ARTICLE

Check for updates

Global transcriptome analysis of allopolyploidization reveals large-scale repression of the D-subgenome in synthetic hexaploid wheat

Akshaya Vasudevan[1,2], Madeleine Lévesque-Lemay[1], Tara Edwards[1] & Sylvie Cloutier [1,2✉]

Synthetic hexaploid wheat (SHW) lines are created as pre-breeding germplasm to diversify the D subgenome of hexaploid wheat and capitalize upon the untapped genetic diversity of the *Aegilops tauschii* gene pool. However, the phenotypes observed in the *Ae. tauschii* parents are not always recovered in the SHW lines, possibly due to inter-subgenome interactions. To elucidate this post-polyploidization genome reprogramming phenomenon, we performed RNA-seq of four SHW lines and their corresponding tetraploid and diploid parents, across ten tissues and three biological replicates. Homoeologue expression bias (HEB) analysis using more than 18,000 triads suggests massive suppression of homoeoalleles of the D subgenome in SHWs. Comparative transcriptome analysis of the whole-genome gene set further corroborated this finding. Alternative splicing analysis of the high-confidence genes indicates an additional layer of complexity where all five splice events are identified, and retained intron is predominant. Homoeologue expression upon resynthesis of hexaploid wheat has implications to the usage and handling of this germplasm in breeding as it relates to capturing the effects of epistatic interaction across subgenomes upon polyploidization. Special considerations must be given to this germplasm in pre-breeding activities to consider the extent of the inter-subgenome interactions on gene expression and their impact on traits for crop improvement.

[1] Agriculture and Agri-Food Canada, Ottawa Research and Development Centre, Ottawa, ON, Canada. [2] Department of Biology, University of Ottawa, Ottawa, ON, Canada. ✉email: sylviej.cloutier@agr.gc.ca

Whole-genome duplication events are one of the prime drivers of speciation[1–3]. The majority of angiosperms have undergone polyploidization in the course of their evolution, and notably, 30% of crop species are deemed to be polyploid based on the extent of duplicated loci in their genome[4]. The grass lineage underwent a minimum of three whole genome duplication events[5], basically making the whole Poaceae family polyploid. Whole genome duplication events are frequently found to be associated with significant evolutionary potential that promotes adaptability to changing environments[6,7]. There are several intriguing questions related to polyploids, from the initial dynamic changes they undergo in the genome for stabilization, until their establishment as discrete populations[8]. Polyploidization creates extensive redundancy within the genome, thereby paving the way for novel alterations that promote the development of new phenotypes and/or adaptation[9].

Bread wheat (*Triticum aestivum* L.), is an allohexaploid crop ($2n = 6x = 42$) comprised of the A, B and D subgenomes. The natural chance hybridizations that occurred ~8000 years ago between cultivated emmer wheat (*Triticum turgidum* L. ssp. *dicoccum*, $2n = 4x = 28$; AABB genome) and Tausch's goatgrass (*Aegilops tauschii* Coss., $2n = 2x = 14$; DD genome), followed by chromosome doubling, resulted in the development of modern bread wheat[10,11]. Presumably, only a limited number of *Ae. tauschii* plants contributed to the evolution of hexaploid wheat, leading to an evolutionary bottleneck called the founder effect[12]; a bottleneck that was further constricted by limited natural flow of genetic variation from diploid to hexaploid species[13,14] and subsequent domestication and breeding.

Wild relatives of crop species are a source of genetic diversity for multiple agriculturally important traits. However, there are certain limitations to their usage, including linkage drag, infertility and poor cross-compatibility[15]. Pre-breeding is a systematic approach that aims to recapture the genetic diversity of crop wild relatives and deploy it into breeding programs. Traditionally, the desired genomic segments from wild crop species are introgressed into elite cultivar backgrounds by repeated backcrossing with the support of state-of-the-art whole-genome genotyping tools for foreground and background selections. SHWs generated from the artificial hybridization between *T. turgidum* ssp. *durum* (AABB) or other subspecies and *Ae. tauschii* (DD), followed by chromosome doubling, serve as effective pre-breeding populations to broaden the genetic diversity of the D subgenome of hexaploid wheat. SHWs developed by the International Maize and Wheat Improvement Center (CIMMYT, Mexico)[16,17] have been extensively utilized in wheat genetic diversity enrichment programs in several countries. Later, other research institutes and wheat improvement programs across the globe also produced SHWs using different sources of *Ae. tauschii* accessions to serve as base populations for commercial wheat breeding.

Ae. tauschii is a proven source of resistance to a broad range of biotic and abiotic stresses[18]. However, parental phenotypes are not always effectively recovered in the synthetic hexaploid lines, possibly due to the dynamic genetic and epigenetic changes and the consequent variation in gene expression[19]. For instance, studies spanning the past two decades have shown that resistance to stem rust pathogen (*Puccinia graminis* f. sp. *tritici*) is suppressed in the hexaploid state by Med15, a component of the Mediator complex encoded by the D subgenome[20,21]. Upon polyploidization, the D homoeologue of the *High-Affinity K+ Transporter 1;5* (*HKT1;5*) shows reduced expression, when compared to diploid parental levels, but regains the diploid parent native expression level under salt stress condition[22]. In certain phenotypes, like grain width, each of the homoeologues is associated differently with the trait of interest and their combined levels of expression determine the resultant phenotype[23].

The alteration across multiple layers of gene regulation upon polyploidization is also considered to improve adaptability of allohexaploid bread wheat when compared to its tetraploid and diploid parents[24,25]. The genetic and epigenetic alterations reported upon polyploidization in wheat are similar to other polyploid plant species. At the structural level, apart from homoeologous exchanges between genomes, both loss and amplification of repetitive DNA due to perceived genome stress have been observed[26]. The recent whole genome sequencing of bread wheat sheds light on the transposon insertions and deletions in the intergenic regions of the A, B and D subgenomes, although the fraction of each transposon family among the homoeologous genomes does not significantly vary[27]. At the epigenetic level, in addition to changes in cytosine methylation status of the DNA, variability in histone methylation[28,29], euchromatin and heterochromatin patterns across chromosomes and prominent differences in the levels of epigenetic regulators like small RNAs were also observed[30,31]. These epigenetic marks on the regulatory elements of the genome modulate gene expression levels in hexaploid bread wheat[32].

The effect of some of these genomic and epigenomic changes occurring upon polyploidization are reflected in the transcriptome. Understanding the pattern of expression changes between the parental ploidy backgrounds and hexaploid wheat provides valuable information to inform interspecific crosses and assist in predicting the phenotypes recovered in the progenies. In this study, we used four SHW lines and their corresponding tetraploid (*T. turgidum*) and diploid (*Ae. tauschii*) parents for a comparative evaluation of transcriptome dynamics. We focused on unraveling the changes in expression patterns of homoeologous genes between the subgenomes in parental (2x and 4x) and synthetic hybrid backgrounds (6x). In addition to quantitative differences, the subtle qualitative differences in transcripts arising from alternative splicing events were also determined.

Results

HEB in SHW lines. In this experiment, RNA-seq data was generated from four SHW lines and their two diploid and two tetraploid parents (Supplementary Table 1), from three biological replicates across ten tissues (Supplementary Fig. 1) for a total of 240 samples. After pre-processing of the sequencing reads for the 240 samples, we obtained a total of 4.9 B reads corresponding to an average of 20.4 M reads per sample, ranging from 4.8 to 70.7 M. The number of raw reads and processed reads for each sample was compiled (Supplementary Tables 2, 3). Approximately 85% of the total reads were uniquely mapped to the Chinese Spring IWGSC RefSeq v2.1[33], ~12% mapped to multiple loci, and ~3% did not map at all (Supplementary Table 4).

