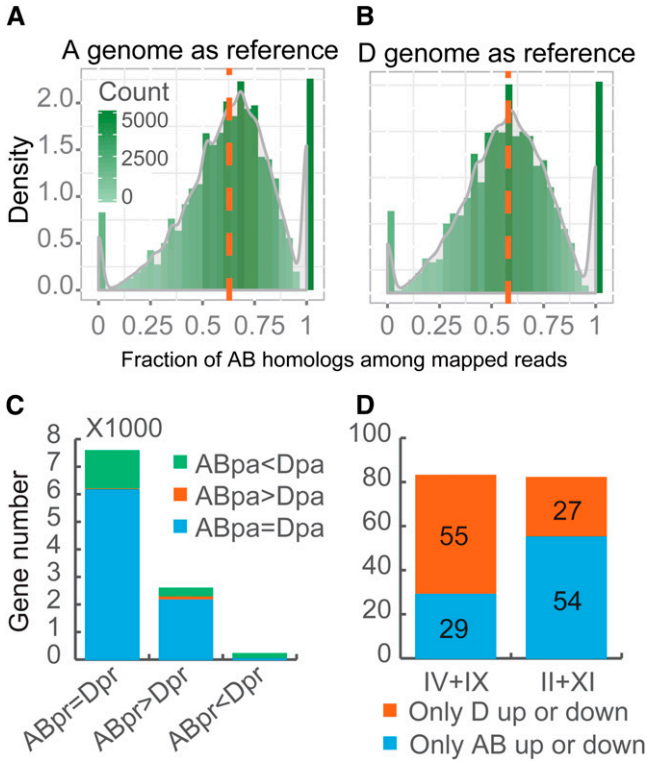


Figure 6. Biased Homoeolog Expression as Measured by SNP Mapping Depths for Young Spikes of S3 Progeny.

(A) and **(B)** Distribution of the fraction of mapped reads that carry the AB homoeolog.

(A) Reads mapped against the *T. urartu* (A) genome. Dotted line indicates the median number, which is greater than 0.5.

(B) Reads mapped against the *Ae. tauschii* (D) genome.

(C) Homoeolog expression regulation during allohexaploidization. *T. turgidum* homoeologs, AB; *Ae. tauschii* homoeolog, D; >, =, and < indicate relative expression levels between homologs in parental lines (pa) or homoeologs in progeny (pr).

(D) Dynamic regulation of AB or D homoeolog expression in nascent allohexaploid wheat and parental genome ELD. Categories II, IV, IX, and XI correspond to those in Figure 4A. up, upregulated in the progeny relative to the parental expression level; down, downregulated in the progeny relative to the parental expression level. For details, see Supplemental Table 7.

eudicots, as well as grass-specific miRNAs, such as miR5200, miR9006, and miR9009 (Figure 7A; Supplemental Table 9). miR169 and miR319 are known to be involved in response to abiotic stresses such as drought, salt, and cold (Liu et al., 2008; Zhang et al., 2009; Lv et al., 2010; Frazier et al., 2011). Among grass-specific miRNAs, nonadditive expression of miR9006 and miR9009 was further confirmed by RNA gel blot (Figure 7B), although the alteration was not as significant as shown by RNA-seq. miR9009 was previously predicted to target *R* gene analogs (RGAs) in wheat (Wei et al., 2009). The higher expression of miR9009 in *Ae. tauschii* relative to *T. turgidum* and its nonadditive repression (hence, upregulation of *R* genes) in S3 plants is especially intriguing given the enhanced powdery mildew resistance of newly synthesized wheat compared with *Ae. tauschii* (Supplemental Figure 1F). In addition, miR5200 has recently been shown to target *FLOWERING LOCUS T* (*FT*) homologs *FTL1/2* in *Brachypodium distachyon* (Wu et al., 2013). The predicted targets AEGTA05442 (Bd-*FTL2* homolog) and AEGTA18142 (Bd-*FTL1* homolog) of miR5200 were present among *Ae. tauschii* gene models (Supplemental Data Set 6). Thus, nonadditive regulation of grass-specific miRNAs may contribute to the robustness of nascent allohexaploid wheat in disease resistance and flowering time flexibility.

By contrast, 12 miRNAs were found to be nonadditively activated in S3 including well studied miRNAs: miR159, miR166, miR167, miR390, and miR396 (Figure 7B; Supplemental Table 9). Targets of these miRNAs were enriched for the GO Biological Process term "response to stress" (P = 2.80E-06, BH multiple test correction), "cell death" (P = 8.58E-06), and "flower development" (P = 4.85E-05). Two miRNAs, miR167 and miR390, were nonadditively activated in young spikes and immature seed, respectively (Supplemental Table 9), and the expression pattern of miR167 was confirmed by RNA gel blot in young spikes (Figure 7B). miR167 was predicted to target auxin response factor (ARF) genes (AEGTA09991, AEGTA09040, and AEGTA30848 for *Ae. tauschii* and TRIUR3_28948, TRIUR3_22410, and TRIUR3_32597 for *T. urartu*; Supplemental Data Set 6). In rice, the miR167-*ARF8*-*GH3.2* signal transduction pathway is involved in auxin signaling (Yang et al., 2006). Notably, a homolog of rice Os-*GH3.1* gene appeared to be nonadditively regulated in S3 (Supplemental Table 2). Therefore, it is plausible that this miR167-*ARF*-*GH3* pathway is conserved and nonadditively regulated in allohexaploid wheat.

Furthermore, we found that a marked number of miRNAs with nonadditive expression were preserved when we used miRNA expression levels in CS to compare with MPVs (Supplemental Table 10). In young spikes, for instance, 75% (15/20) and 66.7% (8/12) of nonadditively repressed and activated miRNAs were present in CS respectively, including miR167, miR172, miR319, miR390, and miR5200.

Like protein-coding genes, a high proportion of miRNA genes displayed parental ELD in nascent allohexaploid wheat (Supplemental Table 11). In young spikes, for example, 23 miRNAs (categories IV and IX: 36.5% of 63 differentially expressed miRNAs) exhibited ELD-ab (P < 0.05, BH multiple test correction), significantly more than those displaying ELD-d (6, 9.5% of categories II and XI). Similar patterns were found in seedlings (25% versus 11.4%). Moreover, among nonadditively expressed miRNAs in young spikes, 5 out of 12 (41.7%) nonadditively activated miRNAs (including miR159, miR167, and miR396) and 9 out of 20 (45%) nonadditively repressed (including miR169 and miR827) showed an ELD-ab expression pattern (Figure 7A; Supplemental Tables 9 and 12). By contrast, only two miRNAs (10%) that were nonadditively repressed belonged to the ELD-d category (miR5200 and miR9018). The predicted targets of ELD-ab miRNAs were functionally enriched for "growth" (FDR = 0.0024), while those of ELD-d miRNAs were enriched for "nucleus" (FDR = 0.00096). Interestingly, miR5200, which targets *FT*-like genes in *B. distachyon* (Wu et al., 2013), showed ELD-ab in seedlings and ELD-d in young spikes, suggestive of its tissue-specific functions in allohexaploid wheat (Supplemental Tables 12 and 13). In addition, comparable percentages of miRNAs with ELD-ab (52%, 12/32) or ELD-d (50%, 3/6) patterns were observed when CS was used as an assumed progeny (Supplemental

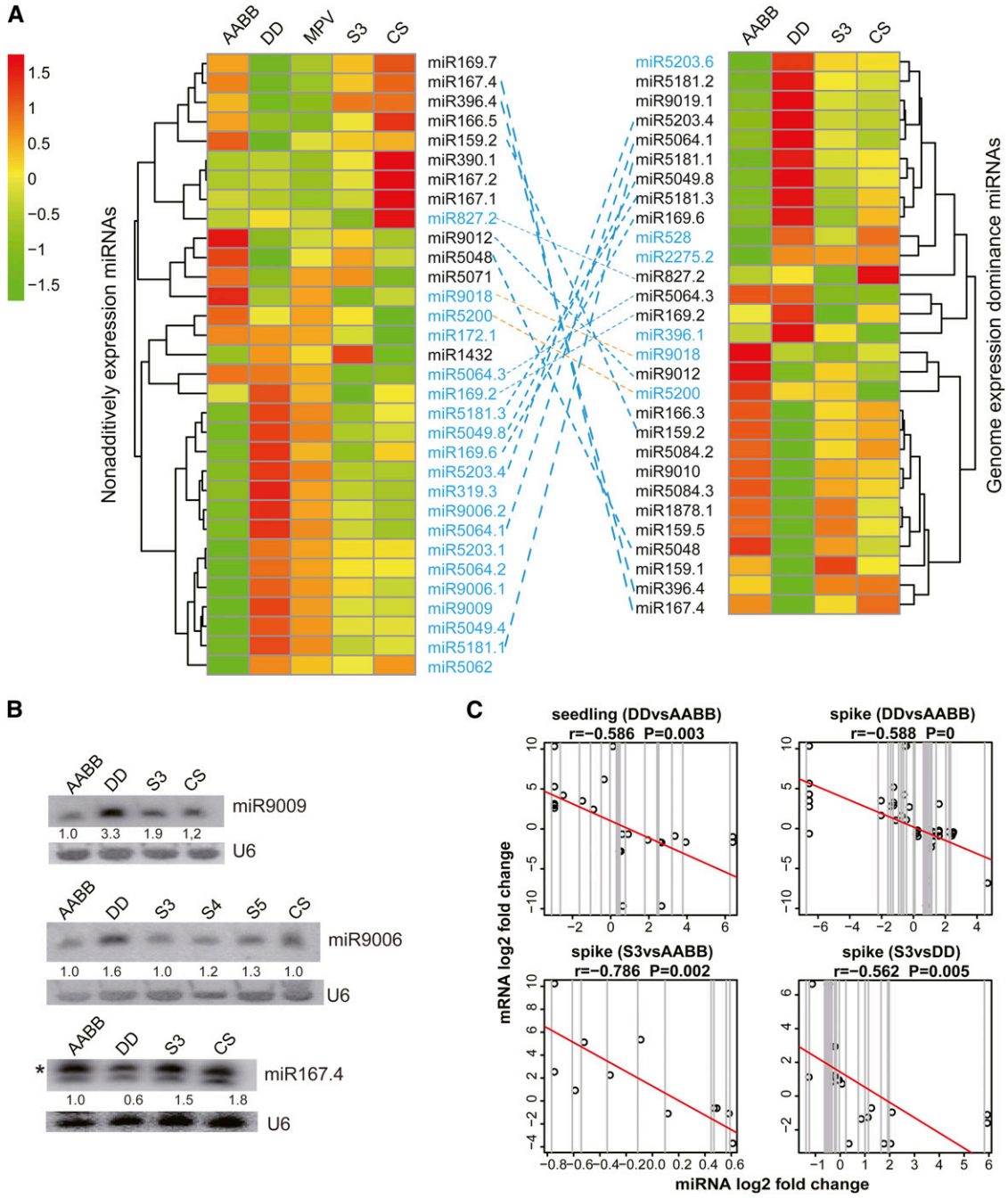


Figure 7. Effect of Allohexaploidization on Expression of miRNAs and Their Targets.

