New Phytologist

Transcriptome asymmetry in synthetic and natural allotetraploid wheats, revealed by RNA-sequencing

Xutong Wang[1], Huakun Zhang[1], Yaling Li[1], Zhibin Zhang[1], Linfeng Li[1,2] and Bao Liu[1]

[1]Key Laboratory of Molecular Epigenetics of the Ministry of Education (MOE), Northeast Normal University, Changchun 130024, China; [2]Department of Biology, Washington University in St Louis, St Louis, MO 63130, USA

Authors for correspondence:
Linfeng Li
Tel: +86 0431 8509 9367
Email: lilf241@nenu.edu.cn

Bao Liu
Tel: +86 0431 8509 9822
Email: baoliu@nenu.edu.cn

Received: *15 June 2015*
Accepted: *20 August 2015*

New Phytologist (2016) **209:** 1264–1277
doi: 10.1111/nph.13678

Key words: allopolyploidization, homeolog expression bias, transcriptome asymmetry, *Triticum–Aegilops* complex, wheat.

Summary

- Allopolyploidization has occurred frequently within the *Triticum–Aegilops* complex which provides a suitable system to investigate how allopolyploidization shapes the expression patterns of duplicated homeologs.
- We have conducted transcriptome-profiling of leaves and young inflorescences in wild and domesticated tetraploid wheats, *Triticum turgidum* ssp. *dicoccoides* (BBAA) and ssp. *durum* (BBAA), an extracted tetraploid (BBAA), and a synthetic tetraploid (S^lS^lAA) wheat together with its diploid parents, *Aegilops longissima* (S^lS^l) and *Triticum urartu* (AA).
- The two diploid species showed tissue-specific differences in genome-wide ortholog expression, which plays an important role in transcriptome shock-mediated homeolog expression rewiring and hence transcriptome asymmetry in the synthetic tetraploid. Further changes of homeolog expression apparently occurred in natural tetraploid wheats, which led to novel transcriptome asymmetry between the two subgenomes. In particular, our results showed that extremely biased homeolog expression can occur rapidly upon the allotetraploidzation and this trend is further enhanced in the course of domestication and evolution of polyploid wheats.
- Our results suggest that allopolyploidization is accompanied by distinct phases of homeolog expression changes, with parental legacy playing major roles in the immediate rewiring of homeolog expression upon allopolyploidization, while evolution and domestication under allotetraploidy drive further homeolog-expression changes toward re-established subgenome expression asymmetry.

Introduction

The role of allopolyploidization (hybridization concomitant with or followed by whole-genome duplication (WGD)) in the diversification and speciation of vascular plants is increasingly appreciated (Soltis *et al.*, 2009; Abbott *et al.*, 2013; Buggs *et al.*, 2014). The abrupt reunion and doubling of two or more divergent genomes in the same nucleus and cytoplasm entail a suite of genomic accommodations to circumvent incompatibilities, whereby rapid and profound changes in genome structure and gene expression are often generated (Wendel, 2000; Comai, 2005; Feldman & Levy, 2005; Chen, 2007; Otto, 2007; Doyle *et al.*, 2008; Hegarty & Hiscock, 2008).

Accumulated evidence indicates that these induced genetic and epigenetic changes may bestow novel phenotypes upon the nascent allopolyploids, which enable or facilitate their immediate establishment, ecological diversification and adaptation to new niches (Adams & Wendel, 2005; Chen, 2007; Jackson & Chen, 2010; Madlung, 2013; Yoo & Wendel, 2014). In the case of *Tragopogon*, for example, the diploid parents, F_1 hybrids, synthetic and natural nascent allotetraploids all showed dramatic differences in inflorescence morphology and many other phenotypes (Soltis *et al.*, 2004; Tate *et al.*, 2006, 2009; Lim *et al.*, 2008; Malinska *et al.*, 2010). Indeed, a series of studies based on multiple molecular approaches have shown that both genetic (i.e. chromosomal rearrangement and gene loss) and epigenetic (i.e. cytosine methylation) changes were generated during and/or soon after the polyploidization process (Tate *et al.*, 2006, 2009; Lim *et al.*, 2008; Buggs *et al.*, 2009, 2010b; Malinska *et al.*, 2011; Koh *et al.*, 2012; Sehrish *et al.*, 2014; Dobesova *et al.*, 2015). Meanwhile, several studies from diverse plant taxa employing microarray- or RNA-seq-based transcriptome profiling have revealed that changes in global gene expression occurred ubiquitously in nascent allopolyploid plants (Pumphrey *et al.*, 2009; Akhunova *et al.*, 2010; Chague *et al.*, 2010; Qi *et al.*, 2012; Li *et al.*, 2014). In addition, it was found that the duplicated genes derived from different parental species (i.e. homeologs) often manifest biased expression in natural allopolyploids of *Senecio* (Hegarty *et al.*, 2006; Hegarty & Hiscock, 2008), *Arabidopsis* (Chang *et al.*, 2010) and *Spartina* (Chelaifa *et al.*, 2010). Importantly, the impacts of polyploidization on the morphologies and genome constitutions of several polyploid crops, including wheat, tobacco, sugarcane, coffee and cotton, have also been documented, and have probably been selected for during the domestication process

(Renny-Byfield & Wendel, 2014). Together, these observations strongly suggest that polyploidization represents a potent driving force which has positively contributed to the evolution and diversification of plants (Soltis *et al.*, 2014) and domestication of crops (Renny-Byfield & Wendel, 2014). However, few investigations have attempted to link the changes of gene expression induced in the immediate aftermath of polyploidization with changes of gene expression that occur on a longer evolutionary timescale.

The *Triticum–Aegilops* complex comprises eight distinct genome groups (A, C, D, M, N, S, T and U) and all but the T genome were involved in multiple independent allopolyploidization events (Zohary & Feldman, 1962; Gill & Friebe, 2002; Marcussen *et al.*, 2014; Li *et al.*, 2015a). To date, 18 naturally occurring allopolyploid species have been described in the complex, of which the A, U and D genomes are the diploid parents to most of the extant polyploid species (Feldman *et al.*, 2012). Of particular significance, allopolyploidization between A and S (also referred to as the B) genomes has led to the formation of wild tetraploid wheat *Triticum turgidum* ssp. *dicoccoides* (BBAA) some 0.36–0.7 million yr ago (Dvořák, 1976; Huang *et al.*, 2002; Dvorak & Akhunov, 2005; Gornicki *et al.*, 2014). Then, the D genome of *Aegilops tauschii* was later introduced into a domesticated form of tetraploid wheat (i.e. *T. turgidum* ssp. *durum*) *c.* 8000 yr ago (Feldman *et al.*, 1995; Willcox, 1998), which has led to the establishment of hexaploid common wheat, *Triticum aestivum* (BBAADD). The evolutionary history of ancient and more recent allopolyploidization events render the *Triticum–Aegilops* complex an excellent system to explore the creative roles of allopolyploidy in general and the immediate effects of genome merging in particular, which together have profoundly impacted species diversification in this complex. For instance, by comparing the nuclear DNA amount of 27 natural and 14 newly synthesized allopolyploids and their diploid parents, Eilam *et al.* (2008) found that rapid genome downsizing occurred during and/or immediately after the formation of allopolyploids. Similarly, it has also been reported that changes in cytosine methylation and 45S rDNA copy number occurred in the newly formed allotetraploids of *Aegilops sharonensis* ($S^{sh}S^{sh}$) × *Triticum monococcum* (AA) and *A. sharonensis* × *Aegilops umbellulata* (UU), and that these genetic and epigenetic alterations exhibited subgenome-specific bias (Shaked *et al.*, 2001; Shcherban *et al.*, 2008). In addition, several studies based on microarray analysis have shown that nonadditive parental gene expression is a common feature in newly synthesized and natural hexaploid wheats (Pumphrey *et al.*, 2009; Akhunova *et al.*, 2010; Chague *et al.*, 2010; Pont *et al.*, 2011; Qi *et al.*, 2012; Chelaifa *et al.*, 2013). More recently, several independent studies based on transcriptome and locus-specific bisulfite sequencing have shown that changes in homeolog expression and cytosine methylation modification can occur immediately after allopolyploidization (Guo & Han, 2014; Leach *et al.*, 2014; Li *et al.*, 2014), and cell type- and developmental stage-dependent subgenome expression dominance was observed in hexaploid common wheat (Pfeifer *et al.*, 2014). Notably, most of the studies in wheat have mainly focused on the phenomenon of allopolyploidization-induced genetic and epigenetic changes *per se*, while important questions regarding the fates of differentially expressed homeologs in natural tetraploid and hexaploid wheats during the course of domestication and longer-term evolution remained uninvestigated.

In this study, we conducted a comprehensive genome-wide analysis of homeolog expression in a synthetic allotetraploid wheat (S^lS^lAA) and its diploid parents, *Triticum urartu* (AA) and *Aegilops longissima* (S^lS^l), to assess to what extent the A- and S^l-subgenome homeologs were rewired and differentially expressed in the newly formed tetraploid wheat. We then analyzed global scale homeolog expression in both wild and cultivated natural tetraploid wheats to assess how the global homeolog expression bias or transcriptome asymmetry that is established in the short term may further change during the course of evolution and domestication. In particular, phenotypic asymmetry conditioned by the A and B subgenomes was reported in the tetraploid wheat (Feldman *et al.*, 2012). We therefore wished to test if the observed phenotypic subgenome asymmetry of tetraploid wheat can be reflected by genome-wide homeolog expression biases. Further, we investigated the unique effect of allohexaploidy vs allotetraploidy on homeolog expression of the BBAA subgenomes that are the counterpart of natural tetraploid wheat. Finally, we addressed how domestication has shaped the pattern of homeolog expression at both the tetraploid and hexaploid levels.