In order to understand the global pattern of gene expression dynamics upon polyploidization, specifically, HEB, and compare them between SHWs and their parental state, genes (homoeologues) belonging to triads, i.e., with one copy in all three subgenomes (18,357 triads[34]), were subjected to the likelihood ratio test[35] (LRT) computation. Here, the differences in gene length between the homoeologues are considered. The null hypothesis, i.e., no expression differences between the homoeologues, was tested for all comparisons. In the SHW lines the AB vs. D homoeologues were compared. To test for the parental expression state, in-silico SHW-like scenarios with corresponding tetraploid and diploid parental expression levels were constituted. Expression biases, predominantly towards the D subgenome in parental level of expression, were observed across all ten tissues and SHWs (Fig. 1). The number of triads showing significant bias towards the subgenome contributed by the diploid (D) parent were drastically reduced in the corresponding SHW lines across all
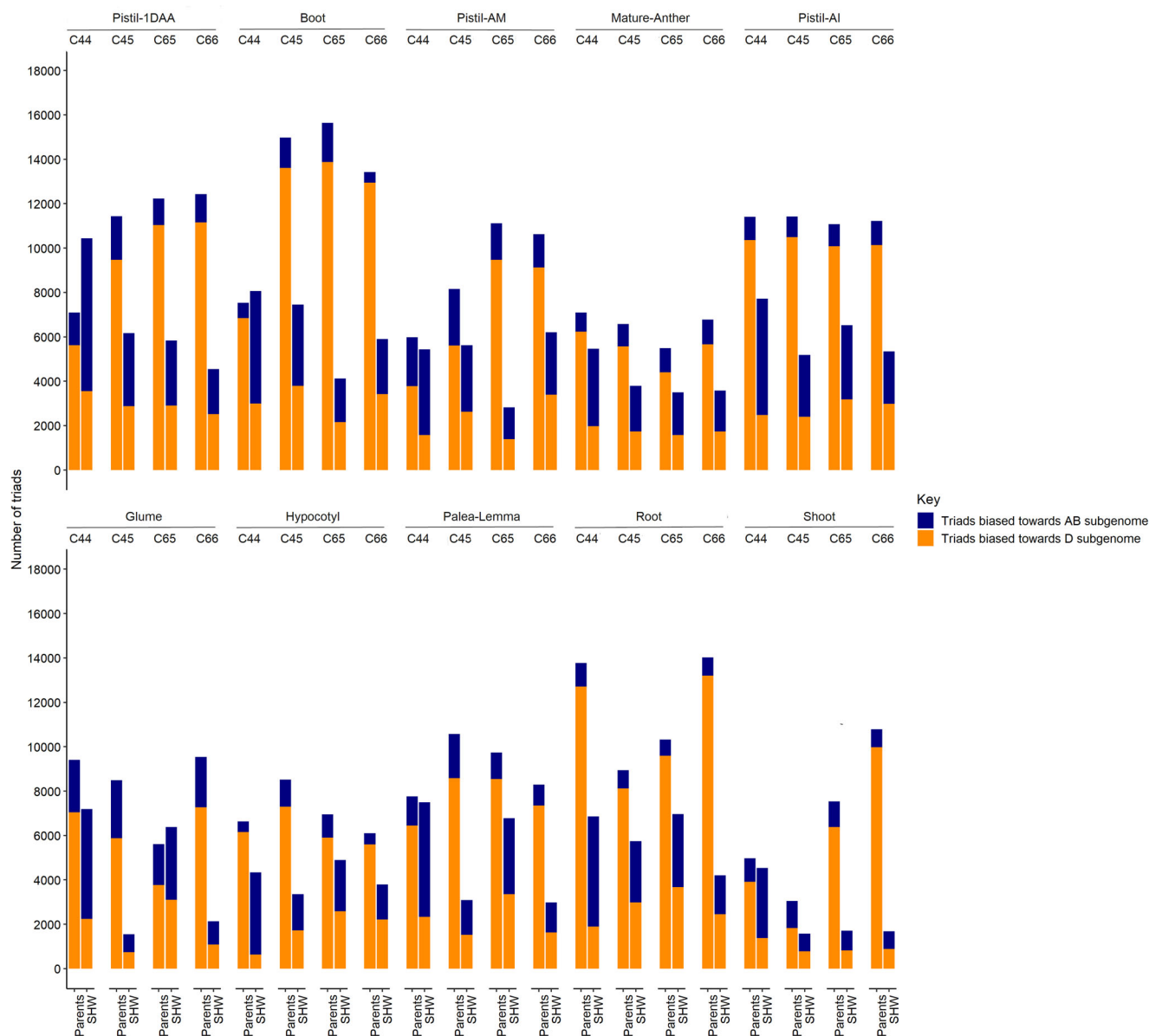
Fig. 1 Number of triads showing significant bias towards AB and D subgenomes in all four SHW lines—C44, C45, C65 and C66 and in the natural state of expression in their parents. The navy blue and dark orange bars represent the number of triads showing significant expression bias towards AB and D subgenomes, respectively. Pistil-1DAA pistil-one day after anthesis, Pistil-AM pistil—when anthers are at mature stage, Pistil-AI pistil—when anthers are at immature stage, Boot head at boot stage.

tissues (Fig. 1 and Supplementary Table 5). The pistil tissue when anthers were green and immature had more than 10,000 triads biased towards the D subgenome, based on parental expression levels in all four SHWs. The heads collected at the boot stage showed a similar pattern in all SHWs but C44 (Fig. 1).

In the shoot tissue of SHW-C66, 9978 and 808 triads showed a significant bias towards the D and AB subgenomes, respectively, based on the parental level of expression. However, in the SHW background, only 890 triads were biased towards the D subgenome (Fig. 2a, b). In anthers, 5667 and 1115 triads were biased towards the D and AB subgenomes, respectively, in the progenitor level of expression and more balanced in the hexaploid condition with 1740 D-biased and 1841 AB-biased triads (Fig. 2c, d). Likewise, 10,128 triads biased towards the D subgenome in the parental expression levels were reduced to 2976 in the pistils (collected when anthers were green and immature) of SHWs (Fig. 2e, f). These results illustrate major shifts in expression bias among hundreds of gene

sets, between the no-interaction scenario, i.e., the parental state of expression and the large-scale gene expression alteration caused by interactions of the subgenomes, i.e., in SHW lines. The number of triads significantly biased towards the D subgenome were higher in the parental genomic background, compared to their expression when they were introgressed into the synthetic hexaploid background. The distribution of the HEB for the other seven tissues for SHW-C66 is provided in Supplementary Figs. 2–8.