AABB, *T. turgidum*; DD, *Ae. tauschii*; S1 to S4, consecutive generation of selfed allohexaploid wheat. ELD-ab, miRNAs with ELD for *T. turgidum*; ELD-d, miRNAs with ELD for *Ae. tauschii*.

(A) Left: Hierarchical clustering of miRNAs displaying nonadditive expression. Blue letters, nonadditively repressed miRNAs; black letters, nonadditively activated miRNAs. Right: Hierarchical clustering analysis of miRNAs displaying parental ELD in young spikes. Blue dotted lines connect ELD-ab miRNAs with those nonadditively expressed miRNAs; yellow dotted lines connect ELD-d miRNAs with those nonadditively expressed miRNAs. Blue letters, ELD-d miRNAs; black letters, ELD-d miRNAs.

(B) RNA gel blot analysis of miRNAs displaying nonadditive expression in young spikes. U6 RNA is used as a loading control.

(C) Inverse correlations between $\log_2$ fold changes of a subset of differentially expressed miRNAs and those of differentially expressed targets. Multiple targets of a single miRNA are indicated by a vertical line. r, Pearson correlation efficient; P values are derived from Wilcoxon paired rank sum test.

Table 14). Overall, miRNAs, like protein-coding genes, displayed preferential ELD-ab expression patterns, which may contribute to the regulation of their target gene expression in nascent allohexaploid wheat.

We then studied the effect of differentially expressed miRNAs on the expression of their targets among the progenitors and newly formed allohexaploid wheat. Out of the 42 miRNAs showing parental differential expression in seedlings, 17 (40.5%) were found to target 22 genes that were differentially expressed between *Ae. tauschii* and *T. turgidum* (Supplemental Table 15). The miRNA $\log_2$ fold changes were negatively correlated with those of target mRNA levels (Pearson correlation, $r = -0.586$, P = 0.003; Figure 7C). A similar negative correlation was observed between 28 out of 57 (49.1%) miRNAs with parental differential expression and their 42 targets with parental differential expression in young spikes ($r = -0.588$, P = 0). For miRNAs differentially expressed between S3 progeny and the two progenitors, differentially expressed targets were found only in young spikes (Supplemental Table 16). Of 18 differentially expressed miRNAs between S3 and *T. turgidum*, 10 targeted 10 differentially expressed genes with a negative correlation between miRNA $\log_2$ fold changes and those of target mRNA levels ($r = -0.786$, P = 0.002; Figure 7C). A similar correlation ($r = -0.562$, P = 0.005) was found between 20 miRNAs and 16 target genes that were differentially expressed between S3 and *Ae. tauschii*. The correlation between miRNA and target gene expression indicates that miRNAs may directly contribute to ELD of protein coding genes in nascent allohexaploid wheat.

siRNAs as a Genetic Buffer and for Regulation of Homoeolog Expression in Nascent Allohexaploid Wheat

siRNAs, especially 24-nucleotide siRNAs that are often derived from repetitive sequences, mediate RNA-dependent DNA methylation and are important to chromatin stability maintenance and gene expression regulation in interspecific hybridization or polyploidization (Ha et al., 2009b; Kenan-Eichler et al., 2011; Lu et al., 2012). The wheat genome contains more repetitive sequences than most other plant genomes, and many of these sequences are derived from TEs. Hence, epigenetic regulation of homoeolog expression is expected during wheat allohexaploidization (Ha et al., 2009b; Kenan-Eichler et al., 2011; Lu et al., 2012).

We found that in all three tissues studied, the ratios of 24-nucleotide siRNAs among *Ae. tauschii* small RNAs were lower than the ratios among those of *T. turgidum* (Supplemental Figures 10A to 10C), while their abundance in the allohexaploid progeny appeared to be similar to that in the *T. turgidum* tetraploid progenitor (Supplemental Figures 10A to 10C). We then analyzed the small RNAs that perfectly matched the D genome sequence. In young spikes, miRNAs and gene-derived small RNAs were identified by mapping to miRBase miRNAs and gene models, while the remaining small RNAs that matched the reference genome sequences were considered to be derived from intergenic regions. We found that the intergenic regions produced the highest number of siRNAs (on average 48.9%), while genic (transcribed region) siRNAs, TE-associated siRNAs, and miRNAs occupied 25.6, 12.5, and 12.9%, respectively (Supplemental Figure 11A). Similar distribution of small RNAs was found in immature seed (Supplemental Figure 11B). We found higher percentages of TE-derived siRNAs in S3 plants than in the D genome donor *Ae. tauschii* (13.0% versus 9.8% in young spikes and 14.2% versus 9.6% in immature seed; Supplemental Figures 11A and 11B) when using the D genome as a reference, suggesting that there could be increased modification of D subgenome repetitive regions in the allohexaploid wheat. Mapping small RNAs to either the A genome or the D genome showed tissue-specific distribution of TE-derived siRNAs. For example, in young spikes, but not in immature seed and seedlings, S3 was found to have higher TE-derived small RNA content than the tetraploid progenitor, when the A genome was used as a reference (Supplemental Figure 12).

To study the effect of siRNAs on genome stability, we surveyed siRNA cluster density along the D genome scaffolds. We defined an siRNA cluster as a region containing a minimum of 10 21- to 24-nucleotide sRNA reads, each separated from the nearest neighbor by a maximum of 200 nucleotides (Ha et al., 2009b; Lu et al., 2012). We identified 79,169 clustered siRNA loci along the D genome sequence in S3 progeny. Among these loci, 61,860 (~78%) were present in all three tissues, and the remaining (17,309, 22%) showed tissue specificity (Figure 8A). Similar siRNA cluster numbers and tissue distribution ratios were found in S3 progeny using the A genome as a reference (Supplemental Figure 13). siRNA clusters were found to be highly transmittable (up to 95%) across early generations of allohexaploid wheat in young spikes, using either the A or D genome as a reference (Figures 8B and 8C). High heritability was also seen in seedlings (>80%) and immature seed (>90%; Supplemental Figures 14A to 14D). The high stability of siRNA clusters may reflect important roles for them in genome stabilization of allohexaploid wheat.

We found that ~70% of gene models colocated with TEs (henceforth, TE-associated genes) within their genic regions (including 2-kb 5′ and 3′ regions) in both the A (5940/8695, 68%) and D (10,303/14,568, 70%) genomes (Figure 8D), much higher than that in *Arabidopsis* (46%; Lu et al., 2012), indicating that wheat genes are more likely to be regulated by RNA-directed DNA methylation. Overall, the mapping densities of siRNAs in the upstream and downstream regions of TE-associated genes were much higher than in transcribed regions (Figures 8E to 8H). Comparing S3 to its progenitors, the siRNA density at D subgenome TE-associated genes in S3 was significantly higher than those in the D genome of *Ae. tauschii* in young spikes (P = 2.17e-17, Wilcoxon paired ranks sum test) and immature seed (P = 1.6e-71; Figures 8E and 8F). By contrast, no such difference was found for A genome TE-associated genes between S3 and *T. turgidum* in young spikes (P = 0.05) and immature seed (P = 0.737; Figures 8G and 8H), indicative of differential modifications of A and D subgenome TE-associated genes in the hexaploid wheat.