Materials and Methods

Plant materials and RNA extraction

The plants used in this study (Fig. 1) included the wild and domesticated tetraploid wheat *Triticum turgidum* ssp. *dicoccoides* (Korn) Thell. (TD, BBAA) and ssp. *durum* Desf. (TTR13, BBAA), respectively; a newly synthesized allotetraploid wheat (AT2, S^lS^lAA) along with its diploid parents, *Triticum urartu* Tum. ex. Gandil. (TMU06, AA), and *Aegilops* longissima Schweinf. ex Muschl. (TL05, S^lS^l). An 'extracted' tetraploid wheat (ETW) with genome BBAA highly similar to the tetraploid component of hexaploid common wheat, *T. aestivum* L. (cv TAA10), was also used in this study (Zhang *et al.*, 2014). The ETW was initially produced and kindly provided by E. Kerber via hybridization between TAA10 and a tetraploid wheat (generating the F_1 pentaploid, genome BBAAD), followed by seven cycles of backcrossing to TAA10 as the recurrent parent (Kerber, 1964). ETW was backcrossed to TAA10 as the recurrent parent two times more and then propagated via self-pollination (to eliminate the D-subgenome chromosomes) in our laboratory for five further generations (Zhang *et al.*, 2014). Thus, in theory, the genome (BBAA) of ETW should be > 99.8% identical to the BBAA subgenomes of its bread wheat donor (TAA10) after the ninth backcross (1–2 ($1/2^9$)) (Zhang *et al.*, 2014). Remarkably, ETW is karyotypically stable as a *bona fide* euploidy based on extensive karyotyping of a large number of independent individuals (Zhang *et al.*, 2014). Similarly, both karyotype and morphological features of the synthetic tetraploid AT2 made it suitable to serve as a proximal model to recapitulate the initial allotetraploidization event(s) leading to the formation of natural tetraploid wheat, *T. turgidum* (Zhang *et al.*, 2013). The synthetic

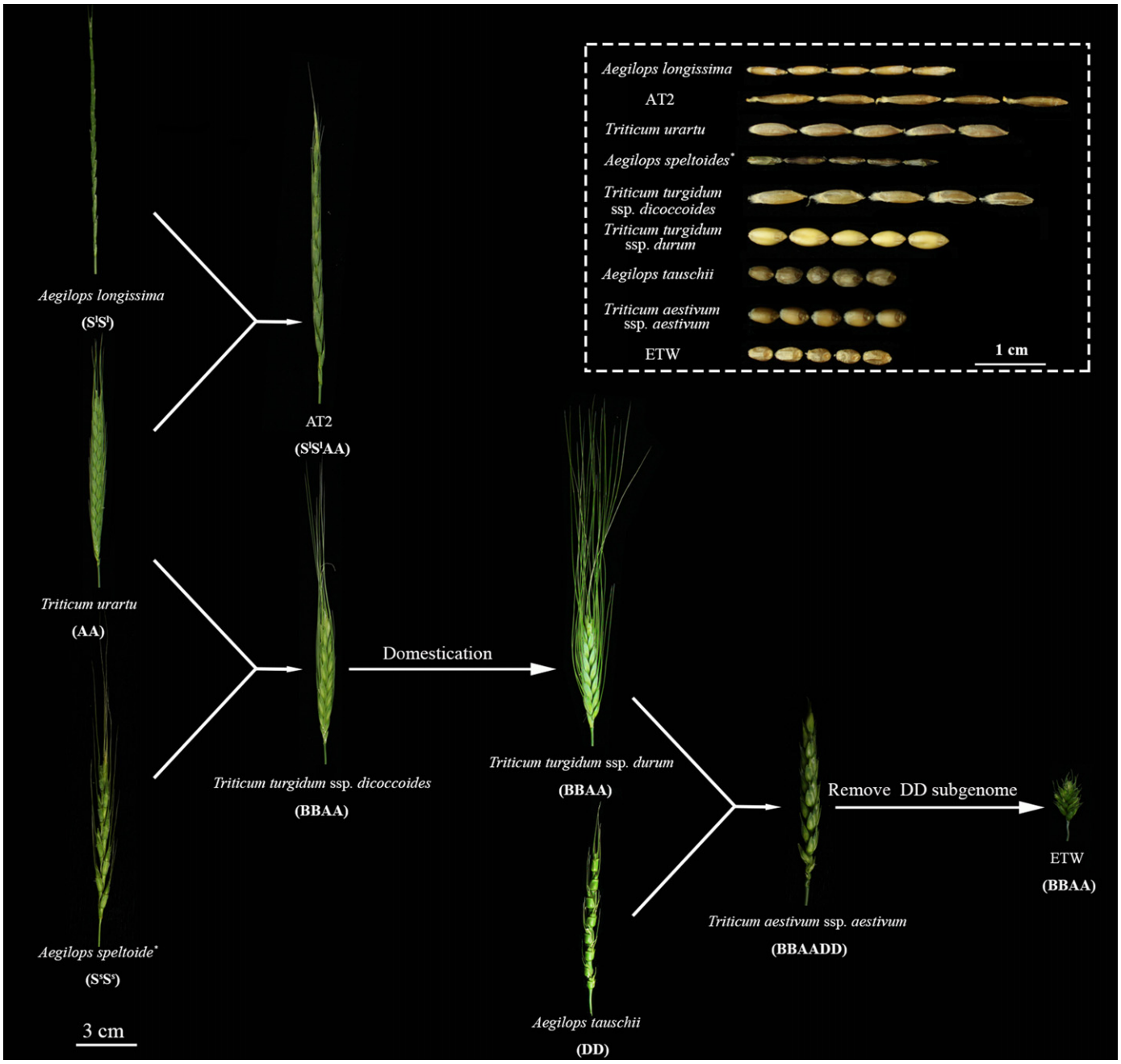

Fig. 1 The spike and seed morphologies of diploid, tetraploid and hexaploid wheats and their polyploidization and domestication processes. *The wild progenitor of the B subgenome of tetra- and hexaploid wheats is still undecided but is closely related to species of the *Sitopsis* section of *Aegilops*, such as *A. speltoides*. ETW, extracted tetraploid wheat.

tetraploid AT2 with a confirmed euploid chromosome number ($2n = 4x = 28$) were used for tissue collections, containing the second leaves and the young inflorescences (*c.* 0.5 cm in length). All plants were grown in a glasshouse, under the same controlled growing conditions: 25 : 20°C, 16 : 8 h, day : night. All the collected tissues were immediately frozen in liquid nitrogen. Total RNAs were isolated from the frozen tissues of leaves and young inflorescences for each sample separately using Trizol (Invitrogen) according to the manufacturer's instructions. The integrity of the extracted RNA was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany).

Transcriptome sequencing and reference sequence assembling

Transcriptome libraries were constructed for each tissue of the diploid and tetraploid wheats separately and then sequenced using Hiseq 2000 (BGI, Shenzhen, China) with standard protocols. Two biological replicates were conducted for each tissue of the samples and sequenced as parallel experiments. Low-quality reads (Phred < 30) were removed from the raw data using FASTX-Toolkit (Gordon & Hannon, 2012). Filtered reads of the diploid species, *A. longissima* and *T. urartu*, were used to

assemble reference sequences using the program Trinity (Grabherr *et al.*, 2011) with default parameters. In total, we obtained *c.* 1.5 billion high-quality reads (100 bp) (Supporting Information Table S1), and all clean reads have been deposited in the SRA database of GenBank (http://www.ncbi.nlm.nih.gov/sra) with the BioProject accession number PRJNA272886. To further evaluate the quality of the *de novo* assemblies, the 'findorf' program (https://github.com/vsbuffalo/findorf) was used to search the existing protein database of *Brachypodium distachyon* (ftp://brachypodium.org/brachypodium.org) and *Oryza sativa* (ftp://ftp.plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs) (Table S2). Thereafter, the program BLAT (Kent, 2002) was employed to compare the length of assembled references of the two diploid species with the following parameters: the matched length of each reference sequence is longer than 300 bp with the gap no more than 100 bp; the assembled reference covers more than half of the total length of the gene. Accordingly, the aligned reference sequences of *T. urartu* were selected as the reference because it contains more common homeologs than that of *A. longissima* (Table S1).

Identification of subgenome-specific homeolog expression patterns in tetraploid wheats vs their ortholog expression in the diploid species

The RNA-seq data of leaves and young inflorescences for each sample were mapped to the assembled references separately using the software BWA (Li & Durbin, 2009), and single nucleotide polymorphisms (SNPs) were diagnosed using SAMtools (Li *et al.*, 2009). To identify genome-specific orthologs from the diploid species *T. urartu* and *A. longissima* and then determine their expression patterns in the synthetic and natural tetraploid wheats, we filtered the obtained homeolog SNPs according to the following criteria: mapping quality of each identified SNP is higher than 30 and depth is > 10; heterozygous SNPs of diploid species *T. urartu* and *A. longissima* and homozygous SNPs of tetraploids were removed because they were not able to distinguish the genome-specific homeologs; heterozygous SNPs that exist across the four tetraploids were selected. All subgenome-specific SNPs are shown in Table S3.

To assess the accumulated orthologous gene expression divergence of the two parental species, an *in silico* parental mix was constructed by combining the filtered reads of *T. urartu* and *A. longissima* at a ratio of 1 : 1. Equally and differentially expressed orthologs were determined according to the relative transcript abundance of the A and S (~B) orthologs in the *in silico* parental mix. That is to say, for a given gene, if expression of the A ortholog was significantly higher than that of S ortholog (exact binomial test, $P < 0.05$), it was designated 'A-biased', and the reverse was designated 'B-biased', while orthologs showing equal expression were categorized as 'not biased' (exact binomial test, $P \geq 0.05$). To further explore how the biased-expressing orthologs were remodeled upon allotetraploidization when they became homeologs in the same nucleus, we assessed the A-homeolog relative expression ratio (A%) of each commonly expressed gene in leaves and young inflorescences of the synthetic and natural tetraploid wheats, respectively. The spectra of the biased-expression homeologs were presented for the *in silico* parental mix and all four tetraploid wheats using ggplot2 (Wickham, 2009). To further evaluate homeologous expression patterns at the individual gene level, we calculated the numbers of genes that showed biased or not-biased expression pattern in the two tissues of diploid parental species and each of the tetraploid wheat genotypes. The comparison of genes numbers showing biased and unbiased homeolog expression within or among the genotypes were visualized using UpSetR (Lex *et al.*, 2014). In addition, it has been reported that the *Q*/*q* gene played a crucial role in the domestication of polyploid wheat (Simons *et al.*, 2006; Zhang *et al.*, 2011), and phenotypic asymmetry of the young inflorescences was observed in the diploid species and tetraploid wheats studied (Fig. 1). We have therefore inspected the expression patterns of both *Q* and *q* alleles in the two tissues of diploid and tetraploid wheats.