In the no-interaction predicted in-silico scenario, the proportion of triads biased towards the D genome were at least 1.5 fold (LogFC 0.6) higher than those biased towards the AB genome, as observed in shoot tissue of SHW-C45, and up to 27 fold (LogFC 4.8) higher, as observed in the head at boot stage tissue of SHW-C66 (Fig. 3). However, this number was drastically reduced in the actual hexaploid background and, an almost equal number of triads were biased towards the subgenomes of both parents (AB and D). On the contrary, the triads showing expression bias in the
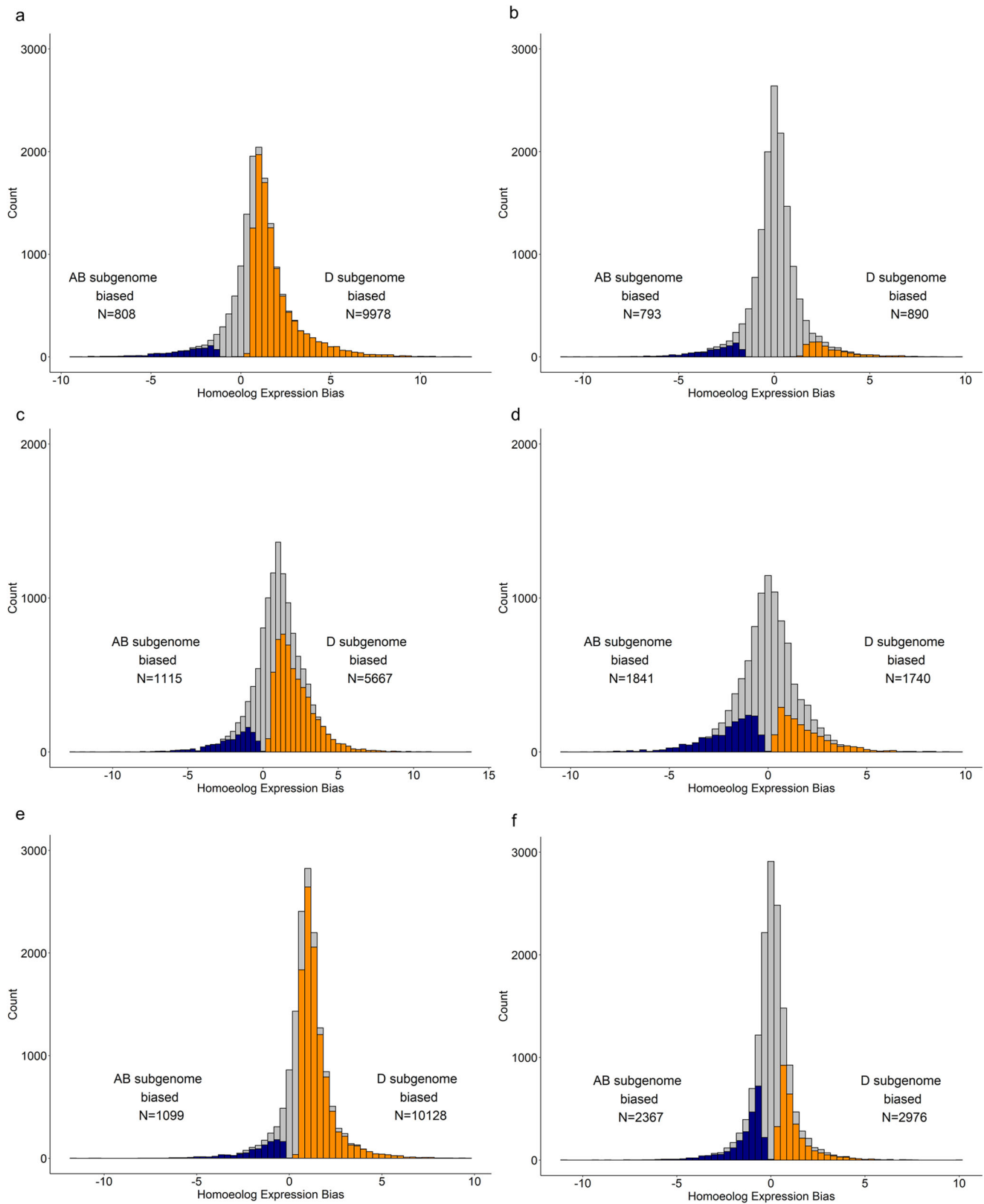
Fig. 2 Distribution of homoeologue expression bias in SHW-C66. (**a**, **c**, **e**) Comparison of the expression bias of triads towards AB and D genomes of tetraploid (PI377655) and diploid (AS2386) parents; (**b**, **d**, **f**) Comparison of the expression bias of triads towards AB and D subgenomes of SHW-C66; (**a**, **b**) Shoot; (**c**, **d**) Mature anther just prior to dehiscence; (**e**, **f**) Pistil when anthers are green and immature. The navy blue and dark orange bars represent the triads showing significant expression bias towards AB and D subgenomes, respectively.
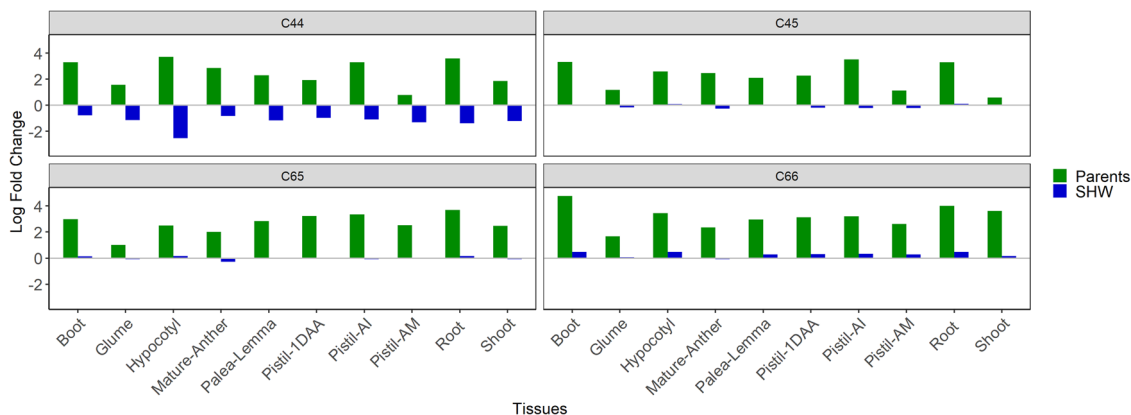
Fig. 3 Comparison of ratios of triads significantly biased towards AB and D subgenomes within a genomic background (parents and SHWs). The ratios are represented as log fold change (LogFC) values. A LogFC < 0 indicates more triads are biased towards the AB subgenome than the D subgenome and a LogFC > 0 indicates more triads are biased towards the D subgenome than the AB subgenome, within a genomic background. C44, C45, C65 and C66 are the four SHWs taken for the study. The green bars represent the ratios in the parental genomic background and the blue bars represent the SHW genomic background. Pistil-1DAA pistil-one day after anthesis, Pistil-AM pistil-when anthers are at mature stage, Pistil-AI pistil-when anthers are at immature stage, Boot head at boot stage.
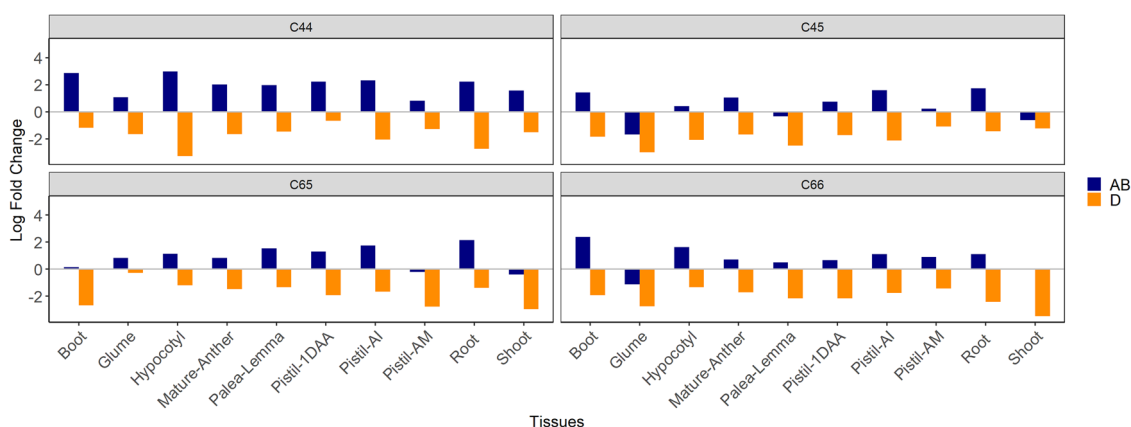
Fig. 4 Comparison of ratios of triads significantly biased towards AB and D subgenomes between genomic backgrounds (parents vs SHWs). The ratios are presented as log fold change (LogFC) values. A LogFC > 0 indicates more triads are biased towards the AB subgenome in the SHW background than the parental background and a LogFC < 0 indicates more triads are biased towards the AB subgenome in the parental background than the SHW, similarly for D subgenome. C44, C45, C65 and C66 are the four SHWs used for the study. The navy blue and dark orange bars represent the triads showing significant expression bias towards AB and D subgenomes, respectively. Pistil-1DAA pistil-one day after anthesis, Pistil-AM pistil-when anthers are at mature stage, Pistil-AI pistil-when anthers are at immature stage, Boot head at boot stage.

opposite direction, i.e., biased towards the AB genome increased in most tissue in the SHWs, with the most pronounced HEB being observed in SHW line C44 (Fig. 3).

The analysis of the trends within the same subgenome background revealed that the triads biased towards the AB subgenome were more numerous in the hexaploid state when compared to the parental state (Fig. 4). All tissues of SHW-C44 had an increased proportion of triads showing expression bias towards the AB subgenome, with up to 7.9-fold (LogFC 2.98) more triads overexpressed in the hypocotyl tissue. In the other three SHW lines, the increase in triads biased towards the AB subgenome were up to 5.2-fold (head at boot stage of C66; LogFC 2.37). The proportion of triads biased towards the D subgenome were reduced by more than eight times as detected across multiple SHW-tissue contexts.

The average magnitude of HEB values for the AB and D biased triads for the SHW and parental background scenarios are summarized in Supplementary Table 5. For the AB-biased triads, the magnitude of HEB values ranged from −1.17 (C65—pistil when anthers are green and immature) to −4.37 (C44—hypocotyl), in the SHW background. The magnitude of HEB of

the AB biased triads varied from −1.22 (C65—head at boot stage) to −3.56 (C66—head at boot stage), in the parental background. Similarly, the magnitude of HEB values of the D-biased triads extended from 1.16 (C44—pistil when anthers are green and immature) to 3.12 (C45—glume), in SHW background, and from 1.32 (C45—head at boot stage) to 4.66 (C45—shoot), in parental background. The average magnitude of HEB values of the AB-biased triads were higher than those of the D-biased triads across all SHW-tissue scenarios, except in the C65 SHW background in pistil when anthers are green and immature. However, no fixed pattern was observed in the parental background (Supplementary Table 5).