We further studied the effect of differentially expressed siRNAs on homoeolog expression levels of neighboring genes. In S3 young spikes, the number of D homoeologs with downregulated siRNAs was higher than that of A genome homoeologs (664 versus 150; $P < 0.001$, χ^2 test), and the same was true of those with upregulated siRNAs (1286 versus 317; $P < 0.001$, χ^2 test; Figure 9A). Similar patterns were found in immature seed (Figure 9B). The significant increase in siRNA density on D homoeologs thus indicates that D genome genes are likely subject to epigenetic repression in nascent allohexaploid wheat. GO analysis showed that D homologs with altered siRNA expression during hexaploidization were enriched for functions in "response to stress" (Figure 9C;

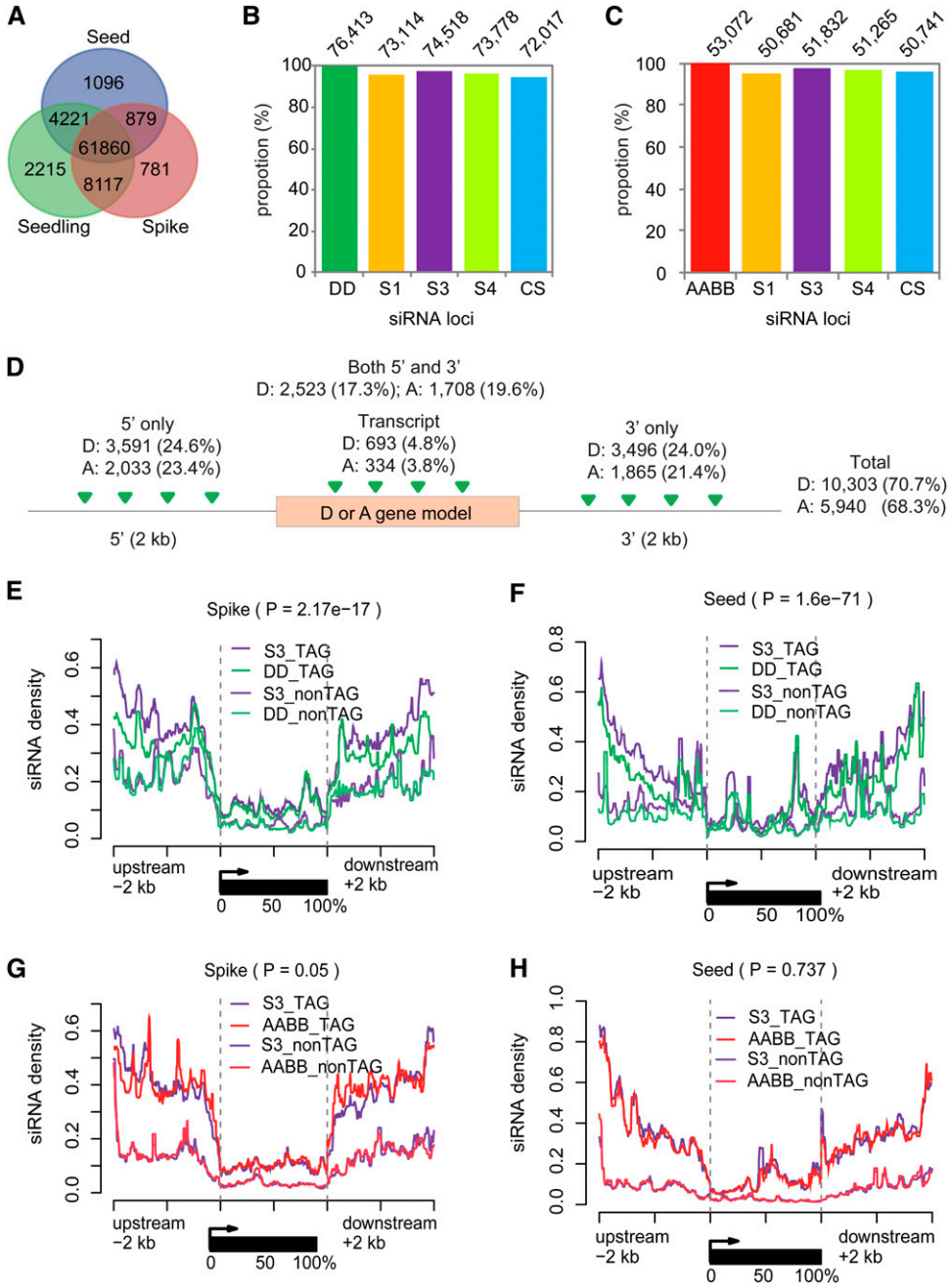


Figure 8. siRNA Expression Patterns during Wheat Allohexaploidization.

(A) Venn diagram showing tissue-specific expression of siRNA clusters in S3 plants using the D genome as reference.
(B) Clustered siRNA loci (on top of the figure) in young spikes identified using the D genome as a reference. The percentages of S1, S3, S4, and CS siRNA clusters were derived against the number of siRNA clusters from *Ae. tauschii* (DD, 76,413).
(C) Clustered siRNA loci (on top of the figure) in young spikes identified using the A genome as a reference. The percentages of S1, S3, S4, and CS siRNA clusters were derived against the number of siRNA clusters from *T. turgidum* (AABB, 53,072).
(D) Proportions of TAGs and locations of TEs (triangles) in transcribed region (gray box) and their 2-kb upstream and downstream regions (extended lines) of *Ae. tauschii* (D) and *T. urartu* (A) gene models.
(E) and **(F)** Small RNA densities (100-bp sliding window) in upstream (2 kb), transcribed, and downstream (2 kb) regions of TAGs and non-TAGs in young spikes **(E)** and immature seed **(F)** on the D genome.
(G) and **(H)** Small RNA densities (100-bp sliding window) in upstream (2 kb), transcribed, and downstream (2 kb) regions of TAGs and non-TAGs in young spikes **(G)** and immature seed **(H)** on the A genome.

FDR = 0.00208, BH multiple test correction). MapMan analysis showed that pathways for protein, RNA, signaling, stress, and transport were among the large gene groups (>50; Figure 9D). Thus, it seems that stress-related genes are intrinsic targets of epigenetic regulation that may offer instant adaptability to allohexaploid wheat, as has been reported in *Arabidopsis* (Chao et al., 2013). In addition, we found that siRNA expression changes at 33 out of 34 TE-associated genes on the D genome were negatively correlated with mRNA expression levels of neighboring genes (Supplemental Table 17; Figures 9E and 9F), indicating that small RNA–mediated epigenetic regulation may indeed play a direct role in regulating homoeolog expression in allohexaploid wheat.

DISCUSSION

Allohexaploid Wheat as a Model for Heterosis Studies

Heterosis has been recognized for centuries (Darwin, 1876). With the accumulation of genomic and epigenetic knowledge, heterosis now is considered to arise from allelic interactions between parental genomes, leading to reprogramming of gene expression patterns favoring enhanced growth, stress tolerance, and fitness in the hybrids. Many studies have used diploids such as maize and rice to decipher the possible molecular basis for heterosis. More recently, polyploid systems that allow experiments not possible in diploids have emerged as model systems to study heterosis (Washburn and Birchler, 2014).

The ubiquity of polyploid plants is strongly suggestive of adaptive advantages to polyploidization in evolution (Comai, 2005; Wood et al., 2009; Mayrose et al., 2011). Common wheat experienced two rounds of polyploidization with the increase of ploidy correlating with enhanced growth vigor and improved adaptation, displaying progressive heterosis. Hexaploid wheat varieties have superior grain quality and adaptability compared with their progenitors. The molecular basis of such a consequence is presumed to have been established in the initial wheat hexaploids and could be elucidated by studying newly synthesized allohexaploid wheat. More importantly, nascent allohexaploid wheats are at their heterotic stage and are good models to study heterosis due to their genetic stability and fixed heterozygosity and hybrid vigor. In addition, synthetic wheat can cross with modern cultivars without obvious reproductive barriers, indicating that there is almost instant genetic stabilization (Mestiri et al., 2010). Meanwhile, they are important resources for introgression of needed traits from wild relatives of wheat (Yang et al., 2009). For all of these reasons, many studies have examined gene expression patterns in nascent and stable allohexaploid wheats (He et al., 2003; Bottley et al., 2006; Pumphrey et al., 2009; Akhunova et al., 2010; Chagué et al., 2010; Qi et al., 2012; Chelaifa et al., 2013; Zhang et al., 2013a). The advent of next-generation sequencing technology and the availability of draft sequences of the A and D genomes (Jia et al., 2013; Ling et al., 2013) provided unprecedented opportunities to examine this important process.

Parental Expression Level Dominance and Heterosis in Allohexaploid Wheat

Comparative genomics studies showed that polyploidization is followed by genome-wide diploidization, during which duplicated genomes experienced differential gene loss or biased fractionation, leading to one subgenome having more genes retained as well as exhibiting higher expression and hence genome dominance (Schnable et al., 2011; Cheng et al., 2012; Pont et al., 2013; Garsmeur et al., 2014). Common wheat experienced allotetraploidization at around 0.5 million years ago and allohexaploidization ~8000 years ago (Huang et al., 2002) and is a good model for studying genetic interactions among the three homoeologous genomes. Despite its relatively short history, genomic changes that may lead to genome dominance appear to have occurred in natural hexaploid wheat, as indicated by conventional genetic evidence (Feldman et al., 2012) and more recent genomic data (Akhunov et al., 2013; Pont et al., 2013).

Thus, common wheat differs from resynthesized wheat, as supported by our expression correlation dendrogram where CS was revealed as an outlier. In other words, newly synthesized allohexaploid wheat and their early generations may represent an early genetic stage of common wheat that may potentially confer hybrid vigor. Compared with natural hexaploid wheat, homoeologs in nascent allohexaploid wheat should represent their initial expression state, possibly maintained by epigenetic mechanisms as have been shown in *Arabidopsis* (Wang et al., 2006). However, we found only a small percentage of nonadditively expressed genes among those expressed across all three tissues, consistent with previous microarray results (Chelaifa et al., 2013). Despite this, nonadditive expression of genes has been readily detected in different hexaploid wheat varieties, stable or nascent, and on different study platforms (Chagué et al., 2010; Qi et al., 2012), suggesting that they may indeed contribute to heterosis in nascent allohexaploid wheat. We speculate that the presence of nonadditively activated expression of expansins and auxin pathway-related genes in young spikes of nascent allohexaploid progeny might correlate with the elongation of spike cells.