Furthermore, two groups of extremely biased expression genes were defined as showing A% > 0.95 or A % < 0.05 expression patterns in either or both tissues of at least one of the allotetraploid wheats. The group I genes are those in which the extremely biased homeolog expression pattern was inherited from the diploid parents; group II genes are those in which the extremely biased homeolog expression pattern was generated *de novo* in an allotetraploid, which showed no parental expression-biased in the *in silico* parental mix. Potential biological functions of biased-expression homeologs in leaves and young inflorescences were determined by searching against the protein database of GenBank using the Blast2GO program (Conesa *et al.*, 2005). Details of the extremely biased expression genes are shown in Table S4. The gene ontology (GO) enrichment was estimated by hypergeometric test and the significance (*P*-value) was adjusted by false discovery rate (FDR) (Benjamini & Hochberg, 1995). The statistical significance of each comparison was tested in R (R Development Core Team, 2014), including the binominal exact test, chi-squared test, hypergeometric test, Fisher's exact test, Kolmogorov–Smirnov (K–S) test and Pearson's product–moment correlation. Regulatory modules underlying changes in homeolog expression were analyzed in each of the tetraploid wheats according to a previous study (Xu *et al.*, 2014), whereby the genes could be classified into three distinct regulatory groups, in each of the two tissues (detailed in the Results section). Furthermore, to assess the expression patterns in both tissues of each of the tetraploid wheats in relation to their intrinsic relative ortholog expression abundance of the diploid parental species, we classified the genes into nine patterns belonging to three categories according to criteria originally proposed in cotton (Yoo *et al.*, 2013) and being further elaborated in a fungal allopolyploid (Cox *et al.*, 2014) (detailed in the Results section).

Results

Assigning transcripts of allotetraploid wheats to subgenome homeologs

In this study, 34 174 and 43 035 contigs were obtained for the diploid parental species, *T. urartu* and *A. longissima*, respectively,

of which 23 200 (67.9%) and 26 985 (62.7%) open reading frames were predicted in the two assembled reference transcriptomes (Table S2). To facilitate comparisons of transcript abundance in the synthetic and natural allotetraploid wheats relative to their diploid parents or progenitors, 69 304 common homeolog-specific SNPs representing 9409 genes were recognized in each of the four tetraploid wheats (synthetic AT2, wild TD, domesticated TTR13 and extracted tetraploid wheat ETW) by mapping their reads to the reference sequences. Accordingly, a total of 7800 and 8900 genes in the leaf and young inflorescence tissues were subjected to homeolog expression analysis, respectively, of which 7450 genes were shared between the two tissues. The correlation coefficients of each biological replicate for the diploid and tetraploid wheats are 0.88 and 0.99 for leaves and young inflorescence, respectively (Table S1, Pearson's correlation test, $P < 2.2e{-}16$).

Tissue-specific ortholog expression difference between the diploid progenitor species

We investigated genome-wide ortholog expression differences in two tissue types, leaves and young inflorescences, between the two diploid parental species, *T. urartu* and *A. longissima*. We found that in leaves, 4852 (61.1%) of the 7944 genes showed differential ortholog expression (i.e. a ratio of ortholog expression $\neq 1:1$ in the *in silico* parental mix), based on binominal test with FDR adjustment (Table 1). For the parental ortholog differentially expressed genes in leaves, a moderate but significant number of genes was more A-biased than S-biased (exact binomial test, $P = 0.002$) in the *in silico* parental mix, indicating that genes of *T. urartu* have overall higher expression levels than those of *A. longissima* in the leaf tissue. By contrast, for young inflorescences, 5397 (60%) of the 9002 genes showed differential expression (Table 1), with significantly more S-biased than A-biased genes (exact binomial test, $P < 2.2e{-}16$) compared with the *in silico* parental mix (Table 1). These results indicate that the evolutionarily accumulated genome-wide gene expression differences between the two diploid species (*T. uratu* and *A. longissima*) show tissue specificity, with overall expression levels in leaves and young inflorescences being reciprocally higher in one species over the other. These tissue-specific differential expression features between the diploid parental species may have significant relevance to changes of homeolog expression in the resulting tetraploid wheats when the diverged parental genomes were merged and doubled in a common nucleus and cytoplasm.

Immediate and longer-term evolutionary changes of homeolog expression in leaf and young inflorescence tissues of synthetic and natural tetraploid wheats

To assess how the parental orthologs would be expressed as homeologs in the synthetic tetraploid wheat (AT2), and by extension, in the three types of natural tetraploid wheats (TD, TTR13 and ETW), we tabulated the relative A-homeolog expression ratio (A% of the total) of each of the 7774 and 8885 commonly expressed orthologous genes in leaves and young inflorescences, respectively (Fig. 2). We made the following major observations: compared with the *in silico* parental mix, the spectra of parental homeolog expression divergences were significantly reduced (became more aggregated) in leaves (K–S test, $P < 0.001$, Fig. 2a), but augmented (became more dispersed) in young inflorescences of all the tetraploid wheats studied (K–S test, $P < 0.001$; Fig. 2b); compared with AT2, the spectra of parental homeolog expression divergence were significantly increased in both tissue types of all three natural tetraploid wheats (K–S test, $P < 0.05$; Fig. 2a,b); among the three natural tetraploid wheats, there was no significant difference in the spectra of parental homeolog expression divergence in leaves (K–S test, $P > 0.05$; Fig. 2a), but significant differences were detected in young inflorescences between any two of the three natural tetraploid wheats (K–S test, $P < 0.001$; Fig. 2b), with ETW exhibiting the largest spectrum (K–S test, $P < 0.001$; Fig. 2b).

We next examined whether the altered spectra of parental homeolog expression divergence would impact the overall subgenome expression dominance in a given tetraploid wheat. We found that: the overall higher ortholog expression levels in leaves of *T. urartu* and in young inflorescences of *A. longissima* (Table 1) were largely retained when they became homeologs in AT2 (Fig. 2), suggesting that the immediate attenuation (in leaf tissue) and augmentation (in young inflorescences) of homeolog expression in AT2 were largely proportional between the two subgenomes; of the three types of natural tetraploid wheats, while they all maintained the A-subgenome expression dominance in leaves as in AT2 (Fig. 2a), they all showed A-subgenome

Table 1 Homeolog expression bias in leaves and young inflorescences of a synthetic tetraploid wheat (S^lS^lAA, AT2) and natural tetraploid wheats, including wild (BBAA, TD) and domesticated (BBAA, TTR13) tetraploid wheats and an extracted tetraploid wheat (BBAA, ETW)

Tissue	Type	*In silico* hybrid (%)	AT2 (%)	TD (%)	TTR13 (%)	ETW (%)
Leaf	A bias	2536 (31.9)	2322 (29.5)	2428 (31.0)	2423 (31.0)	2766 (35.2)
	B bias	2316 (29.2)	2092 (26.6)	2147 (27.5)	2110 (27.0)	2106 (26.8)
	Not biased	3092 (38.9)	3443 (43.8)	3245 (41.5)	3275 (41.9)	2992 (38.0)
	Total	7944 (100)	7857 (100)	7820 (100)	7808 (100)	7864 (100)
YI	A bias	2098 (23.3)	2431 (27.2)	3512 (39.3)	3419 (38.2)	4325 (48.5)
	B bias	3299 (36.6)	2864 (32.0)	2269 (25.4)	2302 (25.7)	2109 (23.6)
	Not biased	3605 (40.0)	3641 (40.7)	3162 (35.4)	3228 (36.1)	2491 (27.9)
	total	9002 (100)	8936 (100)	8943 (100)	8949 (100)	8925 (100)

YI, young inflorescence.

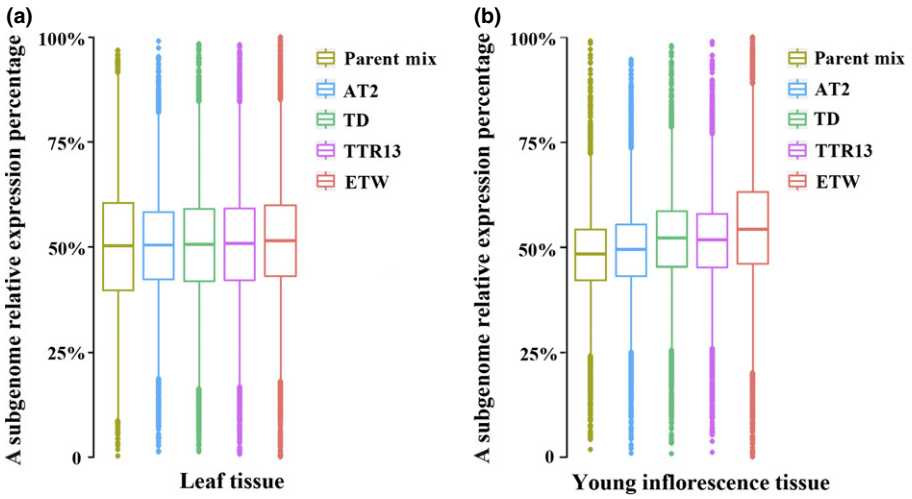

Fig. 2 Global spectra of homeolog expression divergence (based on transcriptome profiling) of a total of 7774 and 8885 genes in leaves (a) and young inflorescences (b), respectively, in the *in silico* parent mix, a synthetic tetraploid (AT2), a wild tetraploid wheat (TD), a domesticated tetraploid (TTR13) and an extracted tetraploid wheat (ETW), revealed by a boxplot that showed the proportion of A homeolog transcript abundance out of the total transcripts.

expression dominance in young inflorescences, which was different from AT2, in which the S subgenome was dominant (Fig. 2b). The latter observation suggests that more extensive evolutionary changes to homeolog expression occurred in young inflorescences than in leaves. Similar phenomena were also observed in the comparison of gene numbers that showed biased or unbiased homeolog expression within and between the genotypes (Fig. 3). We found that the majority of the homeolog expression-biased and unbiased genes were specific to a given genotype in both tissues; a total of 1066 and 1142 genes showed