Role of parental genotypes on HEB. The subset of triads showing similar expression bias patterns across different SHW lines was also analysed. The intersections among the four SHW lines and between the SHW lines sharing common tetraploid or diploid parents is represented in Fig. 5 for the shoot tissue. The largest intersection represented the subset showing no significant bias in all four SHW
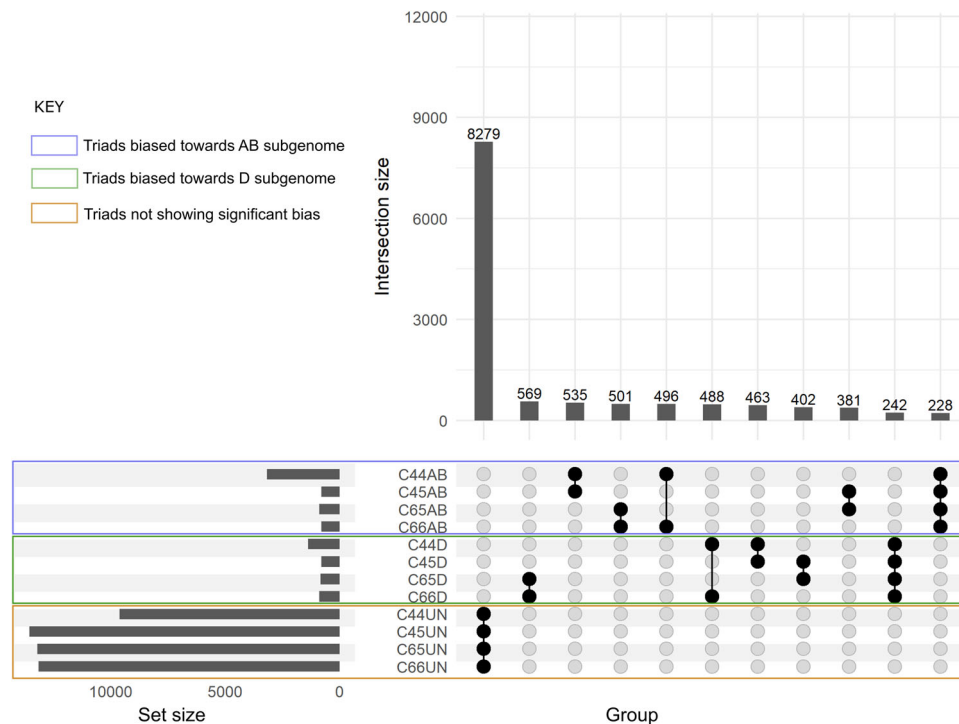
Fig. 5 Subsets of triads showing similar expression pattern across all four SHW lines in shoot tissue. The gray horizontal bars under set size corresponding to C44AB indicates the number of triads biased towards the AB subgenome in the SHW-C44; C44D indicates the number of triads biased towards the D subgenome in the SHW-C44; C44UN indicates the number of triads showing no significant bias in the SHW-C44; Similarly, for SHWs C45, C65 and C66. The vertical bars under intersection size indicate the number of triads showing similar expression patterns between sets. The black dots highlight the two SHW sets compared for the corresponding intersection. The number over the bars represents the number of triads showing similar expression pattern. The triad sets showing significant bias towards the AB or D subgenomes and those not showing any significant bias are highlighted by the blue, green and brown rectangles, respectively.

lines, and this was true across all ten tissues (Fig. 5; Supplementary Figs. 9–17). Shoot tissue had the smallest subset of triads showing similar significant expression bias towards the AB (228 triads) and D (242 triads) subgenomes, across the four SHW lines (Fig. 5). The largest intersection among the SHW lines was found in pistil when anthers are green and immature (Supplementary Fig. 15) and pistil–one day after anthesis, for the AB-biased (1281) and D-biased (1066) triads, respectively (Supplementary Fig. 9).

Role of tissue of expression of homoeologues on HEB. Out of the 18,357 triads analysed for the expression pattern, 69% to 73% were not comprised of any tissue-specific homoeologues at all, in all four SHW lines (Supplementary Table 6). The next largest fractions were triads with all three homoeologues showing same tissue specific expression (8% to 11%) and triads with only one homoeologue showing tissue specificity with the other two being more broadly expressed (7% to 12%). Further, there were more triads with two homoeologues showing expression specificity towards the same tissue and one broadly expressed homoeologue, than one homoeologue alone displaying expression specificity in another tissue (Supplementary Table 6).

Similar constitution was observed in individual SHW-tissue contexts, most commonly, with only one of the three homoeologues being tissue specific or all three homoeologues showing tissue specificity (Supplementary Table 7). The triads with two of the three homoeologues alone showing tissue specificity were lowest in number across the ten tissues in all four SHW lines. In the root and mature anther tissues, the subset of triads comprised of all homoeologues showing tissue specificity were largest in all four SHW lines. Out of the larger subsets of triads with all three

homoeologues showing same tissue specificity, 455 triads in the root were common among all four SHW lines. When the homoeologues of these triads were subjected to gene ontology analysis for understanding functional enrichment, biological process terms including regulation of root morphogenesis and regulation of root meristem growth were enriched in the root. In mature anthers, this number was 427 and the biological process terms including pollen germination, pollen sperm cell differentiation, pollen tube growth, and pollen wall assembly were enriched.

A large number of triads, comprised of either one, two or three tissue-specific homoeologues, showed no significant expression bias, similar to the observation in the entire triad dataset. When the triads showing AB- and D- bias and their tissue specificity patterns were interrogated specifically, a moderate strength of association with Cramer's-V statistic ranging between 0.20 and 0.57 was observed between bias pattern and tissue specificity of the homoeologues in most tissues in all four SHW lines. For instance, more triads showed D-subgenome bias when D-homoeologue alone was tissue-specific, greater number of AB-biased triads were observed when A and B homoelogues showed tissue specificity, and more D-biased triads were observed when A and D or B and D homoeologues alone showed tissue specificity, in most SHW-tissue contexts. However, based on Chi-square test of independence, the association between expression bias and tissue specificity was significant only in mature anthers, root and hypocotyl tissues in all four SHW lines, and C66 showed significant association additionally in pistil-one day after anthesis, pistil-when anthers are green and immature and C44 displayed significant association in all tissues. Although the observed extent of associations were moderate, the relationship between homoeologues' tissue specificity and expression bias cannot be dismissed.

Impact of genome stabilization on HEB. To validate the observed patterns in more stabilized genomes, the expression bias was analyzed in wheat lines that were subjected to domestication and breeding processes using the publicly available RNA-seq data for the landrace Chinese Spring and the cultivar Azhurnaya from Ramírez-Gonzalez et al.[36]. The number of HEB and the magnitude of the expression biases were much smaller in these genotypes compared to the SHW lines (Fig. 6). In Azhurnaya, 2087 to 4244 triads were biased towards the AB subgenome and 2032 to 3739 triads were biased towards the D subgenome, in tissues reported in Fig. 6. In Chinese Spring, 874 to 4512 triads were biased towards the AB subgenome and 812 to 3821 triads were biased towards the D subgenome, across the five tissues used in the analysis. The magnitude of biases was lower in the two domesticated wheat lines compared to the newly created SHW lines.

The influence of domestication and breeding processes on the HEB patterns of the triads were analysed using the SHW, Chinese Spring and Azhurnaya results. Three of our SHW tissues could be matched to the developmental stages of samples in Ramírez-Gonzalez et al.[36] namely pistils when anthers are mature, root and head at boot stage, and were used in this analysis. The majority of the triads (29–55%) maintained similar expression bias patterns in the resynthesized wheat, the landrace (Chinese Spring) and the cultivar (Azhurnaya) in all three tissues (Supplementary Fig. 18). Only 9–14% showed similar bias patterns in the SHW and Chinese Spring, and transitioned to a different bias state in Azhurnaya, possibly reflecting the impact of rigorous selection. Further, 9–30% of the triads, based on the different SHW-tissue scenarios, showed reversal to SHW bias state in Azhurnaya (Supplementary Fig. 18). Taking into consideration the HEB differences for the triads among the SHW lines, we also specifically investigated the subset of triads showing similar bias patterns across all four SHW lines. The major proportion of triads in this subset retained the same expression bias patterns in the SHWs, Chinese Spring and Azhurnaya, and the second largest proportion showed similar patterns in the SHWs and Azhurnaya and not in Chinese Spring, indicating the reversal of bias patterns during crop improvement. The triad set showing similar pattern changes including Chinese Spring and Azhurnaya was the smallest.