By contrast, like in other allopolyploid plants (Rapp et al., 2009; Flagel and Wendel, 2010; Grover et al., 2012), a marked number of genes displayed parental ELD in nascent allohexaploid wheat. Our analysis showed that, aside from genes exhibiting additive expression (roughly half of differentially expressed genes among the progenitors and/or the progeny), the remainder of differentially expressed genes in nascent allohexaploid wheat show predominantly parental ELD, with nonadditive expression in the minority. The maintenance of ELD genes in natural hexaploid wheat (CS) suggests that these genes may have indispensable roles. Our analysis further showed that, in allohexaploid wheat, more genes displayed ELD toward their tetraploid progenitor. These so-called ELD-ab genes were enriched with distinct GO terms when compared with ELD-d genes and include important transcription factors such as key floral development genes *PI*, *AP3*, and *AG* homologs, which may contribute to spike and flower organ renovation in allohexaploid wheat. ELD-ab genes may also underlie the phenotypic similarity between S3 plants and the tetraploid progenitor, especially for plant height, spike shape, and grain length. By contrast, adaptation-related genes appear to display ELD-d and may contribute to stress responses and photoperiod adaptability. This category also comprises a number of genes encoding RGAs, chitinase, and HSP90 that may be involved in

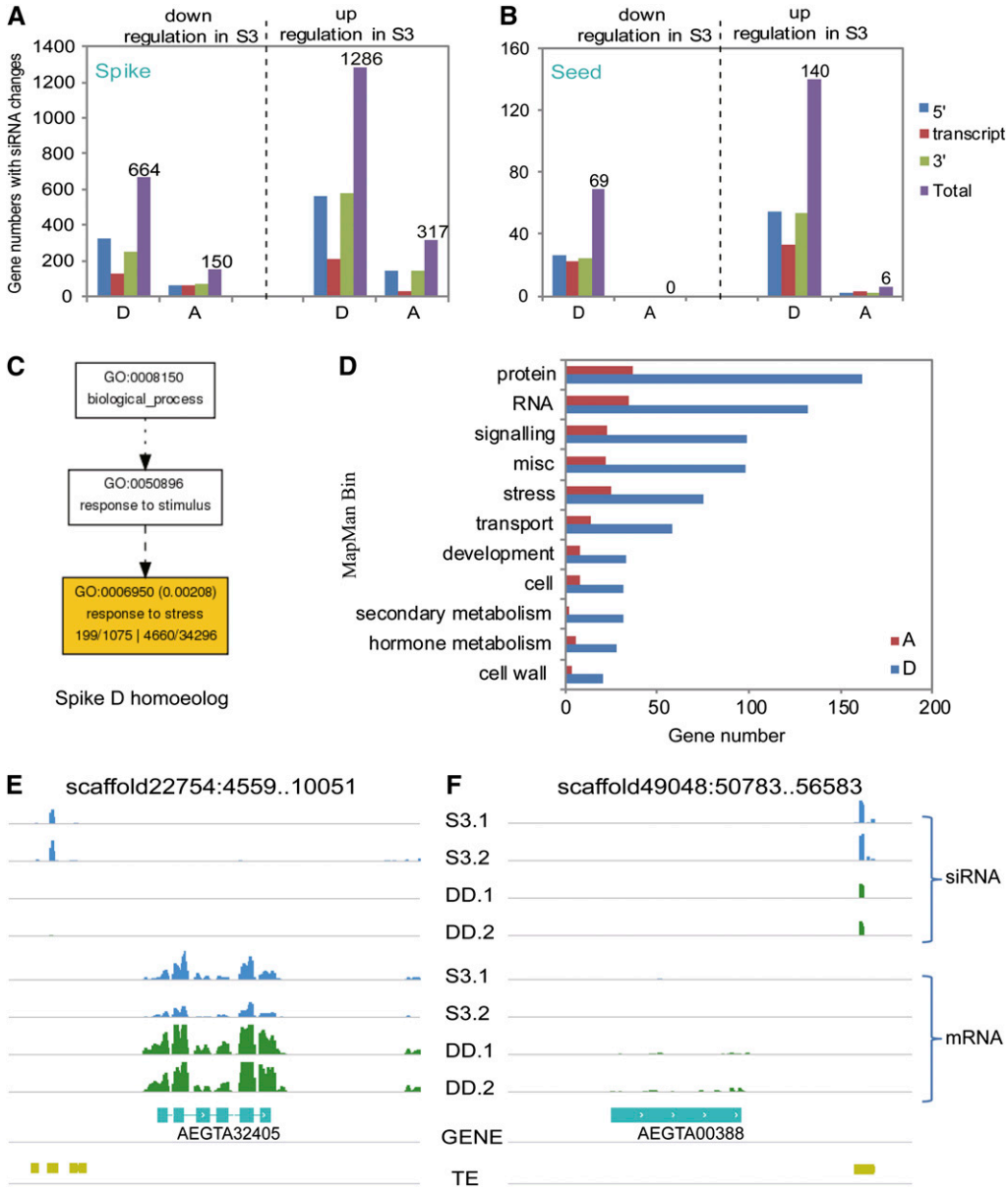


Figure 9. The Effect of Differential Expression of siRNAs on Homoeolog Expression Levels.

(A) and **(B)** Numbers of D homoeologs with siRNAs differentially expressed between S3 and *Ae. tauschii* (DD) or A homoeologs between S3 and *T. turgidum* (AABB) in young spikes **(A)** and immature seed **(B)**.

(C) GO enrichment analysis of D homoeologs in young spikes with differentially expressed siRNAs between S3 and *Ae. tauschii*.

(D) MapMan pathway analysis in young spikes of A homoeologs between S3 and *T. turgidum* or D homoeologs between S3 and *Ae. tauschii* with differentially expressed siRNAs.

(E) and **(F)** Two examples of downregulation of siRNA expression and concerted upregulation of neighboring D homoeolog AEGTA32405 **(E)** and AEGTA00388 **(F)** in young spikes of S3 plants compared with their expression in the diploid progenitor *Ae. tauschii* (DD).

biotic stress responses. Additionally, an *AKT1* homolog might confer salt tolerance to allohexaploid wheat and *LHY* and *CO* homologs may contribute to more flexibility in flowering condition requirements for hexaploid wheat. These data indicate that ELD-d genes are possible major contributors to key adaptation traits such as stress tolerance and flexibility in flowering conditions. Together, the merging of characteristic expression patterns from the two parental lines in the progeny may underlie the enhanced fitness and adaptability of hexaploid wheat.

miRNAs Are Sensitive to Polyploidization and Are Important for Growth Vigor and Adaptation

In *Arabidopsis* interspecific hybrids and allopolyploids, variation in miRNA expression causes nonadditive expression of target

genes that may affect growth vigor and adaptation (Chen, 2007; Ha et al., 2009b; Ng et al., 2012; Chen, 2013). Similarly, up to 25% of studied miRNAs were nonadditively expressed in nascent allohexaploid wheat. The observation of a substantial number of miRNAs displaying nonadditive expression indicates that, like in *Arabidopsis*, miRNAs may play important roles in heterosis for nascent allohexaploid wheat. Despite this, it was difficult to detect nonadditively expressed target genes, possibly due to the limitation of sequence quality of gene models from the draft genomes. However, we did detect negative correlations in fold changes between differentially expressed miRNAs and target gene expression among the progeny and its progenitors in nascent allohexaploid wheat, suggesting direct roles of miRNAs in target gene expression regulation. More interestingly, a marked number of nonadditively expressed miRNAs also displayed parental ELD in the polyploid progeny, demonstrating that similar regulatory mechanisms may apply to miRNAs as protein-coding genes. Furthermore, targets of nonadditively repressed miRNAs (e.g., miR169 and miR319) have been shown in other species to be involved in drought, salt, and cold responses (Frazier et al., 2011; Khraiwesh et al., 2012), while nonadditive expression of grass-specific miRNAs (such as miR5200, miR9006, and miR9009) may possibly contribute to biotic resistance and flowering control via regulating their targets, such as *RGAs* and *FT*. Furthermore, nonadditive expression of miR167 and its putative targets (including several auxin response factors) in young spikes suggests that the miR167-*ARF8*-*GH3.2* auxin signaling pathway might be relevant to spike growth vigor in nascent allohexaploid wheat, a possibility that requires detailed exploration. The sensitivity of miRNA expression suggests that miRNAs are possible

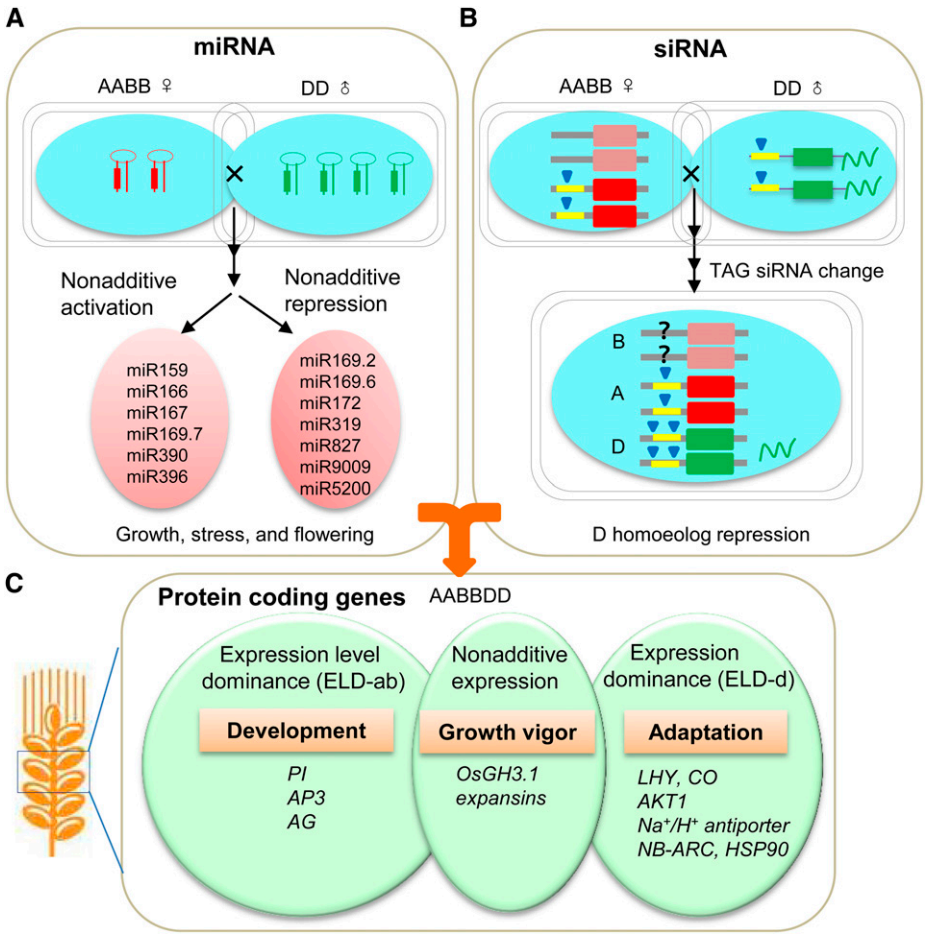


Figure 10. A Small RNA–Dominated Model for Homoeolog Expression Regulation during Wheat Allohexaploidization.