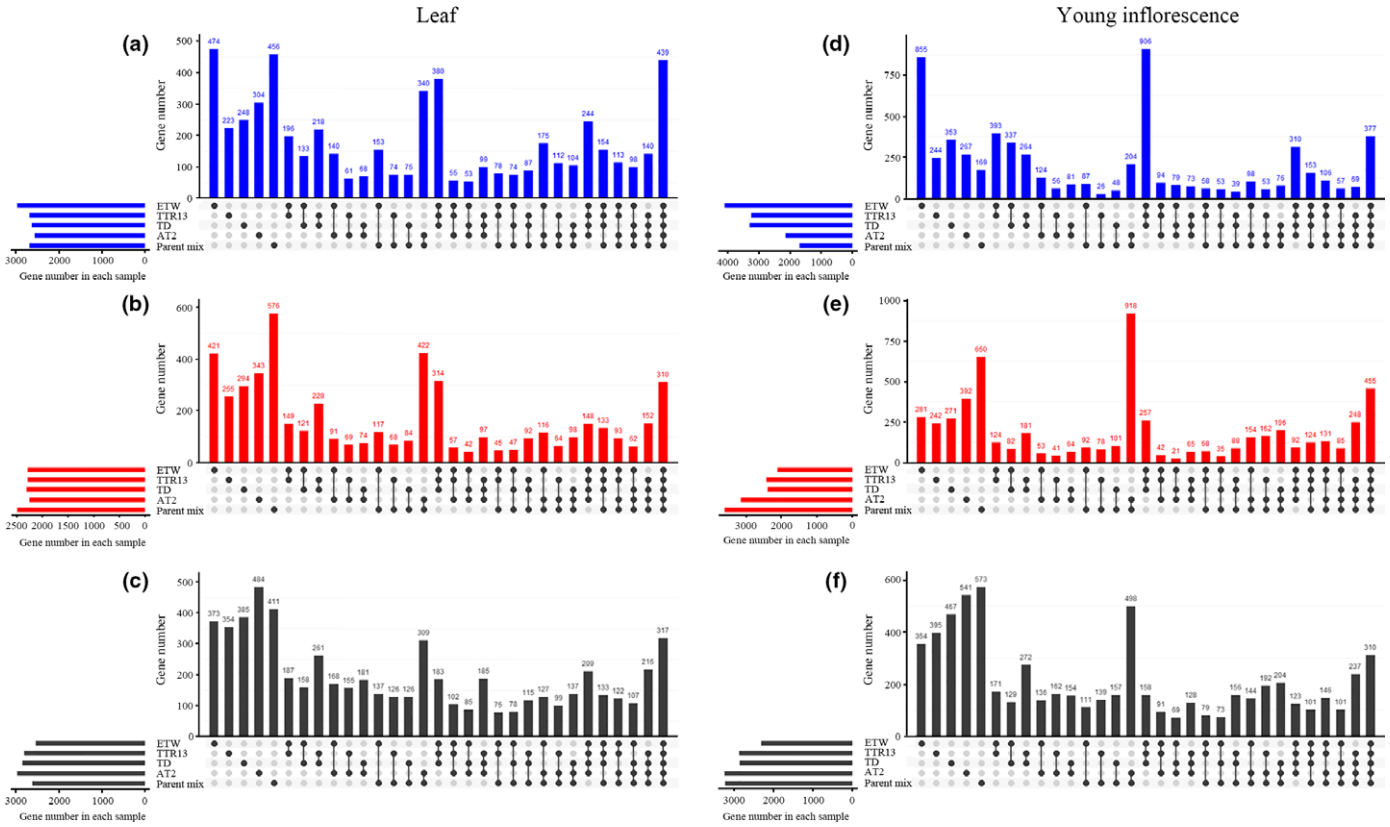

Fig. 3 Quantification of genes that showed biased or unbiased homeolog expression in leaves (left panels) and young inflorescences (right panels) of each diploid parental species (parental mix) and the four tetraploid wheat genotypes. The numbers of genes showing A-biased, B-biased and unbiased expression patterns in one or more of the genotypes are shown in blue (a, d), red (b, e) and black (c, f) colored vertical bars, respectively. Black dot(s) at the bottom of each vertical bar indicate the biased or unbiased homeolog expression identified in a given tissue of each genotype. The lined dots indicate two or more genotypes showing the same (i.e. shared) biased or unbiased homeolog expression pattern. The total numbers of genes that showed a given biased or unbiased homeologous expression pattern in each genotype were tabulated and are represented by horizontal bars on the left of each figure. AT2, a synthetic tetraploid; TD, a wild tetraploid wheat; TTR13, a domesticated tetraploid; ETW, an extracted tetraploid wheat.

the same biased or unbiased expression pattern across all the genotypes (including the diploid parental species and all four tetraploid wheats) in leaves and young inflorescences; a large number of genes showed biased or unbiased expression in a given tissue of only the three natural tetraploids and the trends are more apparent in young inflorescences than in leaves.

Extremely biased homeolog expression in synthetic and natural tetraploid wheats

To further study the phenomenon of homeolog expression rewiring, we investigated those genes that showed extremely biased expression in the two tissue types, leaves and young inflorescences, of each tetraploid wheat. We found that the extremely biased group I genes (defined as A% > 95% or < 5% of the total transcripts; see the Materials and Methods section) were small in number between the two diploid parental species, *T. urartu* and *A. longissima*; only 41 and 28 such genes were identified from the genes we studied in the leaf and young inflorescence tissues, respectively, and only four genes were shared by the two tissues. We traced homolog expression patterns for each of these genes in the respective tissues of the synthetic (AT2) and natural (TD, TTR13 and ETW) tetraploid wheats. We found that the expression patterns of the group I genes in both tissues of the tetraploid wheats were highly conserved; that is, for great majority of these genes, their expression patterns were maintained in both tissues of all four types of tetraploid wheats (Fig. 4a,b, upper panels). Occasionally, reversal in the biased extreme expression direction by the two subgenomes (i.e. from A homeolog expression to S (or B) homeolog expression or vice versa) of the group I genes was observed in some of the tetraploid wheats (Fig. 4a,b, upper panels). We then explored the possibility of *de novo* genesis of extremely biased expression patterns from parental ortholog nonextremely differentially expressed genes (defined as A% > 33% or < 66% of the total transcripts; see the Materials and Methods section) as a result of allotetraploidization, that is, the occurrence of extremely biased group II genes. We found that the *de novo* biased expression genes (group II) were significantly greater in number than the parental inherent biased expression genes (group I), with 106 and 128 genes being identified in leaves and young inflorescences, respectively (Fig. 4a,b, lower panels). Several interesting observations emerged in the expression patterns of the group II genes. First, the biased expression pattern of some of these genes was already established in AT2 in both tissues. Second, some of these rapidly occurring extremely biased genes showed evidence of pattern conservation in all three types of natural tetraploid wheats in leaves, although this was a rare occurrence in young inflorescences. Third, as expected, the natural tetraploid wheats contained a greater number of extremely biased expression genes than did the synthetic wheats, and some of these were conserved among all three natural tetraploid wheats. Finally, in each tetraploid wheat, while leaves showed similar numbers of extremely biased expression genes by both subgenomes (exact binomial test, $P > 0.58$), the A subgenome manifested a significantly larger number of these genes in young inflorescences compared with the S (~B) genome (exact binomial test, $P < 0.01$; Fig. 4a,b, lower panels), consistent with the foregoing observation that, with regard to overall expression, the A subgenome was dominant in all three tetraploid wheats.

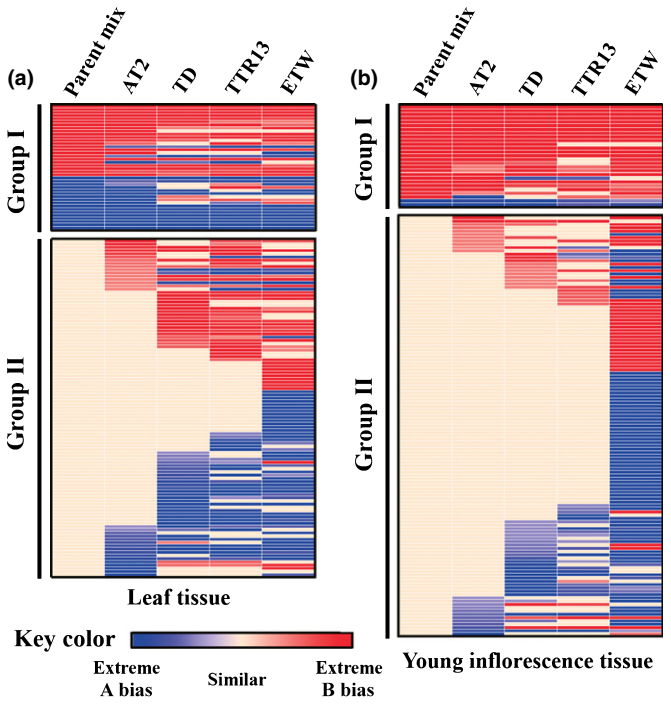

Fig. 4 Expression of the homeolog extremely biased expressing genes (defined as A% > 0.95 or A% < 0.05; see the Materials and Methods section) in leaves (a) and young inflorescences (b) of the *in silico* parent mix and each of the tetraploid wheats. Expression differences less than twofold were defined as similar expression (same color). The group I extremely biased expression genes are those whose expression levels are extremely different between the parental orthologs, while the group II genes are those that showed nonextremely biased expression in the diploid parents but became extremely biased in expression between the homeologs in the tetraploid(s).

Homeolog expression in tetraploid wheats in relation to their diploid parental legacy and evolutionary changes

To assess the expression patterns in both tissues of each of the tetraploid wheats in relation to their intrinsic relative ortholog expression states of the diploid parental species, we classified the genes into nine patterns belonging to three categories according to criteria originally proposed in cotton (Yoo *et al.*, 2013) and further elaborated in a fungal allopolyploid (Cox *et al.*, 2014). Briefly, the three categories were: (1) expression bias group I, which contained three subpatterns, all representing vertical transmission of the relative parental ortholog expression states by the corresponding homeologs in a given tetraploid wheat; (2) expression bias group II, which contained two subpatterns, both representing elimination of parental ortholog expression difference by homeologs in a given tetraploid wheat; (3) expression bias group III, which included four subpatterns, all referring to newly biased expression by homeologs in a given tetraploid wheat from genes that showed no expression difference between the corresponding

parental orthologs or the original ortholog expression bias was reversed to the opposite direction (Yoo *et al.*, 2013; Cox *et al.*, 2014).

We found that in AT2, the three subpatterns of bias group I (i.e. B(S) = A, B(S) > A and B(S) < A) together accounted for 56.3 and 63.1% of the analyzed genes in leaves (4391 of 7798) and young inflorescences (5628 of 8913), respectively (Table 2), suggesting that vertical transmission of the intrinsic parental ortholog expression states (i.e. parental legacy) was a major determinant factor for the relative homeolog expression states in the synthetic tetraploid wheat AT2. Compared with AT2, the proportions of all three patterns of bias group I were substantially reduced in both tissues of the natural tetraploid wheats (Fisher's exact test, $P < 2.2e{-}16$) (Table 2), suggesting enhanced divergent evolution of homeolog expression for substantial numbers of genes over evolutionary time. Unexpectedly, the proportions of each of the two patterns of bias group II were largely conserved in both tissues of all tetraploid wheats, including AT2 (Table 2). Compared with AT2, the proportions of all four patterns of bias group III were increased in both tissues of the natural tetraploid wheats, but the increments were more pronounced in young inflorescence (Table 2). Moreover, of the four patterns of bias group III, although proportions of both directions of the reversed extreme homeolog expression type (i.e. from B > A to B < A or vice versa) showed significant increments (Fisher's exact test, $P < 2.2e{-}16$) in both tissues of the natural tetraploid wheats compared with AT2, in young florescences, the number of genes showing one direction of homeolog expression reversal (i.e. from B > A to B < A) was two to four times greater than the number showing the other direction (Table 2), again consistent with the overall expression dominance by the A subgenome in this tissue of all three natural tetraploid wheats.