The HEB expression analysis across ten different tissues enabled the investigation of bias patterns of the triads across tissues, i.e., through the developmental stages. In all four SHW
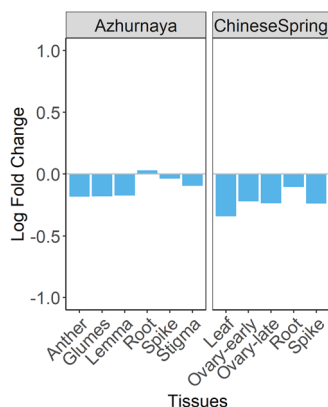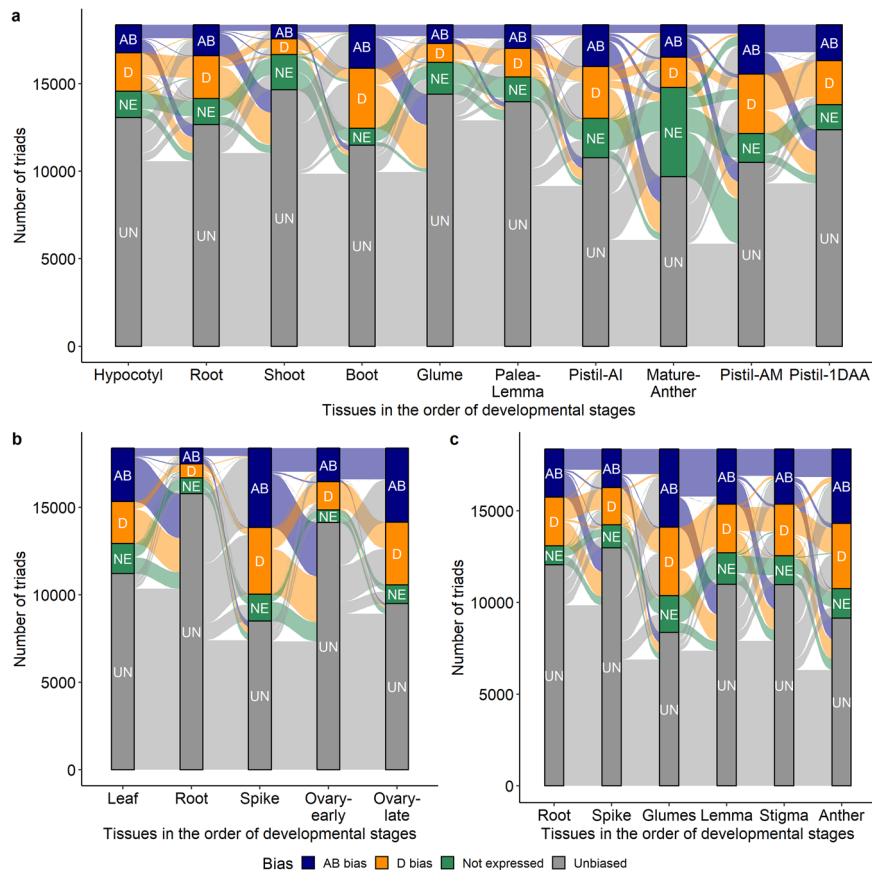
lines, the largest pattern observed was triads showing no significant bias across any tissues. There were anywhere from 1449 to 3014 triads in the SHW lines that showed this pattern (Fig. 7, Supplementary Fig. 19). In all four SHW lines, the next most commonly observed bias trend was where the triads did not show significant bias in any of the tissues but were biased towards the AB or D subgenomes only in the mature anther tissues (Fig. 7, Supplementary Fig. 19). Patterns showing only significant bias in one of the tissues such as pistil—one day after anthesis, pistil when anthers are green and immature and root, were observed in multiple triads. In addition, there were thousands of triad-specific patterns across tissues. Similar trends were also observed in Azhurnaya, and the subset of triads showing significant bias in anther tissues was larger, whereas in Chinese Spring, triads showing significant bias only in spike tissue were common.

Differential expression analysis of complete gene set. In addition, the differential expression analysis of the whole gene set between the parents and SHWs was investigated. The 106,913 high confidence gene models identified in the hexaploid wheat genome as per the IWGSC RefSeq annotation v2.1[33] were used in the analysis. In the genes from the A and B subgenomes, only a few hundred to a maximum of 1015 genes were downregulated in the SHWs compared to their tetraploid parents in SHW-C44 (Fig. 8). In contrast, more than 20,000 D subgenome genes were downregulated in all SHWs compared to their diploid parents (Fig. 8). Most importantly, in all three subgenomes, only 357 to 755 genes and 141 to 372 genes exhibited upregulation in the hexaploid state, in the AB and D subgenomes, respectively.

Qualitative differences in transcripts between parents and SHWs. The RNA-seq data was also mined to characterize the differences in splicing patterns at the different ploidy levels to shed light on the qualitative differences among transcripts. This analysis captured the splice variants from all genes whether differentially-expressed or not. Five major types of alternative splicing events, namely mutually exclusive exons (MXE), alternative 3′ splicing site (A3SS), alternative 5′ splicing site (A5SS), retained intron (RI) and skipped exon (SE) were investigated (Fig. 9). As an example, the number of statistically significant alternatively spliced RNA isoforms detected in each genotype is illustrated for shoot tissue in the facets of the stacked bar chart (Fig. 10a, b). *Retained intron* was the most common alternative splicing event (34.2–79.5%) and *mutually exclusive exons* were least detected (0–7.3%) in all scenarios. The mature anthers and pistil—when anthers are green and immature, showed the least alternative spliced events in comparison with other tissues. The number of alternatively spliced events detected in the other nine tissues are presented in Supplementary Figs. 20–28. Among the multiple alternatively spliced events detected, 69%-78% corresponded to homoeologous genes in triads (Fig. 10c and Supplementary Figs. 20–28). In all comparisons, more alternative splicing events were associated with D homoeologues, except in the shoot tissue of C45 vs parents, where an equal proportion of events was associated with all three subgenomes (Fig. 10c and Supplementary Figs. 20–28). Overall, large-scale quantitative repression of the D subgenome was observed in SHW lines as measured by homoeologous expression biases and, qualitative changes were also more prominent in this subgenome as illustrated by the larger representation of alternative RNA splicing events.

Discussion

The process of introgressing desirable novel alleles and haplotypes underlying phenotypes from genetically diverse parents

Fig. 6 Ratio of triads significantly biased towards AB and D subgenomes in cultivar Azhurnaya and landrace Chinese Spring. The ratios are represented as log fold change (LogFC) values depicted by the blue bars. A LogFC < 0 indicates more triads are biased towards the AB subgenome than the D subgenome and a LogFC > 0 indicates more triads are biased towards the D subgenome than the AB subgenome.

Fig. 7 Expression bias trends of the triads across tissues. Alluvial plots representing the expression bias trends of the triads across the tissues in the (**a**) SHW-C66 (**b**) Landrace Chinese Spring and (**c**) Cultivar Azhurnaya; AB bias (AB), represented by navy blue bars: triads significantly biased towards the AB subgenome; D bias (D), represented by dark orange bars: triads significantly biased towards the D subgenome; Not expressed (NE), represented by green bars: the homoeologues of the triads are not expressed; Unbiased (UN), represented by gray bars: no significant expression bias observed.

known as pre-breeding is a viable strategy to develop cultivars with improved disease resistance, quality and agronomic traits, including climate resilience traits. However, hybridizations among diverse parents often result in poor penetrance and expressivity of traits sourced from donor parents. SHWs are being generated to access the genetic diversity present in the *Ae. tauschii* genepool (D genome donor) to overcome this allelic diversity bottleneck of bread wheat, wherein poor recovery of parental traits were reported when their genomic background changes from diploid (DD) to hexaploid (AABBDD)[19]. Hence, a critical understanding of the genome-wide impact of polyploidization in crop species is needed to design strategies to enhance the potential of pre-breeding. In this study, we investigated and unraveled the global expression level alterations in SHW lines and their corresponding tetraploid and diploid progenitors to portray the transcriptome dynamics taking place during initial polyploidization events. The four parents used in the study were diverse in geographical origin. Specifically, *T. turgidum* cv. Langdon is from North America (Langdon) and PI377655 is from former Yugoslavia, while *Ae. tauschii* AS2386 and AS2399 are from Iran[37,38]. The primary SHW plants used in the experiment had been selfed for at least four generations; these selfings potentially facilitated the stabilization of the genome compared to nascent polyploids[39], allowing the capture of heritable transcriptomic variations.