AABB, *T. turgidum*; DD, *Ae. tauschii*; AABBDD, nascent allohexaploid progeny; ELD-ab, genes with expression level similar to that in *T. turgidum*; ELD-d, genes with expression level similar to that in *Ae. tauschii*.

(A) miRNA regulation module. Nonadditively expressed miRNAs target genes involved in growth, stress, and flowering.

(B) siRNA regulation module. Note increased siRNA expression on the D subgenome and reduced transcript production (wavy lines). Rectangles represent genes; yellow lines represent repeat sequences; blue triangles, siRNAs. Question marks indicate unknown conditions in the B genome.

(C) Expression patterns of protein-coding genes, their distinct functions (development, growth vigor, and adaptation), and key representative genes in nascent allohexaploid wheat.

key regulators that are critical for growth vigor and adaptation in nascent allohexaploid wheat.

Genetic Buffering of siRNAs and Homoeolog Expression Regulation in Nascent Hexaploid Wheat

siRNA-mediated epigenetic mechanisms are shown to be involved in the maintenance of genome stability of interspecific hybrids and allopolyploids (Ha et al., 2009b; Kenan-Eichler et al., 2011; Greaves et al., 2012). Nascent allohexaploid wheats are considered to maintain nearly intact D and AB homoeologous genomes upon polyploidization; hence, epigenetic regulation of gene expression is expected. Like in other species, siRNAs may play important roles in genome stability when polyploidization reactivates retrotransposons in nascent allohexaploid wheat, a notion that is supported by the high transmission rates for siRNA clusters among early generations of nascent allohexaploid wheat (Kashkush et al., 2003). However, this observation differs from that reported by Kenan-Eichler et al. (2011), who found that the abundance of TE-associated siRNAs was strongly decreased upon polyploidization. Such differences may result from different reference sequences used in the two studies.

An accurate assessment of TEs and TE-associated siRNAs in intergenic regions may require higher quality genome assemblies than currently available. However, the majority of the coding regions in the A and D genomes, including 5′ and 3′ untranslated regions and neighboring regions, are well assembled (Jia et al., 2013; Ling et al., 2013). Compared with the long intergenic TEs, TE sequences that colocate with genes are usually short miniature inverted repeat elements or residual sequences of retrotransposon long terminal repeats, or small long terminal repeat retrotransposons (Feschotte et al., 2002; Gao et al., 2012; Han et al., 2013; Wei et al., 2014). We reanalyzed the composition of TEs in our TE-associated genes (TAGs) and found that most TAG TEs are indeed quite small, <300 bp (Supplemental Figure 15). Thus, we have a relatively high degree of confidence in our assessment of siRNAs associated with TAG TEs. Considering that more than 70% of wheat genes have TEs in their neighboring regions, siRNA-mediated DNA methylation may also contribute to interactive homoeolog expression through associated TE sequences. We found in young spikes and immature seed of S3 plants higher siRNA density at upstream or downstream regions of TE-associated D homologs than at those of A homologs, which may account for biased repression of the D homoeologs in nascent allohexaploid progeny. The negative correlation between siRNA density changes and expression levels of neighboring genes further supports the role of small RNA–mediated epigenetic regulation on homoeolog expression in nascent hexaploid wheat.

SNP-assisted analysis of homoeolog expression level showed more AB homoeologs with higher (A+B combined) expression than D homoeologs in nascent allohexaploid wheat. This seems to conflict with the results from Pont et al. (2013), who considered the D subgenome to be supradominant over the AB subgenome. However, it is not clear whether D homoeologs utilized in their analysis also display higher expression levels. Additional expression data are needed to reconcile this contradiction. Interestingly, we found that biased AB homoeolog expression was mostly achieved during allopolyploidization by dynamic regulation of both AB and D homoeologs, instead of being simply inherited from progenitors. Thus, the merging of the AB genomes with the D genomes in allohexaploid wheat brought in further regulation on both AB and D homoeologs, adding an extra layer of regulatory plasticity.

From Nascent Hexaploid Wheat to Modern Cultivars

Allohexaploid wheat combines the AB genomes from tetraploid wheat with the D genome from *Ae. tauschii*, resulting in the union of genomes from varieties previously adapted to different environments and thus providing the potential for further adaptation to a wider range of growth environments. Based on our data, we propose a gene regulatory scheme for wheat allohexaploidization (Figure 10). We divide the process into three modules. In the miRNA module, nonadditive expression of important miRNAs (both conserved and grass specific) may exert effects on their targets involved in plant growth, stress, and flowering. In the siRNA module, increased siRNA density in S3 may result in biased downregulation of D homoeologs, which involves dynamic regulation of both D and AB homoeologs. In the protein-coding gene module, genes with expression levels similar to those of *Ae. tauschii* (ELD-d) may contribute to flowering-related adaptations (speculatively based on *LHY* and *CO*), disease resistance (based on *R* genes), and salt tolerance (based on *AKT1*). Meanwhile, genes with expression levels similar to those of *T. turgidum* (*PI*, *AP3*, and *AG*) and the nonadditively activated genes may contribute to spike development and growth vigor. Overall, the molecular underpinnings established during the early allopolyploidization events laid the groundwork for the successful advent of common wheat.

METHODS

Plant Materials

The tetraploid wheat *Triticum turgidum* ssp *dicoccon* accession PI94655 (AABB, 2n=4x=28) and diploid goat grass *Aegilops tauschii* ssp *strangulata* (DD, 2n=2x=14) accession AS2404 were used as parental lines to generate allohexaploid wheat (Zhang et al., 2010). *T. turgidum* was used as a maternal parent and pollinated with *Ae. tauschii* anthers. S1 seeds were obtained from triploid F1 hybrids (ABD) after spontaneous chromosome doubling and were grown as the first generation of allohexaploid plants (AABBDD) (Zhang et al., 2010). Spontaneous chromosome doubling resulted from the union of unreduced gametes (Zhang et al., 2007). No embryo rescue procedure or hormone treatment was applied for the production of the triploid F1 hybrid. S1 euploid plants with 2n=42, usually paired as 21 bivalents at meiotic metaphase I, were selected by cytogenetic observation and self-pollinated to produce the S2, S3, and S4 generations. Genomic in situ hybridization and fluorescence in situ hybridization were used to analyze their chromosomes constitutions, as previously described (Hao et al., 2013).

Sample Preparation and Sequencing Library Construction

Seedlings were grown in a controlled environment as described (Sparks and Jones, 2009), and 7-d-old plants were harvested. Spikes at heading stage and immature seeds of plants 11 DAF were collected from a pool of 10 to 12 plants. All materials were stored in liquid nitrogen until RNA extraction. Tissues were independently collected two times creating two biological duplicates for the maternal parent PI94655, the paternal parent AS2404, the S3 generation plants, and Chinese Spring. Other samples (S1, S2, and S4) were harvested once. mRNA-seq and sRNA-seq libraries were constructed according to the

manuals as provided by Illumina and were sequenced on HiSeq2000. In brief, total RNAs were isolated using TRIzol reagent (Invitrogen) and were then treated with RNase-free DNase I (New England Biolabs) to remove any contaminating genomic DNA. mRNA was extracted using Dynabeads oligo (dT) (Dynal; Invitrogen). Double-stranded cDNAs were synthesized using reverse transcriptase (Superscript II; Invitrogen) and random hexamer primers. The cDNAs were fragmented by nebulization, and the standard Illumina protocol was followed thereafter to develop mRNA-seq libraries.

Alignment of RNA-seq Reads to the Reference Genomes

Reference genome sequences and annotations were downloaded from the wheat genome website (A genome, ftp://climb.genomics.cn/pub/10.5524/100001_101000/100050/A/Assembly/; D genome, ftp://climb.genomics.cn/pub/10.5524/100001_101000/100054/D/Assembly/). Reads were aligned to each genome using the default parameters for TopHat (version 2.09) (Trapnell et al., 2009; Kim et al., 2013), which provides sensitive and accurate alignment results, even for highly repetitive genomes. Reads were also aligned to the unannotated Chinese Spring gene fragments that have been used as reference for calling homoeologous SNPs in Chinese Spring and assignment to A, B, and D subgenomes (downloaded from ftp://ftpmips.helmholtz-muenchen.de/plants/wheat/UK_454/SNPs/) using the Burrows-Wheeler Alignment tool software (Li et al., 2010). The RSeQC v2.3.6 software package (Wang et al., 2012) was used to calculate read-mapping statistics for each sample from BAM files generated from SAMtools (Li et al., 2009).

Annotation of Orthologous Genomic Regions

The A genome gene model was aligned against the genomic sequence of the D draft genome sequence with a BLAST (Altschul et al., 1997) cutoff E-value of e-10 to identify orthologous genomic regions between two reference genomes. The annotated A genome genes in the D genome located in orthologous genomic regions and showing sequence identity of more than 90% with the D gene model were considered to be the homologous genes between A and D.