Regulatory modules underpinning homeolog expression trended from convergence to divergence in natural tetraploid wheats

To address the regulatory mechanisms responsible for the homeolog expression rewiring in synthetic and natural tetraploid wheats, we compared homeolog expression in the tetraploid wheats with parental ortholog expression states in the *in silico* parental mix from a different angle. Specifically, based on previously established criteria (Xu *et al.*, 2014) we classified the genes into three distinct regulatory groups, I, II and III. Regulatory group I consists of those in which the expression differences of homeologs in a given tetraploid were significantly reduced relative to the *in silico* parental mix, and hence trended towards the 1 : 1 ratio; regulatory group II are those in which the expression differences of parental homeologs were significantly increased in a given tetraploid relative to the *in silico* parental mix, and hence trended away from the 1 : 1 ratio; and regulatory group III are those in which the expression differences of parental homeologs in a given tetraploid wheat were statistically the same as in the *in silico* parental mix.

We found that in leaves of the synthetic tetraploid wheat (AT2), 21.8%, 21.3% and 56.8% of the 7840 analyzed genes belonged to regulatory groups I, II and III expression patterns, respectively, while in the young inflorescences, 12.4%, 20.5% and 67.1% of the 8900 analyzed genes belonged to the three groups, respectively (Table 3). The relatively higher proportion of regulatory group I genes in leaves than in young inflorescences (*c.* 2 : 1) mirrored the significantly reduced homeolog expression bias in leaves but not inflorescences of AT2 (Fig. 2). The significantly higher proportions of genes belonging to regulatory group III in both tissues of AT2 suggest a strong effect of common

Table 2 Expression pattern of the homeolog in synthetic tetraploid wheat ($S^{l}S^{l}AA$, AT2), *Triticum turgidum* ssp. *dicoccoides* (BBAA, TD), *T. turgidum* ssp. *durum* (BBAA, TTR13) and extracted tetraploid wheat (BBAA, ETW)

Tissue	Category	Expression in parent	Expression in allotetraploid	AT2 (%)	TD (%)	TTR13 (%)	ETW (%)
Leaf	Bias group I	B(S) = A	B(S) = A	1797 (23.0)	1526 (19.6)	1524 (19.7)	1408 (18.1)
		B(S) > A	B(S) > A	1217 (15.6)	870 (11.2)	858 (11.1)	799 (10.3)
		B(S) < A	B(S) < A	1377 (17.7)	1046 (13.5)	1040 (13.4)	1145 (14.7)
	Bias group II	B(S) > A	B(S) = A	755 (9.7)	826 (10.6)	835 (10.8)	795 (10.2)
		B(S) < A	B(S) = A	880 (11.3)	887 (11.4)	904 (11.7)	780 (10.0)
	Bias group III	B(S) = A	B(S) > A	581 (7.5)	661 (8.5)	662 (8.5)	671 (8.6)
		B(S) = A	B(S) < A	592 (7.6)	758 (9.8)	753 (9.7)	894 (11.5)
		B(S) > A	B(S) < A	329 (4.2)	602 (7.7)	601 (7.7)	700 (9.0)
		B(S) < A	B(S) > A	270 (3.5)	595 (7.7)	578 (7.5)	594 (7.6)
	Total number of genes			7798 (100)	7771 (100)	7755 (100)	7786 (100)
YI	Bias group I	B(S) = A	B(S) = A	2140 (24.0)	1562 (17.5)	1594 (17.9)	1248 (14.2)
		B(S) > A	B(S) > A	2096 (23.5)	1167 (13.1)	1195 (13.4)	1045 (11.9)
		B(S) < A	B(S) < A	1392 (15.6)	1150 (12.9)	1128 (12.7)	1281 (14.5)
	Bias group II	B(S) > A	B(S) = A	920 (10.3)	1030 (11.6)	1042 (11.7)	797 (9.0)
		B(S) < A	B(S) = A	571 (6.4)	567 (6.4)	590 (6.6)	443 (5.0)
	Bias group III	B(S) = A	B(S) > A	631 (7.1)	722 (8.1)	722 (8.1)	674 (7.6)
		B(S) = A	B(S) < A	772 (8.7)	1255 (14.1)	1227 (13.8)	1567 (17.8)
		B(S) > A	B(S) < A	262 (2.9)	1079 (12.1)	1045 (11.7)	1410 (16.0)
		B(S) < A	B(S) > A	129 (1.4)	375 (4.2)	372 (4.2)	353 (4.0)
	Total number of genes			8913 (100)	8907 (100)	8915 (100)	8818 (100)

YI, young inflorescence.

trans-regulation in both subgenomes in the newly synthesized tetraploid wheat. Compared with AT2, the proportions of regulatory group III genes of both leaves and young inflorescences (ranging from 36.6% in young inflorescences of ETW to 45.7% in leaves of TTR13; Table 3) were significantly reduced in all the natural tetraploid wheats (Fisher's exact test, $P < 2.2e{-}16$); concomitantly, the proportions of regulatory groups I and II genes of both leaf and young inflorescence (ranging from 54.2% in leaf of TTR13 to 63.4% in young inflorescence of ETW, Table 3) were significantly increased in all the natural tetraploid wheats (Fisher's exact test, $P < 2.2e{-}16$) compared with AT2. These results suggest that, to a large extent, the initially strong convergent regulation on subgenome expression following allotetraploidization has been replaced by divergent regulation in the course of evolution and domestication. Notably, the relative proportions of the regulatory group I and II genes were significantly different between the two tissues in all the tetraploid wheats (including AT2). Specifically, while in leaves these ratios were close to 1 : 1 (exact binomial test, $P > 0.4$), in young inflorescences they were close to or lower than 1 : 2 (Table 3), and the two ratios were significantly different (exact binomial test, $P < 2.2e{-}16$). Notably, this aspect of between-tissue difference was most pronounced in ETW (Table 3).

Genes manifesting homeolog-biased expression in leaves and young inflorescences of tetraploid wheat showed different biological functions

To explore whether the homeolog equal- or biased-expressing genes in the various tetraploid wheats might have different biological functions or are involved in specific biological processes, we performed a gene ontology slim (GOslim) analysis for the three gene categories (i.e. A-biased, B-biased, not biased) in leaves and young inflorescences. We found that in leaves, A-biased genes were significantly enriched in several cellular components, such as plastid and plasma membrane, and were involved in catabolic and some metabolic processes; B-biased genes were specifically enriched in cytosol and responses to stress or stimuli; and unbiased genes showed little enrichment for a specific biological process. In young inflorescences, all three gene categories showed hardly any enrichment, suggesting the greater number of affected genes are involved in diverse biological functions (Table S5). Of the homeolog-biased expressing genes, the *Q*/*q* gene is of particular interest. We found that the wild-type allele of this gene, that is, *q*, showed relative lower expression levels (based on RNA-seq reads) in leaves of *T. urartu* (genome A) and in young inflorescences of *A. longissima* (genome S^l) (Table S6). In the synthetic tetraploid wheat (AT2), the homeolog biases were eliminated, possibly as a result of the globally reduced homeolog expression divergence (Fig. 2). In the wild tetraploid wheat (TD), the *q* allele showed A-subgenome biased expression in both leaves and young inflorescences; this trend was conserved but obviously augmented in both the domesticated tetraploid wheat and extracted tetraploid wheat (Table S6), probably as a result of mutation of the *q* allele of the A subgenome to *Q*, which showed higher expression than *q* (Simons *et al.*, 2006; Zhang *et al.*, 2011).

Table 3 Summary of the numbers of genes belonging to each of the regulatory groups and their relative proportions in synthetic tetraploid wheat (S^lS^lAA, AT2), *Triticum turgidum* ssp. *dicoccoides* (BBAA, TD), *T. turgidum* ssp. *durum* (BBAA, TTR13) and extracted tetraploid wheat (BBAA, ETW)

Tissue	Regulatory group*	AT2 (%)	TD (%)	TTR13 (%)	ETW (%)
Leaf	Group I	1701 (21.8)	2156 (27.6)	2103 (27.1)	2045 (26.3)
	Group II	1665 (21.3)	2155 (27.6)	2105 (27.1)	2372 (30.5)
	Group III	4432 (56.8)	3494 (44.7)	3547 (45.7)	3369 (43.3)
YI	Group I	1101 (12.4)	1607 (18.0)	1635 (18.3)	1513 (17.2)
	Group II	1831 (20.5)	3483 (39.1)	3390 (38.0)	4073 (46.2)
	Group III	5981 (67.1)	3817 (42.8)	3890 (43.6)	3232 (36.6)

*See details in the Materials and Methods section.
YI, young inflorescence.

Discussion

The impacts of hybridization and WGD on the evolution and diversification of plants have been increasingly recognized (Rieseberg & Willis, 2007; Soltis *et al.*, 2009, 2012; Matsushita *et al.*, 2012; Abbott *et al.*, 2013; Song & Chen, 2015). Previous studies from diverse plant taxa have shown that changes in genome structure and gene expression often result in subgenome and tissue specificity in the allopolyploids (Flagel & Wendel, 2010; Buggs *et al.*, 2011). In fact, a recent finding regarding allopolyploid genome evolution is that the two or more subgenomes constituting a given allopolyploid are asymmetric in their contributions both structurally and functionally; this feature has been suggested to play an important role in the final outcomes of diploidization – a universal evolutionary fate of all polyploid organisms (Doyle *et al.*, 2008; Birchler, 2012; Freeling *et al.*, 2012; Roulin *et al.*, 2012). Even in structurally *bona fide* neoallopolyploid species, functional subgenomic asymmetry is evident. For example, in polyploid wheat, several studies based on morphological features have suggested that subgenomic asymmetry is common and relevant to its enhanced growth vigor and adaptation relative to its diploid or lower-level polyploid parents (Li *et al.*, 2014, 2015b). It was also reported recently that at the hexaploid level (common wheat), the D subgenome was dominant over both the A and B subgenomes, while at the tetraploid level (*durum* wheat) the A subgenome was dominant over the B subgenome in genomic stability (Pont *et al.*, 2013; Li *et al.*, 2014). However, so far no study has systemically investigated transcriptome asymmetry in polyploid wheat at either the tetraploid or hexaploid level.