The triad gene sets (genes present in 1:1:1 condition in ABD homeologue chromosomes) form the vital portion of the wheat genome playing key functional roles, while the genes present as other homoeologue copy number variations perform other

functions[40]. In all ten developmentally distinct tissues in the four SHW lines, the majority of the genes from the D subgenome were suppressed in SHWs, when compared with their parental levels of expression. Li et al.[41] in their transcriptome analysis of resynthesized wheat, have also shown the shift in expression dominance while comparing parental and hexaploid states. Changes in expression bias among homoeologues upon whole genome duplication have been observed in other allopolyploid crops upon resynthesis, including in rapeseed[42] and cotton[43], reflecting the subgenome dominance in these species as well. However, in bread wheat, it was proposed that a balance is prevalent among the subgenomes, without any subgenome dominance[34,40]. Nevertheless, tissue-, growth stage-, and environmental cue- dependent dominance have been reported as the cause of developmental process and environmental adaptations[44,45]. In this study, based on the observations in the SHW lines, the expression level dominance in the parental states is reduced by polyploidization, navigating the evolutionary process towards achieving an overall balanced allohexaploid, while the homoeologue-specific variability is capitalized by the developmental or environmental requirements. In domesticated wheat genotypes, the expression of D homoeologues were found to be less repressed than the A and B homoeologues[36,40]. Hence, it can be considered that the initial broad-scale repressive effect on D homoeologues is to establish a balance in expression between genomes derived from different parents (AB and D), and not for the subjugation of the D subgenome. Further, the expression bias patterns established in newly synthesized polyploids are maintained over multiple generations through the processes of domestication and breeding, as
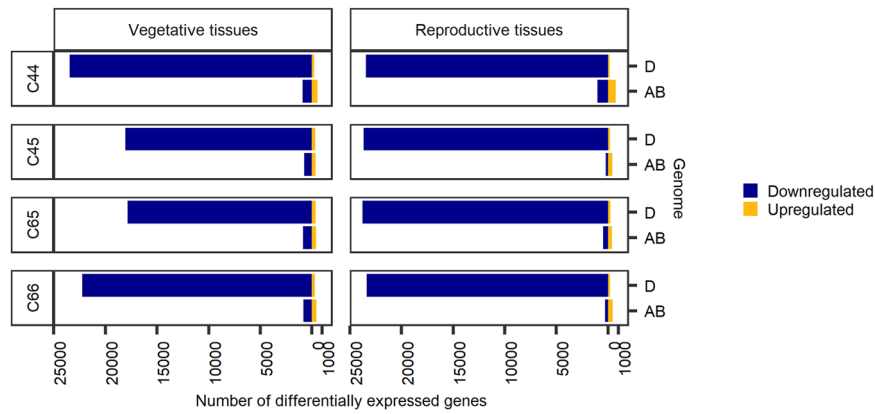
Fig. 8 Differentially expressed genes of the AB and D subgenomes of the vegetative and reproductive tissue pools upon comparison of the SHW lines (C44, C45, C65 and C66) with their parents. The blue bars represent the number of downregulated genes and the yellow bars represent the number of upregulated genes.
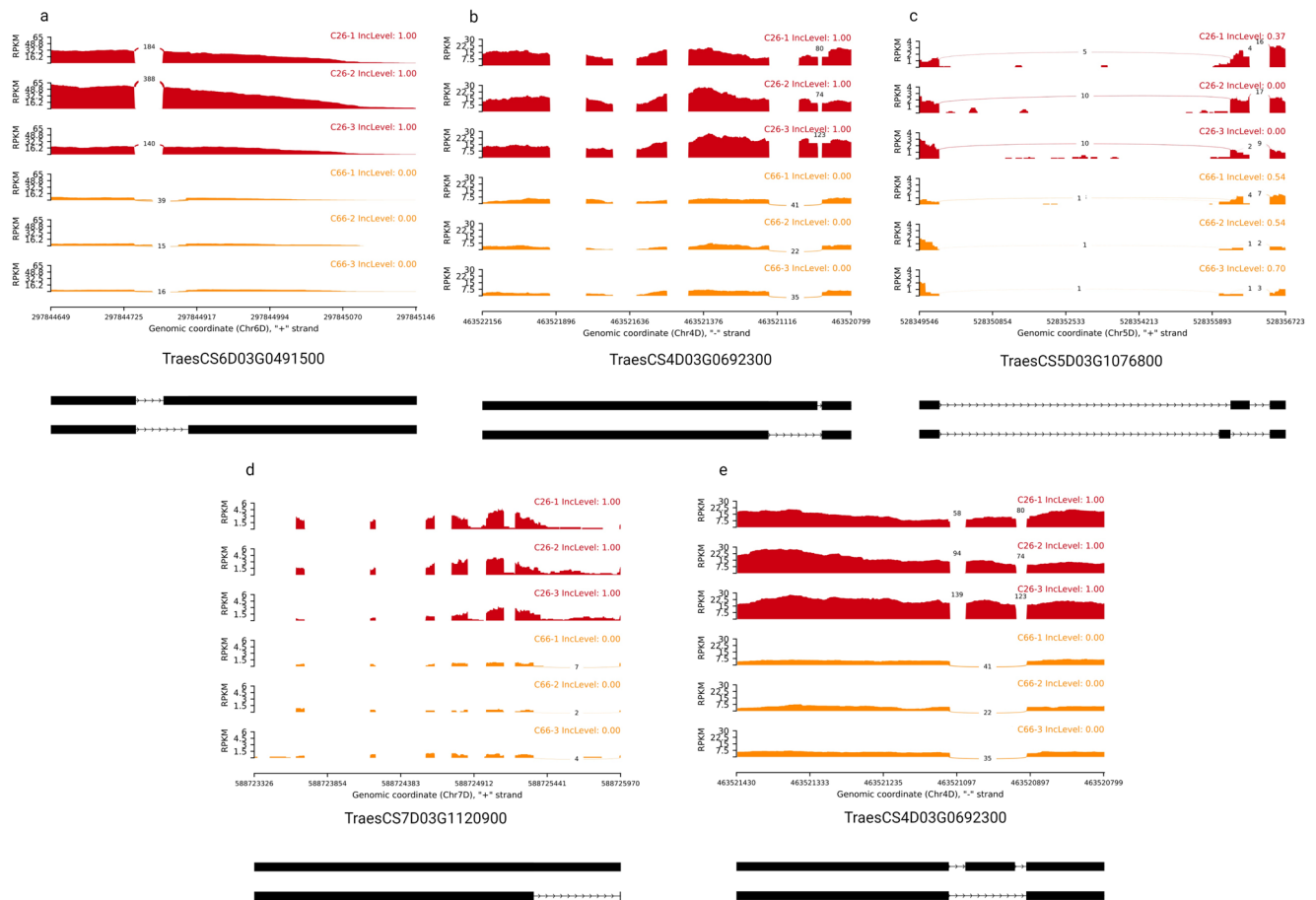

Fig. 9 Sashimi plots representing examples of the five types of alternative splicing events in the shoot tissue of *Ae. tauschii* C26 line (diploid parent) vs SHW line C66. (**a**) Alternative 3′ splicing site; (**b**) Alternative 5′ splicing site; (**c**) Mutually exclusive exons; (**d**) Retained intron; (**e**) Skipped exon. The red plots represent the *Ae. tauschii* C26 parent transcripts and the yellow plots represent the SHW line C66 transcripts.

observed in the RNA-seq data from more stabilized genomes such as Chinese Spring (a landrace) and Azhurnaya (a cultivar), sourced from the open-source databases corresponding to Ramírez-Gonzalez et al.[36]. Hence, the initial novel expression patterns (repression of D subgenome) observed in newly synthesized hexaploid wheat, are stabilized and inherited over multiple generations.