Quantification and Differential Gene Expression of RNA-seq Data Sets

HTSeq software (www-huber.embl.de/users/anders/HTSeq/) was used to count the read numbers mapped to each orthologous genomic region. For all comparisons, read counts were normalized to the aligned RPKM (Mortazavi et al., 2008) to obtain the relative levels of expression. Differential expression analysis was performed using R packages of DESeq (Anders and Huber, 2010) for comparisons among samples with two biological replicates (the two progenitors, S3, and CS) or DEGseq (Wang et al., 2010) for any of the comparisons without biological replicates (S1, S2, and S4). For DEGseq package–based differential expression analysis, we followed the "MA-plot-based method with random sampling model," which can be used for simulating technical variation. Due to the lack of biological replicate for S1, S2, and S4 RNA-seq samples, fold changes >2 between compared samples were considered as simulated biological variation in the DEGseq analysis. A similar analytical strategy has been adopted previously (He et al., 2013).

Dendrogram of samples was generated using the function flashClust and the Pearson correlation between biological replicates was calculated using the function cor in R based on the RPKM values. The Venn diagrams in this study were prepared using the function vennDiagram in R based on the gene list for each tissue type.

Homoeolog-Specific Bias Detection

Picard-tools (http://picard.sourceforge.net/) was used to preprocess the alignment results for each sample, and GATK2 software (McKenna et al., 2010) was used to perform SNP calling. SNPs were identified by comparing sequencing reads of all homologous genes between AB and D with the following criteria: all reads uniquely match both parental genomes; all reads from one parent contain the same nucleotide at the SNP position and differ from that at the same position of another parent. AB and D homoeolog-specific bias in hexaploids was identified by determining whether there were significant deviations from the binomial test (i.e., the allele ratio in hybrids deviated from 0.5) of parental alleles defined by SNPs as described previously (Wang et al., 2012).

Small RNA Library Construction and Sequencing

To determine small RNA populations in wheat hexaploids and their progenitors, we made 25 small RNA libraries from 7-d-old seedlings, spikes at heading date stage, and immature seeds at 11 DAF of the following plants: *T. turgidum*, *Ae. tauschii*, the first self-pollinated generation (S1), S2, S3, S4, and Chinese Spring. Small RNAs were prepared using a previously published protocol (Kasschau et al., 2007) and subjected to high-throughput sequencing.

Small RNA Classification

After trimming adaptor sequences at the 5′ and 3′ ends of sequenced reads, the cellular structural RNAs, such as rRNAs, snoRNAs, and snRNAs, were removed using in-house Python scripts. Clean reads of 20 to 25 nucleotides were aligned to draft genomic sequences of *Ae. tauschii* or *T. urartu* by Bowtie software (0.12.9 release) with parameter settings for perfect match (Supplemental Table 2). sRNA-seq reads were then compared with plant mature miRNA sequences downloaded from miRBase (Release20, http://www.mirbase.org/) (Kozomara and Griffiths-Jones, 2011) and wheat miRNAs characterized by Jia et al. (2013) and Wei et al. (2009). Since there is no complete wheat miRNA data set in the miRBase, we provided additional serial names for miRNAs whose number has been taken by other species in the miRBase (Supplemental Data Set 5). The remaining sRNA reads were then used to identify siRNA clusters. An siRNA cluster was defined as a region containing a minimum of 10 sRNA reads covered, each separated from the nearest neighbor by a maximum of 200 nucleotides. Reads that were aligned to repetitive elements identified by RepeatMasker (Chen, 2004) using Repbase as the reference library were classified as TE-associated sRNAs.

Differential Expression Analysis of Small RNAs

Differentially expressed miRNAs or siRNAs clusters between tissues or among hexaploids and parents were identified using DESeq software (Anders and Huber, 2010). For all comparisons, the resulting P values were adjusted using Benjamini and Hochberg's approach for controlling the false discovery rate. miRNA and siRNAs clusters with an adjusted P value < 0.05 were assigned as differentially expressed.

miRNA Target Prediction

Prediction of miRNA target genes of was performed by psRobot_tar scripts in psRobot (Wu et al., 2012).

GO and MapMan Enrichment Analysis

GO classification of each gene model was performed using BLAST software. GO enrichment analysis was implemented by the GOseq R package in which gene length bias was corrected. AgriGO, a Web-based tool and database for gene ontology analysis (http://bioinfo.cau.edu.cn/agriGO/) (Mestiri et al., 2010) was also used in this study. MapMan (http://mapman.gabipd.org/) was used to identify processes and pathways of specific gene sets (Sreenivasulu et al., 2008). GO terms with corrected FDR < 0.05 were considered significantly enriched.

miRNA Gel Blot Analysis

Total RNA was isolated from spikes using TRIzol (Invitrogen). Small RNA (20 μg) was separated in 15% polyacrylamide gels with 7 M urea and blotted on Hybond N^+ membranes (Amersham). Probes of 21- to 22-mer DNA oligonucleotides corresponding to the antisense strand of miRNAs were labeled with [γ-^{32}P]ATP by T4 polynucleotide kinase (New England Biolabs). RNA gel blot analysis was performed using a protocol described previously (Lu et al., 2012). The spliceosomal RNA U6 was used as the loading control.

Accession Numbers

The small RNA and RNA-seq data are deposited in the Gene Expression Omnibus with accession numbers GSM1364784 to GSM1364808 and GSM1364759 to GSM1364783, respectively.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. The Growth Vigor and Morphological Resemblance of Nascent Allohexaploid Wheat to the Cultivated Tetraploid Parent *T. turgidum*.

Supplemental Figure 2. RNA-seq Read Mapping Counts against the Three Reference Genomes.

Supplemental Figure 3. Comparative Analysis of Reference Gene Models from *T. urartu* (AA), *Ae. tauschii* (DD), and Chinese Spring (CS).

Supplemental Figure 4. Proportions of Genes with Different Expression Levels in the Three Tissues of Nascent Allohexaploid Wheat, Their Progenitors, and Chinese Spring.

Supplemental Figure 5. Expression Level Distribution of Genome-specific and Homologous Genes between A and D Genomes.

Supplemental Figure 6. Functional Categories of Genes Showing Tissue-Specific Expression.

Supplemental Figure 7. Nonadditively Expressed Genes in Nascent Allohexaploid Wheat in Seedlings and Immature Seed.

Supplemental Figure 8. Transgressive Inheritance of Nonadditively Expressed Genes in Young Spikes in Nascent Allohexaploid Wheat (S1, S3, and S4) and Their Preservation in Chinese Spring (CS).

Supplemental Figure 9. Higher AB Homoeolog Expression Shown by SNP Density in Young Spikes of S3 Progeny Using the Chinese Spring Genome as a Reference.

Supplemental Figure 10. Size Distribution of Small RNAs in Nascent Allohexaploid Wheat, the Progenitors, and Chinese Spring.

Supplemental Figure 11. Classification and Distribution of 20- to 25-Nucleotide Small RNAs in Nascent Allohexaploid Wheat, the Progenitors, and Chinese Spring Using the D Genome as a Reference.

Supplemental Figure 12. Classification and Distribution of 20- to 25-Nucleotide Small RNAs in Nascent Allohexaploid Wheat, the Progenitors, and Chinese Spring Using the A Genome as a Reference.

Supplemental Figure 13. Venn Diagram Showing the Tissue-Specific Expression of siRNA Clusters in S3 Using the *T. urartu* (AA) Genome as a Reference.

Supplemental Figure 14. Transgressive Inheritance of siRNA Clusters among Early Generations of Nascent Allohexaploid Wheat.

Supplemental Figure 15. Length Distribution of TEs Colocated with Genes (TAGs).

Supplemental Table 1. Differentially Expressed Genes among Nascent Allohexaploid Wheat and Its Progenitors

Supplemental Table 2. Nonadditively Expressed Genes in GO Term "Regulation of Cell Size" in Young Spikes of S3 Plants.

Supplemental Table 3. Transgressive Heritance of Nonadditively Expressed Genes in S1, S3, and S4 Young Spikes, and Their Preservation in CS.

Supplemental Table 4. The 12 Differential Expression States of Genes among S3 Progeny and Its Progenitors and Their Preservation in Chinese Spring.

Supplemental Table 5. List of Genes with ELD That Were Assigned to the MapMan "Transport" Bin.

Supplemental Table 6. SNP-Tagged Homoeolog Expression Patterns in Young Spikes of S3 Nascent Allohexaploid Wheat and Comparison with Those in its Progenitors

Supplemental Table 7. Contribution of Homoeolog Expression Regulation to ELD Genes in Nascent Allohexaploid Progeny.

Supplemental Table 8. Differential Expression of miRNAs in Progenitors and Allohexaploid Progeny.

Supplemental Table 9. Nonadditively Expressed miRNAs in S3 Allohexaploid Wheat.

Supplemental Table 10. Nonadditively Expressed miRNAs that Were Present in Young Spikes of S3 and Preserved in CS.

Supplemental Table 11. The 12 Differential Expression Bins of miRNAs among S3 Progeny and Its Progenitors and Their Preservation in Natural Hexaploid Chinese Spring.

Supplemental Table 12. List of miRNAs with ELD-ab and ELD-d in Young Spikes of S3 Allohexaploid Plants.

Supplemental Table 13. List of miRNAs with ELD-ab and ELD-d in Seedlings of S3 Allohexaploid Progeny.

Supplemental Table 14. List of miRNAs with ELD-ab and ELD-d in Young Spikes of S3 Allohexaploid Plants and Their Preservation in CS.

Supplemental Table 15. Negative Correlation between Log_2 Fold Changes of Differentially Expressed miRNAs and Differentially Expressed Target Genes between the Two Progenitors (DD: AABB).

Supplemental Table 16. Negative Correlation between Log_2 Fold Changes of Differentially Expressed miRNAs and Differentially Expressed Target Genes in Young Spikes between S3 and the Two Progenitors.

Supplemental Table 17. Negative Correlation between Log_2 Fold Changes of Differentially Expressed siRNAs and Differentially Expressed D Homoeologs in Young Spikes of *Ae. tauschii* (DD) and S3 Progeny.