Here, we assessed the relative global ortholog expression levels of two tissues (leaves and young inflorescences) between two putative diploid progenitors of natural tetraploid wheats (TD, TTR13 and ETW) and exact parents of the synthetic tetraploid wheat (AT2). Our results showed that > 60% of the orthologs from the two diploid species, *T. urartu* and *A. longissima*, are differentially expressed in each of the two tissues. Moreover, the two diploid species bear clear tissue specificity; overall ortholog expression levels in leaves and young inflorescences are

reciprocally higher in one species than in the other. It has been proposed that natural selection has profound effects on the evolution of gene expression both within and between species (Gilad *et al.*, 2006; Romero *et al.*, 2012). Under this scenario, we suspect that the ortholog differential expression between the two diploid species might have evolved as a result of differential adaptation during the evolutionary process. We noted that differential expression between the diploid parents has profound impacts on the homeolog expression repatterning in the resulting allotetraploids. For example, in the synthetic allotetraploid wheat (AT2), although global expression difference between the A- and S-subgenome homeologs (i.e. subgenome transcriptome expression divergence) was significantly attenuated in leaves and augmented in young inflorescences compared with the *in silico* parental mix, the overall higher homeolog expression levels by subgenome A (from *T. urartu*) in leaves and by subgenome S (from *A. longissima*) in young inflorescences were largely maintained (Table 1; Fig. 3). This suggests that the attenuation or augmentation of genome-wide homeolog expression immediately following allopolyploidization in AT2 has been largely determined by the original parental homeolog transcript stoichiometry. Therefore, the transcriptomic asymmetry (i.e. overall expression level dominance by one subgenome over the other for a given tissue) observed in AT2 is primarily decided by parental legacy, and superimposed on this, it was further modified proportionally by transcriptome shock (Hegarty *et al.*, 2006; Buggs *et al.*, 2011). Using microarray analysis, we found that formation of AT2 is indeed associated with a strong transcriptome shock at total gene expression level (without distinguishing between homeologs), which disrupts the intrinsic expression patterns of parental expression-conserved genes and is manifested as transgressive expression in AT2 compared with the parental mix (H. Zhang, X. Gou, A. Zhang, X. Wang, Y. Dong, L. Li & B. Liu, unpublished). Clearly, the strong transcriptome shock-induced disruption and remodeling of gene expression is responsible for the observed changes in the spectra of global homeolog expression divergences in AT2 (Fig. 2). Importantly, as also discussed earlier, these changes in homeolog expression spectra have been largely proportional to the original relative parental homeolog transcript stoichiometry (i.e. parental legacy). Although the importance of parental legacy in impacting homeolog expression in the resulting allopolyploids they parented was first proposed by Leslie D. Gottlieb decades ago (Roose & Gottlieb, 1976; Gottlieb, 2003), the issue has not been examined in detail until recently (Buggs *et al.*, 2014; Soltis *et al.*, 2014). This is at least in part the result of the difficulties in unequivocally distinguishing legacy vs novelty when differential homeolog expression is observed in naturally formed allopolyploids, because the exact parental individuals and their inherent heterozygosity probably can never be ascertained under natural settings. Nevertheless, the issue has been elegantly addressed in *Tragopogon mirus* and *Tragopogon miscellus*, two naturally formed allotetraploid species in of *Tragopogon* (Asteraceae), which are only *c.* 40 generations away from the allotetraploidization events leading to their formation, and with exact diploid progenitor populations still extant (Soltis *et al.*, 2014). It was found that parental legacy indeed plays a significant role in homeolog expression in the allotetraploid species, but novel expression patterns generated upon allotetraploidization and evolved after speciation apparently account for the majority of differential homeolog expression (Tate *et al.*, 2006; Buggs *et al.*, 2009, 2010a, 2011). Because we used a synthetic allotetraploid wheat (AT2) with exact parental individuals which themselves had been extensively selfed before constructing the allotetraploids, the issue of parental legacy can be accurately assessed, and our observations are fully in line with the findings in *Tragopogon* (Tate *et al.*, 2006; Buggs *et al.*, 2009, 2010a, 2011).

It should be noted that, in the young inflorescences, the patterns of transcriptome asymmetry in the tetraploid wheats are associated with their morphological asymmetry. For example, it was observed that the inflorescence morphology of tetraploid wheat (BBAA) is under the control of the A subgenome (Feldman *et al.*, 2012). Indeed, our results revealed that the A subgenome showed overall expression level dominance in all three natural tetraploids. More importantly, we found that the inflorescence morphologies of synthetic tetraploid AT2 resemble those of *A. longissima* than those of *T. urartu*. As expected, the S subgenome showed overall expression level dominance over the A subgenome in the synthetic tetraploid AT2. In fact, subgenomic functional asymmetry has been shown or suggested to play important roles at both the initial stage of establishment and the evolutionary success of allopolyploids relative to their diploid parents. In cotton, for example, it was found that homeolog biased expression is related to the enhanced ability of tetraploids to cope with various abiotic stresses than diploid cotton species (Liu & Adams, 2007; Bardil *et al.*, 2011; Dong & Adams, 2011; Kim & Chen, 2011). In wheat, it was also proposed that structural and functional subgenomic asymmetry may play a critical role in rapid cytological and genetic diploidization in newly formed allopolyploid wheat, attributes that are critical to their successful establishment (Feldman *et al.*, 2012). As mentioned earlier, phenotypic asymmetry was evident at both the tetra- and hexaploid levels in wheat (Peng *et al.*, 2003; Abbo *et al.*, 2006; Shitsukawa *et al.*, 2007). For instance, the significant contribution by the D subgenome to increased salt tolerance in common wheat was found to exert the effects immediately upon allohexaploid formation, and the effects were then largely conserved in the course of domestication of hexaploid common wheat (Yang *et al.*, 2014).

It is expected that the rapidly emerging genomic or transcriptomic asymmetry following allopolyploid speciation will continue to evolve under natural and/or human selections, as clearly documented in cotton (Renny-Byfield & Wendel, 2014) and *Tragopogon* (Buggs *et al.*, 2011). Here, we compared the global homeolog expression patterns between the synthetic and natural tetraploid wheats. We found that, relative to the synthetic allotetraploid wheat (AT2), the spectra of homeolog expression divergence were significantly enlarged in all three types of natural tetraploid wheats (TD, TTR13 and ETW) in both leaves and young inflorescences, although the degrees of both the enlarged homeolog expression difference and subgenome transcriptome asymmetry (the overall homeolog-specific expression levels by the A vs B subgenomes) were highly variable depending on both

genotype and tissue. Notably, we found that the synthetic (AT2) and natural (TD, TTR13 and ETW) tetraploid wheats showed different overall biased expression levels in the young inflorescence tissue: while AT2 showed overall higher homeolog expression levels of subgenome A in leaves and subgenome S in young inflorescences, both tissues of all three natural tetraploid wheats showed subgenome A expression dominance. One possible explanation for this observation might be that natural and/or artificial selections have favored higher A-subgenome expression, which is consistent with previous suggestions (Feldman *et al.*, 2012; Pont *et al.*, 2013; Li *et al.*, 2014). It thus appears that there are two broad phases of subgenome transcriptome asymmetry, phase one being conditioned by parental legacy and modified by transcriptome shock and phase two being cumulative modifications of subgenome expression during the evolutionary process under natural and/or human selections; these two phases can be underpinned by dramatically different regulatory modules. Taken together, our findings, coupled with previous observations in hexaploid wheat and other plant taxa, suggested that the status of original parental ortholog expression have profound effects on the subgenome asymmetry of the resulting allopolyploids, and which might have played crucial roles at both the initial stages of establishment and the longer-term evolutionary process of the allopolyploid species.

Several studies in cotton indicated that the domestication process has dramatically rewired the fiber transcriptome and led to reprogrammed resource allocation toward increased fiber growth (Chaudhary *et al.*, 2008; Yoo & Wendel, 2014). In our case, however, we found that the overall homeolog expression levels of the leaves and young inflorescences showed no significant differences between the wild (TD) and cultivated (TTR13) tetraploid wheats. These results are consistent with previous observations in hexaploid common wheat in which the pattern of homeolog expression is highly conserved during the domestication process (Li *et al.*, 2014). However, we noted that the number and extent of extremely biased expression genes are obviously increased in the domesticated tetraploid wheat (TTR13) compared with the wild tetraploid wheat (TD). In particular, the extents of overall subgenome expression dominance (by subgenome A in both tissues) were greatest in the ETW, with a genome composition (BBAA) basically identical to that of the BBAA subgenomes of hexaploid common wheat (Zhang *et al.*, 2014; Liu *et al.*, 2015), suggesting a strong and distinct impact of domestication under the allohexaploid trajectory on the evolution of subgenome gene expression. Furthermore, globally similar transcriptomes do not rule out the possibility that expression of specific genes or pathways has been targeted by domestication. For example, the four homeologs of the MADS-box genes (i.e. *WLHS1*, *PI*, *AP3* and *AG*) are differentially regulated by genetic and epigenetic mechanisms in common wheat (Shitsukawa *et al.*, 2007; Li *et al.*, 2014). Our results from the transcriptome profiling also revealed that the *Q*/*q* gene exhibited A subgenome-biased expression in both the leaf and young inflorescence tissues of all the natural tetraploid wheats, including the wild ssp. *dicoccoides* (TD)(Table S6). Previous studies have established that the wild-type *q* allele in diploid and wild tetraploid wheats confers a speltoid spike phenotype characterized by a spear-shaped spike with an elongated rachis, while the mutant *Q* allele (which showed a higher expression level than *q*) conditions the free-threshing character and square spike phenotype in domesticated wheats (Simons *et al.*, 2006; Zhang *et al.*, 2011). Our results thus indicated that the A-subgenome *q* allele already showed a higher expression than the B-subgenome allele before its mutation to *Q*, and artificial selection acting on the *Q* allele further increases the A-subgenome biased expression in the domesticated tetraploid wheat (TTR13) and common wheat (being reflected by ETW) (Table S6). In addition, we noted that eight pentatricopeptide repeat (PPR) protein-coding genes showed different extremely biased expression patterns in the four tetraploid wheat genotypes studied, especially in the extracted tetraploid ETW. It has been demonstrated that the PPR protein family has profound effects on organelle biogenesis, organellar gene expression and cytoplasmic male sterility (Bentolila *et al.*, 2002; Alice & Small, 2014). The polyploid wheats reunited two or three distinct nuclear genomes but only inherited the cytoplasm from the B-subgenome diploid progenitor (Gornicki *et al.*, 2014; Li *et al.*, 2015a). Together, it appears that nucleus–organelle interactions upon and after polyploidization might have played a part in remodeling the homeolog expression patterns, and artificial selection has acted primarily on a small number of domestication-related genes and pathways rather than on a genome-wide scale in wheat. In addition, the observations of distinct homeolog expression patterns between the domesticated natural tetraploid wheat (TTR13) and ETW confirmed the earlier proposition based on microarray analysis that there exist distinct differences between the tetraploid and hexaploid evolutionary trajectories on the evolution of subgenome transcriptome expression (Zhang *et al.*, 2014), an issue that merits further investigations.