Polyploidization events involve genome reshuffling and can alter the gene dosage either immediately after duplication, or over

a long period of time (million year scale) during evolution[46]. Here we showed the instantaneous effects (within five generations of selfing since initial crossing) of polyploidization in hexaploid wheat. The shift in expression balance can be attributed to gene dosage (allele copies) because biologically-significant stoichiometry of macromolecule complexes and regulatory genes impacting downstream loci are critical for the fitness of an organism[47]. The difference among expression bias of the same gene sets between tissues may be attributed to the expression
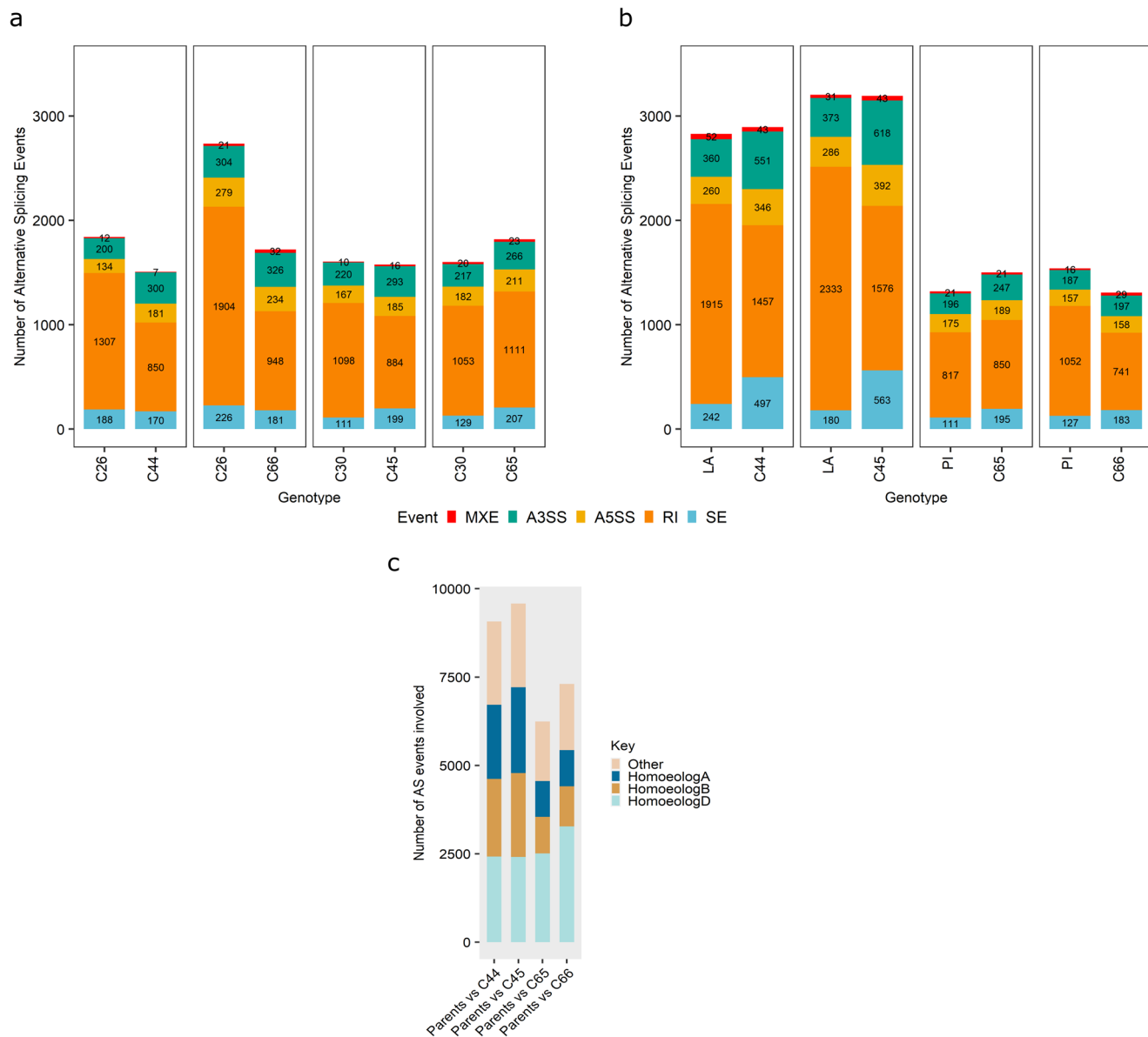
Fig. 10 Number of alternative splicing events detected in the diploid (C26, C30)/tetraploid (LA, PI) parents compared to the SHWs (C44, C45, C65, C66) in shoot tissue. (**a**) Diploid parent vs. SHW line; (**b**) Tetraploid parent vs. SHW. MXE mutually exclusive exons, A3SS alternative 3′ splicing, A5SS alternative 5′ splicing, RI retained intron, SE skipped exon, LA Langdon, PI PI377655; (**c**) Proportion of alternative splicing events involving the homoeologues in triads and other genes.

contribution of the duplicates that might differ in a tissue-specific pattern, especially in the reproductive cells[48,49]. Notably, five of the ten tissues used in the study possibly had a higher fraction of haploid gametophytic cells, referred to as the reproductive pool. In yet another dimension, polyploidization events increase the complexity of the gene regulatory network, through the introduction of inter-genomic trans-regulatory mechanisms[50]. This could result in homoeoallele-specific expression as observed in previous studies in wheat[45] ($2n = 6x = 42$), high-ploidy sugarcane[51] ($2n = 8x–12x = 80–120$) and peanut[52] ($2n = 4x = 40$). Further, the individual developmental state of the plant and the genotype may also have an influence over allele specificity[51,53], further explaining the homoeoallele expression variation across the ten developmentally-distinct tissues and four genetically different SHW lines used in our investigation.

Gene expression dynamics is the translational component linking observed phenotypic variation and its underlying

genotype. Several triads showing similar expression bias patterns in the SHWs, Chinese Spring and Azhurnaya, indicate this subset remains unaffected by artificial selection. The other triads having altering HEB patterns in the resynthesized, domesticated and elite wheat backgrounds, imply the selection for specific expression bias patterns during domestication and crop improvement, which are potentially coupled to other genes regulating the characteristics of modern-day bread wheat. The limitation of this comparison is that the SHW lines used in the study are not the direct progenitors of neither Chinese Spring or Azhurnaya and the expression data are from different studies. Hence, further interrogation with multiple wheat landraces and elite wheat cultivars will shed more light on the selection for HEB.

While investigating the bias trends of triads across the tissues, the number of triads showing unbiased expression pattern across all tissues were the largest, aligning with the observation that major subsets of triads did not show significant bias in all four

SHW lines in all ten tissues. The HEB specific to biotic[45] and abiotic[54] stress response have been reported in wheat and allo-polyploid cotton. The results from this study, showing pre-ponderance of triad-specific bias trends across developmental stages, suggest the spatial and temporal variability of HEB, leading to different bias trends through developmental stages, in addition to stress responsive patterns. In allopolyploid cotton, a study analysing the fiber development, revealed the changes in chromatin architecture in the different developmental stages and their influence on gene regulatory networks (GRNs) leading to HEB[55]. This highlights the potential role of dynamic variation in 3D chromatin architecture in bringing about spatiotemporal differences in HEB. In addition, this also conveys how the expression plasticity in allopolyploids are leveraged throughout development.

Beyond the effect of the environmental stressors on expression bias, variability in expression bias in the different tissues has also been reported in several alloplyploid species, such as rapeseed[56], cotton[57], coffee[58] and white clover[59]. Previous work in hexaploid wheat has also shown that triads with varied expression bias patterns across tissues are more tissue specific[36]. In this study using SHW transcriptomes, the moderate association between homoeologue expression bias and tissue-specificity of homoeologues in the triads, elucidate the potential role of expression specificity of homoeologues in driving the direction of the expression biases. Hence, the tissue, developmental stage, environmental cues and parental genotype, all determine the bias patterns of the triads through exploiting the additional plasticity in the hexaploid wheat genomes when com-pared to their tetraploid and diploid progenitors.