Supplemental Data Set 1. RNA-seq Data Statistics and Genome Mapping Rates.

Supplemental Data Set 2. Small RNA Sequence Statistics and Genome Mapping Information.

Supplemental Data Set 3. List of Nonadditively Expressed Genes in Seedlings, Young Spikes, and Immature Seed.

Supplemental Data Set 4. List of 12 Bins of Differentially Expressed Genes Classified according to Yoo et al. (2013).

Supplemental Data Set 5. Mature miRNAs Used in This Study and Their Expression Levels in Seedlings, Spikes, and Seed.

Supplemental Data Set 6. Targets of Nonadditively Expressed miRNAs Predicted Using D or A Gene Models as References.

ACKNOWLEDGMENTS

We thank Boulos Chalhoub of INRA, France for critical reading the article. We also thank Jerome Salse, Xiaowu Wang, Johnathan Wendel, and Jizeng Jia for constructive communications and the two anonymous reviewers for helpful suggestions. This work is supported in part by the National High-Tech Project ("863," 2011AA100104, 2012AA10A308, and 2011AA100103), the NSFC (91331117 and 31071420), and the National Key Program on Transgenic Research (2013ZX08009-001).

AUTHOR CONTRIBUTIONS

L.M., D.L., and A.L. designed the experiments. X.Z., M.H., S.G., and L.Z. performed the experiments. A.L., J.W., J.Y., X.J., J.Y.W., L.Y., X.S., R.Z., L.W., and Y.Z. analyzed the data. A.L. and L.M. wrote the article. L.M., D.L., and A.L. contributed equally to this work.

Received February 20, 2014; revised March 28, 2014; accepted April 28, 2014; published May 16, 2014.

REFERENCES

Akhunov, E.D., et al. (2013). Comparative analysis of syntenic genes in grass genomes reveals accelerated rates of gene structure and coding sequence evolution in polyploid wheat. Plant Physiol. **161:** 252–265.

Akhunova, A.R., Matniyazov, R.T., Liang, H., and Akhunov, E.D. (2010). Homoeolog-specific transcriptional bias in allopolyploid wheat. BMC Genomics **11:** 505.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. **11:** R106.

Arnaud, D., Chelaifa, H., Jahier, J., and Chalhoub, B. (2013). Re-programming of gene expression in the genetically stable bread allo-hexaploid wheat. In Polyploid and Hybrid Genomics, Z.J. Chen and J.A. Birchler, eds (Hoboken, NJ: John Wiley & Sons), pp. 195–211.

Bingham, E.T., Groose, R.W., Woodfield, D.R., and Kidwell, K.K. (1994). Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. Crop Sci. **34:** 823–829.

Bottley, A., Xia, G.M., and Koebner, R.M.D. (2006). Homoeologous gene silencing in hexaploid wheat. Plant J. **47:** 897–906.

Brenchley, R., et al. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature **491:** 705–710.

Chagué, V., Just, J., Mestiri, I., Balzergue, S., Tanguy, A.M., Huneau, C., Huteau, V., Belcram, H., Coriton, O., Jahier, J., and Chalhoub, B. (2010). Genome-wide gene expression changes in genetically stable synthetic and natural wheat allohexaploids. New Phytol. **187:** 1181–1194.

Chao, D.Y., Dilkes, B., Luo, H., Douglas, A., Yakubova, E., Lahner, B., and Salt, D.E. (2013). Polyploids exhibit higher potassium uptake and salinity tolerance in Arabidopsis. Science **341:** 658–659.

Chelaifa, H., et al. (2013). Prevalence of gene expression additivity in genetically stable wheat allohexaploids. New Phytol. **197:** 730–736.

Chen, N. (2004). Using RepeatMasker to identify repetitive elements in genomic sequences. Curr. Protoc. Bioinformatics **Chapter 4:** 10.

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.

Cheng, F., Wu, J., Fang, L., Sun, S.L., Liu, B., Lin, K., Bonnema, G., and Wang, X.W. (2012). Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. PLoS ONE **7:** e36442.

Colmer, T.D., Flowers, T.J., and Munns, R. (2006). Use of wild relatives to improve salt tolerance in wheat. J. Exp. Bot. **57:** 1059–1078.

Comai, L. (2005). The advantages and disadvantages of being polyploid. Nat. Rev. Genet. **6:** 836–846.

Comai, L., Tyagi, A.P., Winter, K., Holmes-Davis, R., Reynolds, S.H., Stevens, Y., and Byers, B. (2000). Phenotypic instability and rapid gene silencing in newly formed arabidopsis allotetraploids. Plant Cell **12:** 1551–1568.

Darwin, C.R. (1876). The Effects of Cross and Self Fertilisation in the Vegetable Kingdom. (London: John Murray).

Degner, J.F., Marioni, J.C., Pai, A.A., Pickrell, J.K., Nkadori, E., Gilad, Y., and Pritchard, J.K. (2009). Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. Bioinformatics **25:** 3207–3212.

Du, H., Wu, N., Fu, J., Wang, S., Li, X., Xiao, J., and Xiong, L. (2012). A GH3 family member, OsGH3-2, modulates auxin and abscisic acid levels and differentially affects drought and cold tolerance in rice. J. Exp. Bot. **63:** 6467–6480.

Dubcovsky, J., and Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. Science **316:** 1862–1866.

Feldman, M., Levy, A.A., Fahima, T., and Korol, A. (2012). Genomic asymmetry in allopolyploid plants: wheat as a model. J. Exp. Bot. **63:** 5045–5059.

Feschotte, C., Jiang, N., and Wessler, S.R. (2002). Plant transposable elements: where genetics meets genomics. Nat. Rev. Genet. **3:** 329–341.

Flagel, L.E., and Wendel, J.F. (2010). Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation. New Phytol. **186:** 184–193.

Frazier, T.P., Sun, G., Burklew, C.E., and Zhang, B. (2011). Salt and drought stresses induce the aberrant expression of microRNA genes in tobacco. Mol. Biotechnol. **49:** 159–165.

Gan, Q., Schones, D.E., Ho Eun, S., Wei, G., Cui, K., Zhao, K., and Chen, X. (2010). Monovalent and unpoised status of most genes in undifferentiated cell-enriched Drosophila testis. Genome Biol. **11:** R42.

Gao, D.Y., Chen, J.F., Chen, M.S., Meyers, B.C., and Jackson, S. (2012). A highly conserved, small LTR retrotransposon that preferentially targets genes in grass genomes. PLoS ONE **7:** e32010.

Garsmeur, O., Schnable, J.C., Almeida, A., Jourda, C., Hont, A., and Freeling, M. (2014). Two evolutionarily distinct classes of paleopolyploidy. Mol. Biol. Evol. **31:** 448–454.

Greaves, I.K., Groszmann, M., Ying, H., Taylor, J.M., Peacock, W.J., and Dennis, E.S. (2012). Trans chromosomal methylation in Arabidopsis hybrids. Proc. Natl. Acad. Sci. USA **109:** 3570–3575.

Groose, R.W., Talbert, L.E., Kojis, W.P., and Bingham, E.T. (1989). Progressive eterosis in autotetraploid alfalfa - studies using 2 types of inbreds. Crop Sci. **29:** 1173–1177.

Grover, C.E., Gallagher, J.P., Szadkowski, E.P., Yoo, M.J., Flagel, L.E., and Wendel, J.F. (2012). Homoeolog expression bias and expression level dominance in allopolyploids. New Phytol. **196:** 966–971.

Ha, M., Kim, E.D., and Chen, Z.J. (2009a). Duplicate genes increase expression diversity in closely related species and allopolyploids. Proc. Natl. Acad. Sci. USA **106:** 2295–2300.

Ha, M., Lu, J., Tian, L., Ramachandran, V., Kasschau, K.D., Chapman, E.J., Carrington, J.C., Chen, X., Wang, X.J., and Chen, Z.J. (2009b). Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids. Proc. Natl. Acad. Sci. USA **106:** 17835–17840.

Han, Y.J., Qin, S.S., and Wessler, S.R. (2013). Comparison of class 2 transposable elements at superfamily resolution reveals conserved and distinct features in cereal grass genomes. BMC Genomics **14:** 71.

Hao, M., Luo, J., Zhang, L., Yuan, Z., Yang, Y., Wu, M., Chen, W., Zheng, Y., Zhang, H., and Liu, D. (2013). Production of hexaploid triticale by a synthetic hexaploid wheat-rye hybrid method. Euphytica **193:** 347–357.

He, G., Chen, B., Wang, X., Li, X., Li, J., He, H., Yang, M., Lu, L., Qi, Y., Wang, X., and Wang Deng, X. (2013). Conservation and divergence of transcriptomic and epigenomic variation in maize hybrids. Genome Biol. **14:** R57.

He, P., Friebe, B.R., Gill, B.S., and Zhou, J.M. (2003). Allopolyploidy alters gene expression in the highly stable hexaploid wheat. Plant Mol. Biol. **52:** 401–414.

Huang, S.X., Sirikhachornkit, A., Su, X., Faris, J., Gill, B.S., Haselkorn, R., and Gornicki, P. (2002). Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. Proc. Natl. Acad. Sci. USA **99:** 8133–8138.

Jackson, S., and Chen, Z.J. (2010). Genomic and expression plasticity of polyploidy. Curr. Opin. Plant Biol. **13:** 153–159.

Jain, M., Kaur, N., Tyagi, A.K., and Khurana, J.P. (2006). The auxin-responsive GH3 gene family in rice (Oryza sativa). Funct. Integr. Genomics **6:** 36–46.

Jia, J., et al.; International Wheat Genome Sequencing Consortium (2013). *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. Nature **496:** 91–95.

Kashkush, K., Feldman, M., and Levy, A.A. (2003). Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nat. Genet. **33:** 102–106.

Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-wide profiling and analysis of Arabidopsis siRNAs. PLoS Biol. **5:** e57.