Acknowledgements

We are grateful to Prof. Moshe Feldman of the Weizmann Institute of Science, Israel, for providing the initial seeds for the various wheat genotypes. We also thank Prof. Kenneth Olsen, Sara Wright and all other members of the Olsen laboratory for their constructive comments that helped to improve the manuscript. This work was supported by the National Natural Science Foundation of China (31290210) and the Program for Introducing Talents to Universities (B07017).

Author contributions

B.L. and L.L. planned and designed the research. X.W., H.Z., Y.L. and Z.Z. performed experiments and analyzed data. B.L. and L.L. wrote the manuscript.

References

Abbo S, Gopher A, Peleg Z, Saranga Y, Fahima T, Salamini F, Lev-Yadun S. 2006. The ripples of "The Big (agricultural) Bang": the spread of early wheat cultivation. *Genome* **49**: 861–863.

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJ, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R *et al.* 2013. Hybridization and speciation. *Journal of Evolutionary Biology* **26**: 229–246.

Adams KL, Wendel JF. 2005. Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology* **8**: 135–141.

Akhunova AR, Matniyazov RT, Liang H, Akhunov ED. 2010. Homoeolog-specific transcriptional bias in allopolyploid wheat. *BMC Genomics* **11**: 505.

Alice B, Small I. 2014. Pentatricopeptide repeat proteins in plants. *Annual Review of Plant Biology* **65**: 415–442.

Bardil A, de Almeida JD, Combes MC, Lashermes P, Bertrand B. 2011. Genomic expression dominance in the natural allopolyploid *Coffea arabica* is massively affected by growth temperature. *New Phytologist* **192**: 760–774.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)* **57**: 289–300.

Bentolila S, Alfonso AA, Hanson MR. 2002. A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. *Proceedings of the National Academy of Sciences, USA* **99**: 10887–10892.

Birchler JA. 2012. Insights from paleogenomic and population studies into the consequences of dosage sensitive gene expression in plants. *Current Opinion in Plant Biology* **15**: 544–548.

Buggs RJ, Chamala S, Wu W, Gao L, May GD, Schnable PS, Soltis DE, Soltis PS, Barbazuk WB. 2010a. Characterization of duplicate gene evolution in the recent natural allopolyploid *Tragopogon miscellus* by next-generation sequencing and Sequenom iPLEX MassARRAY genotyping. *Molecular Ecology* **19**: 132–146.

Buggs RJ, Doust AN, Tate JA, Koh J, Soltis K, Feltus FA, Paterson AH, Soltis PS, Soltis DE. 2009. Gene loss and silencing in *Tragopogon miscellus* (Asteraceae): comparison of natural and synthetic allotetraploids. *Heredity* **103**: 73–81.

Buggs RJ, Elliott NM, Zhang L, Koh J, Viccini LF, Soltis DE, Soltis PS. 2010b. Tissue-specific silencing of homoeologs in natural populations of the recent allopolyploid *Tragopogon mirus*. *New Phytologist* **186**: 175–183.

Buggs RJ, Wendel JF, Doyle JJ, Soltis DE, Soltis PS, Coate JE. 2014. The legacy of diploid progenitors in allopolyploid gene expression patterns. *Philosophical Transactions of the Royal Society B: Biological* **369**: 20130354.

Buggs RJ, Zhang L, Miles N, Tate JA, Gao L, Wei W, Schnable PS, Barbazuk WB, Soltis PS, Soltis DE. 2011. Transcriptomic shock generates evolutionary novelty in a newly formed, natural allopolyploid plant. *Current Biology* **21**: 551–556.

Chague V, Just J, Mestiri I, Balzergue S, Tanguy AM, Huneau C, Huteau V, Belcram H, Coriton O, Jahier J *et al.* 2010. Genome-wide gene expression changes in genetically stable synthetic and natural wheat allohexaploids. *New Phytologist* **187**: 1181–1194.

Chang PL, Dilkes BP, McMahon M, Comai L, Nuzhdin SV. 2010. Homoeolog-specific retention and use in allotetraploid *Arabidopsis suecica* depends on parent of origin and network partners. *Genome Biology* **11**: R125.

Chaudhary B, Hovav R, Rapp R, Verma N, Udall JA, Wendel JF. 2008. Global analysis of gene expression in cotton fibers from wild and domesticated *Gossypium barbadense*. *Evolution & Development* **10**: 567–582.

Chelaifa H, Chague V, Chalabi S, Mestiri I, Arnaud D, Deffains D, Lu Y, Belcram H, Huteau V, Chiquet J *et al.* 2013. Prevalence of gene expression additivity in genetically stable wheat allohexaploids. *New Phytologist* **197**: 730–736.

Chelaifa H, Monnier A, Ainouche M. 2010. Transcriptomic changes following recent natural hybridization and allopolyploidy in the salt marsh species *Spartina* x *townsendii* and *Spartina anglica* (Poaceae). *New Phytologist* **186**: 161–174.

Chen ZJ. 2007. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annual Review of Plant Physiology* **58**: 377–406.

Comai L. 2005. The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* **6**: 836–846.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674–3676.

Cox MP, Dong T, Shen G, Dalvi Y, Scott DB, Ganley AR. 2014. An interspecific fungal hybrid reveals cross-kingdom rules for allopolyploid gene expression patterns. *PLoS Genetics* **10**: e1004180.

Dobesova E, Malinska H, Matyasek R, Leitch AR, Soltis DE, Soltis PS, Kovarik A. 2015. Silenced rRNA genes are activated and substitute for partially eliminated active homeologs in the recently formed allotetraploid, *Tragopogon mirus* (Asteraceae). *Heredity* **114**: 356–365.

Dong S, Adams KL. 2011. Differential contributions to the transcriptome of duplicated genes in response to abiotic stresses in natural and synthetic polyploids. *New Phytologist* **190**: 1045–1057.

Doyle JJ, Flagel LE, Paterson AH, Rapp RA, Soltis DE, Soltis PS, Wendel JF. 2008. Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics* **42**: 443–461.

Dvořák J. 1976. The relationship between the genome of *Triticum urartu* and the A and B genomes of *Triticum aestivum*. *Canadian Journal of Genetics and Cytology* **18**: 371–377.

Dvorak J, Akhunov ED. 2005. Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the Aegilops-Triticum alliance. *Genetics* **171**: 323–332.

Eilam T, Anikster Y, Millet E, Manisterski J, Feldman M. 2008. Nuclear DNA amount and genome downsizing in natural and synthetic allopolyploids of the genera *Aegilops* and *Triticum*. *Genome* **51**: 616–627.

Feldman M, Levy AA. 2005. Allopolyploidy – a shaping force in the evolution of wheat genomes. *Cytogenetic and Genome Research* **109**: 250–258.

Feldman M, Levy AA, Fahima T, Korol A. 2012. Genomic asymmetry in allopolyploid plants: wheat as a model. *Journal of Experimental Botany* **63**: 5045–5059.

Feldman M, Lupton FGH, Miller TE. 1995. Wheats. In: Smartt J, Simmonds NW, eds. *Evolution of crop plants, 2nd edn*. London, UK: Longman Scientific, 184–192.

Flagel LE, Wendel JF. 2010. Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation. *New Phytologist* **186**: 184–193.

Freeling M, Woodhouse MR, Subramaniam S, Turco G, Lisch D, Schnable JC. 2012. Fractionation mutagenesis and similar consequences of mechanisms removing dispensable or less-expressed DNA in plants. *Current Opinion in Plant Biology* **15**: 131–139.

Gilad Y, Oshlack A, Rifkin SA. 2006. Natural selection on gene expression. *Trends in Genetics* **22**: 456–461.

Gill B, Friebe B. 2002. Cytogenetics, phylogeny and evolution of cultivated wheats. In: Rajaram S, Curtis BC, Gómez Macpherson H, eds. *Bread wheat: improvement and production. FAO Plant production and protection series no. 30.* Rome, Italy: Food Agriculture Organization, 71–88.

Gordon A, Hannon G. 2012. *FASTX-toolkit, FASTQ/A short-reads pre-processing tools.* [WWW document] URL http://hannonlab.cshl.edu/fastx_toolkit/ [accessed 2 February 2010]

Gornicki P, Zhu H, Wang J, Challa GS, Zhang Z, Gill BS, Li W. 2014. The chloroplast view of the evolution of polyploid wheat. *New Phytologist* **204**: 704–714.

Gottlieb L. 2003. Plant polyploidy: gene expression and genetic redundancy. *Heredity* **91**: 91–92.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**: 644–652.

Guo X, Han F. 2014. Asymmetric epigenetic modification and elimination of rDNA sequences by polyploidization in wheat. *Plant Cell* **26**: 4311–4327.

Hegarty MJ, Barker GL, Wilson ID, Abbott RJ, Edwards KJ, Hiscock SJ. 2006. Transcriptome shock after interspecific hybridization in senecio is ameliorated by genome duplication. *Current Biology* **16**: 1652–1659.

Hegarty MJ, Hiscock SJ. 2008. Genomic clues to the evolutionary success of polyploid plants. *Current Biology* **18**: R435–R444.

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, 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. *Proceedings of the National Academy of Sciences, USA* **99**: 8133–8138.

Jackson S, Chen ZJ. 2010. Genomic and expression plasticity of polyploidy. *Current Opinion in Plant Biology* **13**: 153–159.

Kent WJ. 2002. BLAT-the BLAST-like alignment tool. *Genome Research* **12**: 656–664.

Kerber E. 1964. Wheat: reconstitution of the tetraploid component (AABB) of hexaploids. *Science* **143**: 253–255.

Kim E-D, Chen ZJ. 2011. Unstable transcripts in Arabidopsis allotetraploids are associated with nonadditive gene expression in response to abiotic and biotic stresses. *PLoS ONE* **6**: e24251.

Koh J, Chen S, Zhu N, Yu F, Soltis PS, Soltis DE. 2012. Comparative proteomics of the recently and recurrently formed natural allopolyploid *Tragopogon mirus* (Asteraceae) and its parents. *New Phytologist* **196**: 292–305.

Leach LJ, Belfield EJ, Jiang C, Brown C, Mithani A, Harberd NP. 2014. Patterns of homoeologous gene expression shown by RNA sequencing in hexaploid bread wheat. *BMC Genomics* **15**: 276.

Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. 2014. UpSet: visualization of intersecting sets. *Visualization and Computer Graphics, IEEE Transactions on* **20**: 1983–1992.

Li A, Liu D, Wu J, Zhao X, Hao M, Geng S, Yan J, Jiang X, Zhang L, Wu J *et al.* 2014. mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell* **26**: 1878–1900.

Li AL, Geng SF, Zhang LQ, Liu DC, Mao L. 2015b. Making the bread: insights from newly synthesized allohexaploid wheat. *Molecular Plant* **8**: 847–859.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.

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

Li LF, Liu B, Olsen KM, Wendel JF. 2015a. A re-evaluation of the homoploid hybrid origin of *Aegilops tauschii*, the donor of the wheat D-subgenome. *New Phytologist* **208**: 4–8.

Lim KY, Soltis DE, Soltis PS, Tate J, Matyasek R, Srubarova H, Kovarik A, Pires JC, Xiong Z, Leitch AR. 2008. Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (Asteraceae). *PLoS ONE* **3**: e3353.

Liu C, Yang X, Zhang H, Wang X, Zhang Z, Bian Y, Zhu B, Dong Y, Liu B. 2015. Genetic and epigenetic modifications to the BBAA component of common wheat during its evolutionary history at the hexaploid level. *Plant Molecular Biology* **88**: 53–64.

Liu Z, Adams KL. 2007. Expression partitioning between genes duplicated by polyploidy under abiotic stress and during organ development. *Current Biology* **17**: 1669–1674.

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

Malinska H, Tate JA, Matyasek R, Leitch AR, Soltis DE, Soltis PS, Kovarik A. 2010. Similar patterns of rDNA evolution in synthetic and recently formed natural populations of *Tragopogon* (Asteraceae) allotetraploids. *BMC Evolutionary Biology* **10**: 291.

Malinska H, Tate JA, Mavrodiev E, Matyasek R, Lim KY, Leitch AR, Soltis DE, Soltis PS, Kovarik A. 2011. Ribosomal RNA genes evolution in *Tragopogon*: a story of new and old world allotetraploids and the synthetic lines. *Taxon* **60**: 348–354.

Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, Jakobsen KS, Wulff BB, Steuernagel B, Mayer KF, Olsen OA. 2014. Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **345**: 1250092.

Matsushita SC, Tyagi AP, Thornton GM, Pires JC, Madlung A. 2012. Allopolyploidization lays the foundation for evolution of distinct populations: evidence from analysis of synthetic *Arabidopsis* allohexaploids. *Genetics* **191**: 535–547.

Otto SP. 2007. The evolutionary consequences of polyploidy. *Cell* **131**: 452–462.

Peng J, Ronin Y, Fahima T, Roder MS, Li Y, Nevo E, Korol A. 2003. Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proceedings of the National Academy of Sciences, USA* **100**: 2489–2494.

Pfeifer M, Kugler KG, Sandve SR, Zhan B, Rudi H, Hvidsten TR, International Wheat Genome Sequencing Consortium, Mayer KF, Olsen OA. 2014. Genome interplay in the grain transcriptome of hexaploid bread wheat. *Science* 345: 1250091.

Pont C, Murat F, Confolent C, Balzergue S, Salse J. 2011. RNA-seq in grain unveils fate of neo- and paleopolyploidization events in bread wheat (*Triticum aestivum* L.). *Genome Biology* **12**: R119.

Pont C, Murat F, Guizard S, Flores R, Foucrier S, Bidet Y, Quraishi UM, Alaux M, Dolezel J, Fahima T *et al.* 2013. Wheat syntenome unveils new evidences of contrasted evolutionary plasticity between paleo- and neoduplicated subgenomes. *Plant Journal* **76**: 1030–1044.

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

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

R Development Core Team. 2014. *The R project for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL www.R-project.org [accessed 9 June 2015].

Renny-Byfield S, Wendel JF. 2014. Doubling down on genomes: polyploidy and crop plants. *American Journal of Botany* **101**: 1711–1725.

Rieseberg LH, Willis JH. 2007. Plant speciation. *Science* **317**: 910–914.

Romero IG, Ruvinsky I, Gilad Y. 2012. Comparative studies of gene expression and the evolution of gene regulation. *Nature Reviews Genetics* **13**: 505–516.

Roose M, Gottlieb L. 1976. Genetic and biochemical consequences of polyploidy in *Tragopogon*. *Evolution* **30**: 818–830.

Roulin A, Auer PL, Libault M, Schlueter J, Farmer A, May G, Stacey G, Doerge RW, Jackson SA. 2012. The fate of duplicated genes in a polyploid plant genome. *Plant Journal* **73**: 143–153.

Sehrish T, Symonds VV, Soltis DE, Soltis PS, Tate JA. 2014. Gene silencing via DNA methylation in naturally occurring *Tragopogon miscellus* (Asteraceae) allopolyploids. *BMC Genomics* **15**: 701.

Shaked H, Kashkush K, Ozkan H, Feldman M, Levy AA. 2001. Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* **13**: 1749–1759.

Shcherban AB, Badaeva ED, Amosova AV, Adonina IG, Salina EA. 2008. Genetic and epigenetic changes of rDNA in a synthetic allotetraploid. *Aegilops sharonensis* × *Ae. umbellulata*. *Genome* **51**: 261–271.

Shitsukawa N, Ikari C, Shimada S, Kitagawa S, Sakamoto K, Saito H, Ryuto H, Fukunishi N, Abe T, Takumi S. 2007. The einkorn wheat (*Triticum monococcum*) mutant, maintained vegetative phase, is caused by a deletion in the *VRN1* gene. *Genes & Genetic Systems* **82**: 167–170.

Simons KJ, Fellers JP, Trick HN, Zhang Z, Tai Y-S, Gill BS, Faris JD. 2006. Molecular characterization of the major wheat domestication gene *Q*. *Genetics* **172**: 547–555.

Soltis DE, Buggs RJA, Barbazuk WB, Chamala S, Chester M, Gallagher JP, Schnable PS, Soltis PS. 2012. Rapid and repeated evolution in the early stages of polyploidy: genomic and cytogenetic studies of recent polyploidy in *Tragopogon*. In: Soltis PS, Soltis DE, eds. *Polyploidy and genome evolution*. Berlin, Germany: Springer, 271–292.

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, Depamphilis CW, Wall PK, Soltis PS. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* **96**: 336–348.

Soltis PS, Liu X, Marchant DB, Visger CJ, Soltis DE. 2014. Polyploidy and novelty: Gottlieb's legacy. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**: 20130351.

Soltis DE, Soltis PS, Pires JC, Kovarik A, Tate JA, Mavrodiev E. 2004. Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): cytogenetic, genomic and genetic comparisons. *Biological Journal of the Linnean Society* **82**: 485–501.

Song Q, Chen ZJ. 2015. Epigenetic and developmental regulation in plant polyploids. *Current Opinion in Plant Biology* **24**: 101–109.

Tate JA, Joshi P, Soltis KA, Soltis PS, Soltis DE. 2009. On the road to diploidization? Homoeolog loss in independently formed populations of the allopolyploid *Tragopogon miscellus* (Asteraceae). *BMC Plant Biology* **9**: 80.

Tate JA, Ni Z, Scheen AC, Koh J, Gilbert CA, Lefkowitz D, Chen ZJ, Soltis PS, Soltis DE. 2006. Evolution and expression of homeologous loci in *Tragopogon miscellus* (Asteraceae), a recent and reciprocally formed allopolyploid. *Genetics* **173**: 1599–1611.

Wendel JF. 2000. Genome evolution in polyploids. *Plant Molecular Biology* **42**: 225–249.

Wickham H. 2009. *ggplot2: elegant graphics for data analysis.* New York, NY, USA: Springer.

Willcox G. 1998. Archaeobotanical evidence for the beginnings of agriculture in Southwest Asia. In: Damania AB, Valkoun J, Willcox G, Qualset CO, eds. *The origins of agriculture and crop domestication.* Aleppo, Syria: ICARDA, 25–38.

Xu C, Bai Y, Lin X, Zhao N, Hu L, Gong Z, Wendel JF, Liu B. 2014. Genome-wide disruption of gene expression in allopolyploids but not hybrids of rice subspecies. *Molecular Biology and Evolution* **31**: 1066–1076.

Yang C, Zhao L, Zhang H, Yang Z, Wang H, Wen S, Zhang C, Rustgi S, von Wettstein D, Liu B. 2014. Evolution of physiological responses to salt stress in hexaploid wheat. *Proceedings of the National Academy of Sciences, USA* **111**: 11882–11887.

Yoo MJ, Szadkowski E, Wendel JF. 2013. Homoeolog expression bias and expression level dominance in allopolyploid cotton. *Heredity* **110**: 171–180.

Yoo MJ, Wendel JF. 2014. Comparative evolutionary and developmental dynamics of the cotton (*Gossypium hirsutum*) fiber transcriptome. *PLoS Genetics* **10**: e1004073.

Zhang H, Bian Y, Gou X, Dong Y, Rustgi S, Zhang B, Xu C, Li N, Qi B, Han F. 2013. Intrinsic karyotype stability and gene copy number variations may have laid the foundation for tetraploid wheat formation. *Proceedings of the National Academy of Sciences, USA* **110**: 19466–19471.

Zhang H, Zhu B, Qi B, Gou X, Dong Y, Xu C, Zhang B, Huang W, Liu C, Wang X, et al. 2014. Evolution of the BBAA component of bread wheat during its history at the allohexaploid level. *Plant Cell* **26**: 2761–2776.

Zhang Z, Belcram H, Gornicki P, Charles M, Just J, Huneau C, Magdelenat G, Couloux A, Samain S, Gill BS. 2011. Duplication and partitioning in evolution and function of homoeologous *Q* loci governing domestication characters in polyploid wheat. *Proceedings of the National Academy of Sciences, USA* **108**: 18737–18742.

Zohary D, Feldman M. 1962. Hybridization between amphidiploids and the evolution of polyploids in the wheat (*Aegilops-Triticum*) group. *Evolution* **16**: 44–61.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 The information of Illumina reads obtained from the diploid and tetraploid wheats of this study

Table S2 The detailed information of the assembled contigs from the two diploid species *A. longissima* (S^lS^l, TL05) and *T. urartu* (AA, TMU06)

Table S3 A and B subgenome-specific SNPs identified in each assembled contig

Table S4 Detailed information of the extremely biased expression genes in the leaf and young inflorescence tissues

Table S5 Detailed information of GOslim and the enrichment analysis for the three gene categories (A-biased, B-biased, not biased) in leaves and young inflorescences

Table S6 Expression pattern of *Q* and *q* alleles in leaves and young inflorescences of the five wheat genotypes

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.