Kenan-Eichler, M., Leshkowitz, D., Tal, L., Noor, E., Melamed-Bessudo, C., Feldman, M., and Levy, A.A. (2011). Wheat hybridization and polyploidization results in deregulation of small RNAs. Genetics **188:** 263–272.

Khraiwesh, B., Zhu, J.K., and Zhu, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochim. Biophys. Acta **1819:** 137–148.

Kihara, H. (1944). Discovery of DD analyser, one of the ancestors of *T. vulgare*. Agric. Hortic. (Tokyo) **19:** 889–890.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. **14:** R36.

Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. **39:** D152–D157.

Leitch, A.R., and Leitch, I.J. (2008). Genomic plasticity and the diversity of polyploid plants. Science **320:** 481–483.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics **25:** 2078–2079.

Li, R., et al. (2010). De novo assembly of human genomes with massively parallel short read sequencing. Genome Res. **20:** 265–272.

Ling, H.-Q., et al. (2013). Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature **496:** 87–90.

Liu, C.G., Calin, G.A., Volinia, S., and Croce, C.M. (2008). MicroRNA expression profiling using microarrays. Nat. Protoc. **3:** 563–578.

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.

Lv, D.K., Bai, X., Li, Y., Ding, X.D., Ge, Y., Cai, H., Ji, W., Wu, N., and Zhu, Y.M. (2010). Profiling of cold-stress-responsive miRNAs in rice by microarrays. Gene **459:** 39–47.

Madlung, A. (2013). Polyploidy and its effect on evolutionary success: old questions revisited with new tools. Heredity (Edinb.) **110:** 99–104.

Matsuoka, Y. (2011). Evolution of polyploid triticum wheats under cultivation: the role of domestication, natural hybridization and allopolyploid speciation in their diversification. Plant Cell Physiol. **52:** 750–764.

Mayrose, I., Zhan, S.H., Rothfels, C.J., Magnuson-Ford, K., Barker, M.S., Rieseberg, L.H., and Otto, S.P. (2011). Recently formed polyploid plants diversify at lower rates. Science **333:** 1257.

McFadden, E.S., and Sears, E.R. (1946). The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J. Hered. **37:** 81–89, 107.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. **20:** 1297–1303.

Mestiri, I., Chagué, V., Tanguy, A.M., Huneau, C., Huteau, V., Belcram, H., Coriton, O., Chalhoub, B., and Jahier, J. (2010). Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity. New Phytol. **186:** 86–101.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods **5:** 621–628.

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.

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.

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.

Pont, C., et al. (2013). Wheat syntenome unveils new evidences of contrasted evolutionary plasticity between paleo- and neoduplicated subgenomes. Plant J. **76:** 1030–1044.

Pumphrey, M., Bai, J., Laudencia-Chingcuanco, D., Anderson, O., and Gill, B.S. (2009). Nonadditive expression of homoeologous genes is established upon polyploidization in hexaploid wheat. Genetics **181:** 1147–1157.

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell **80:** 847–857.

Qi, B., Huang, W., Zhu, B., Zhong, X., Guo, J., Zhao, N., Xu, C., Zhang, H., Pang, J., Han, F., and Liu, B. (2012). Global transgenerational gene expression dynamics in two newly synthesized allohexaploid wheat (*Triticum aestivum*) lines. BMC Biol. **10:** 3.

Rapp, R.A., Udall, J.A., and Wendel, J.F. (2009). Genomic expression dominance in allopolyploids. BMC Biol. **7:** 18.

Riddle, N.C., Jiang, H., An, L., Doerge, R.W., and Birchler, J.A. (2010). Gene expression analysis at the intersection of ploidy and hybridity in maize. Theor. Appl. Genet. **120:** 341–353.

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carré, I.A., and Coupland, G. (1998). The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. Cell **93:** 1219–1229.

Schnable, J.C., Springer, N.M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proc. Natl. Acad. Sci. USA **108:** 4069–4074.

Soltis, D.E., Albert, V.A., Leebens-Mack, J., Bell, C.D., Paterson, A.H., Zheng, C., Sankoff, D., Depamphilis, C.W., Wall, P.K., and Soltis, P.S. (2009). Polyploidy and angiosperm diversification. Am. J. Bot. **96:** 336–348.

Sparks, C.A., and Jones, H.D. (2009). Biolistics transformation of wheat. Methods Mol. Biol. **478:** 71–92.

Sreenivasulu, N., Usadel, B., Winter, A., Radchuk, V., Scholz, U., Stein, N., Weschke, W., Strickert, M., Close, T.J., Stitt, M., Graner, A., and Wobus, U. (2008). Barley grain maturation and germination: metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. Plant Physiol. **146:** 1738–1758.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25:** 1105–1111.

Vaucheret, H. (2006). Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes Dev. **20:** 759–771.

Wang, J., Tian, L., Lee, H.S., Wei, N.E., Jiang, H., Watson, B., Madlung, A., Osborn, T.C., Doerge, R.W., Comai, L., and Chen, Z.J. (2006). Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. Genetics **172:** 507–517.

Wang, J., Tian, L., Madlung, A., Lee, H.S., Chen, M., Lee, J.J., Watson, B., Kagochi, T., Comai, L., and Chen, Z.J. (2004). Stochastic and epigenetic changes of gene expression in Arabidopsis polyploids. Genetics **167:** 1961–1973.

Wang, L., Feng, Z., Wang, X., Wang, X., and Zhang, X. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics **26:** 136–138.

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics **28:** 2184–2185.

Washburn, J.D., and Birchler, J.A. (2014). Polyploids as a 'model system' for the study of heterosis. Plant Reprod. **27:** 1–5.

Wei, B., Cai, T., Zhang, R., Li, A., Huo, N., Li, S., Gu, Y.Q., Vogel, J., Jia, J., Qi, Y., and Mao, L. (2009). Novel microRNAs uncovered by deep sequencing of small RNA transcriptomes in bread wheat (*Triticum aestivum* L.) and *Brachypodium distachyon* (L.) Beauv. Funct. Integr. Genomics **9:** 499–511.

Wei, L.J., Gu, L.F., Song, X.W., Cui, X.K., Lu, Z.K., Zhou, M., Wang, L.L., Hu, F.Y., Zhai, J.X., Meyers, B.C., and Cao, X.F. (2014). Dicer-like 3 produces transposable element-associated 24-nt siRNAs that control agricultural traits in rice. Proc. Natl. Acad. Sci. USA **111:** 3877–3882.

Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B., and Rieseberg, L.H. (2009). The frequency of polyploid speciation in vascular plants. Proc. Natl. Acad. Sci. USA **106:** 13875–13879.

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:** W22–W28.

Wu, L., Liu, D.F., Wu, J.J., Zhang, R.Z., Qin, Z.R., Liu, D.M., Li, A.L., Fu, D.L., Zhai, W.X., and Mao, L. (2013). Regulation of FLOWERING LOCUS T by a microRNA in *Brachypodium distachyon*. Plant Cell **25:** 4363–4377.

Yang, J.H., Han, S.J., Yoon, E.K., and Lee, W.S. (2006). 'Evidence of an auxin signal pathway, microRNA167-ARF8-GH3, and its response to exogenous auxin in cultured rice cells'. Nucleic Acids Res. **34:** 1892–1899.

Yang, W., Liu, D., Li, J., Zhang, L., Wei, H., Hu, X., Zheng, Y., He, Z., and Zou, Y. (2009). Synthetic hexaploid wheat and its utilization for wheat genetic improvement in China. J. Genet. Genomics **36:** 539–546.

Yao, H., Dogra Gray, A., Auger, D.L., and Birchler, J.A. (2013). Genomic dosage effects on heterosis in triploid maize. Proc. Natl. Acad. Sci. USA **110:** 2665–2669.

Yoo, M.J., Szadkowski, E., and Wendel, J.F. (2013). Homoeolog expression bias and expression level dominance in allopolyploid cotton. Heredity (Edinb.) **110:** 171–180.

Zhang, H., et al. (2013a). Persistent whole-chromosome aneuploidy is generally associated with nascent allohexaploid wheat. Proc. Natl. Acad. Sci. USA **110:** 3447–3452.

Zhang, H.K., Bian, Y., Gou, X.W., Dong, Y.Z., Rustgi, S., Zhang, B.J., Xu, C.M., Li, N., Qi, B., Han, F.P., Wettstein, D.v., and Liu, B. (2013b). Intrinsic karyotype stability and gene copy number variations may have laid the foundation for tetraploid wheat formation. **110:** 19466–19471.

Zhang, J., Xu, Y., Huan, Q., and Chong, K. (2009). Deep sequencing of Brachypodium small RNAs at the global genome level identifies microRNAs involved in cold stress response. BMC Genomics **10:** 449.

Zhang, L., Yang, Y., Zheng, Y., and Liu, D. (2007). Meiotic restriction in emmer wheat is controlled by one or more nuclear genes that continue to function in derived lines. Sex. Plant Reprod. **20:** 159–166.

Zhang, L.Q., Liu, D.C., Zheng, Y.L., Yan, Z.H., Dai, S.F., Li, Y.F., Jiang, Q., Ye, Y.Q., and Yen, Y. (2010). Frequent occurrence of unreduced gametes in *Triticum turgidum-Aegilops tauschii* hybrids. Euphytica **172:** 285–294.

Zhao, N., Xu, L., Zhu, B., Li, M., Zhang, H., Qi, B., Xu, C., Han, F., and Liu, B. (2011a). Chromosomal and genome-wide molecular changes associated with initial stages of allohexaploidization in wheat can be transit and incidental. Genome **54:** 692–699.

Zhao, N., Zhu, B., Li, M., Wang, L., Xu, L., Zhang, H., Zheng, S., Qi, B., Han, F., and Liu, B. (2011b). Extensive and heritable epigenetic remodeling and genetic stability accompany allohexaploidization of wheat. Genetics **188:** 499–510.