In addition to comparison of expression patterns among homo-eoalleles, the differential expression analysis at the whole genome level between the parents and SHW lines performed to characterize the expression changes of the genes between different genomic backgrounds, revealed the large-scale repression of the D subgenome genes. The introgression of chromosomal segments from species in the secondary and tertiary gene pools has resulted in transcriptomic changes in both wheat[60–63] and other polyploid species[64,65]. Speci-fically, transcriptomic analysis of wheat × *Ambylopyrum muticum* introgression lines indicated the suppressed expression of genes in the introgressed segments[60]. Similarly, a higher proportion of genes in the introgressed region were down regulated in the expression studies of wheat-barley addition lines[61]. Introgression of alien chromatin segments into a species leads to the suppression of alien transcripts by the host genome; though how the host genome dis-tinguishes the alien genome remains unclear. Combining the D genome from *Ae. tauschii* with the AB genome in *T. turgidum* results in analogous responses, hinting that the AB genome acts as the native genome and the D genome as the introduced foreign chromatin in SHWs.

The outcome of a polyploidization event at the genomic, and consequently at the phenotypic level is dependant on the genetic background of the lines involved in the hybridization event[66,67]. The assessment of multiple SHW lines derived from diverse tet-raploid and diploid accessions showed variability in response to FHB infection, which could not be envisaged based on the FHB responses of the parent[68]. Similar variability based on genotypic background effects has been observed for yield-related traits of wheat lines with rye introgressions[69].

The expression level differences from polyploidization to achieve dosage-balance, to establish homoeollele-specific expression or in exhibiting genomic background effects, are all potentially the man-ifestation of the genetic and epigenetic changes that occur upon polyploidization. The chromosome-level analysis of synthetic lines and their tetraploid and diploid parents using molecular cytogenetic techniques have depicted structural differences in chromatin regions[70], and genome sequence level analysis have also revealed

similar outcomes[71,72]. In an epigenetic perspective, microRNAs (miRNAs), the key players in post transcriptional gene regulation, were expressed in a non-additive manner in resynthesized wheat[41]. The modified expression of miRNAs possibly underlies the down-stream shift in expression bias and repression in the mRNA data. Small interfering RNAs (siRNAs) derived from double stranded RNA molecules were involved in the methylation of DNA sequen-ces, specifically transposable elements, and also in establishing repressive heterochromatin marks[73]. Comparison of a newly syn-thesized hexaploid wheat and its *Ae. tauschii* parent revealed increased accumulation of siRNAs on the D homoeologues in the hexaploid, in contrast to the observations in *T. turgidum* vs hex-aploid A subgenomes[41]. The increased accumulation of siRNAs could be the underlying cause for the repression of D homeologues in hexaploids[41]. DNA methylation is an epigenetic feature for gene expression regulation and genome stability through silencing of transposable elements (TEs). The ploidy-level changes drive mod-ification of methylation patterns in CG, CHG and CHH sequence contexts in wheat, involving the TEs as well. As TEs occupy the major proportion of the wheat genome, such epigenetic repro-gramming of the TE fraction modulates expression of genes in its vicinity because TEs harbor promoter and enhancer motifs[74,75]. Among the multiple histone marks in the genome, an increase of the repressive histone-3 lysine-27 dimethylations (H3K27me2) was found to be positively correlated with the increase in ploidy in wheat[76]. Furthermore, the H3K27me3 marks underlie the subgenome-specific activity of regulatory elements[77], and the introduction of the D subgenome triggers distinct histone mod-ifications and consequent inter-subgenome regulatory interactions[29]. The chromatin accessibility, which is controlled by association between transcription factors and nucleosomes, is lower in the D subgenome of hexaploids compared to the *Ae. tauschii* background, also leading to extensive reduction in gene expression[78]. In addition to these genetic and epigenetic changes, alteration of chromatin architecture upon polyploidization also warrants investigation in hexaploid wheat. For instance, changes in topologically-associated domains (TAD) and A/B compartmentalization have been reported in other polyploid crop species like watermelon[79] and cotton[80].

Following gene transcription, precursor mRNAs are modified using the splicing machinery to produce mature RNAs. During this process, the splicing machinery can retain different combinations of introns and exons by alternative splicing (AS). This AS mechanism can diversify the transcriptional and translational products of a gene, thereby altering its function in the developmental frameworks of space (different tissues) and time (vegetative vs reproductive stages)[81]. There are tissue-specific isoforms, and abiotic[82] and biotic[83] stress-induced AS events. However, the impact of polyploidization and the resulting genomic shock on AS have only recently been investigated. In a study of hexaploid wheat and its tetraploid and diploid parents and relatives, Yu et al.[84] also found that retained intron was the most common AS mechanism observed. The dom-inance of intron retention has been observed across plant species, including Arabidopsis[85], rice, sorghum and maize[86] and cotton[87]. The influence of polyploidization on differential splicing events has been observed in allopolyploid rapeseed, along with homoeolog-specific differences in AS[88]. The characterization of AS in the tran-scriptome derived from embryogenesis tissues in wheat and its progenitors, had a preponderance of alternative 3' splicing events, however, the difference has been attributed to AS-event specificity of the analysis tool and percent spliced in thresholds utilized[89].

In conclusion, the global transcriptome dynamics analysis in SHWs compared to its tetraploid and diploid parents revealed expression bias among homoeologues resulting in large-scale suppression of the D subgenome in the SHWs as inferred from more than 18,000 triads and expression from ten tissues. Quali-tative changes observed among transcripts in the form of novel

splice variants upon polyploidization also majorly impacted the D genome homoeologues. These quantitative (transcript abundance) and qualitative (splice variants) expression changes seem to be dependent on the genomic composition of the parents and the developmental stage (tissues) of the wheat lines. To harness the genetic diversity available in the D genome progenitors (*Ae. tauschii*) and other wild relatives for wheat improvement through interspecific hybridization, the potentiality of modified expression in the progeny must be taken into consideration to recover the desired phenotype in the polyploid background.

Methods

Plant materials. Four SHW lines (C44, C45, C65 and C66) and their two tetraploid (PI377655–C16, Langdon–C19) and two diploid (AS2386–C26, AS2399–C30) parents, for a total of eight genotypes, were used in the study (Supplementary Table 1). The following ten tissue samples (Supplementary Fig. 1) were collected from the above mentioned eight lines: (1) head at boot stage, (2) shoot collected at heading, (3) lemma and palea at heading, (4) glume from first and second floret of a spike, (5) pistil—when anthers are green and immature, (6) pistil—when anthers are yellow and just prior to dehiscence, (7) pistil—one day after anthesis, (8) yellow anther just prior to dehiscence, (9) hypocotyl and (10) root, where the latter two tissues were collected from young seedlings grown on Whatman filter papers. Three biological replicates per tissue from all the eight genotypes were used in the study (ten tissues × eight genotypes × three replicates = 240 samples).

RNA extraction and library preparation. The RNeasy Plant Mini Kit (Qiagen Inc., Germantown, MD, USA) was used for all tissues except for the mature anther and pistil when anthers are green and immature, which were isolated using miRNeasy mini kit (Qiagen Inc). The extracted RNAs were quantified initially with the Implen Nanophotometer (Implen Inc, Westlake Village, CA, USA) and the quality and concentration were assessed with the Agilent RNA 6000 Nano assay (Agilent Technologies, Santa Clara, CA, USA). The construction of cDNA libraries and sequencing were carried out at the Centre d'expertise et de services Génome Québec (Montreal, QC, Canada). Briefly, mRNA enrichment was performed using NEBNext Poly(A) Magnetic Isolation Module (New England BioLabs, Ipswich, MA, USA), followed by cDNA synthesis and library preparation with the NEBNext RNA First Strand Synthesis, NEBNext Ultra Directional RNA Second Strand Synthesis modules and the NEBNext Ultra II DNA Library Prep Kit (New England BioLabs). Libraries were quantified using PicoGreen dsDNA assay kit, and quality was analysed using LabChip GX (PerkinElmer, MA, USA). The normalized libraries were clustered on an Illumina cBot, and 100-bp paired-end reads were generated on a HiSeq 4000 platform (Illumina Inc., CA, USA) for the mature anthers and the pistils collected when the anthers were still immature, i.e., prior to pollination. The other eight tissues were sequenced as 150-bp paired-end reads on the NovaSeq 6000 platform (Illumina Inc.).

Processing and alignment of sequencing reads. The sequencing reads were pre-processed using Trimmomatic[90] and aligned to the recently updated version of the bread wheat genome pseudomolecules (RefSeq v2.1; Annotation v2.1)[33] using Spliced Transcripts Alignment to a Reference (STAR)[91], after genome indexing, with the maximum intron length set to 10,000 bp, the maximum number of allowed mismatches as six, and the other parameters left at their default settings.

Estimation of homoeologue expression bias (HEB) and likelihood ratio test (LRT). In the bread wheat genome, nearly 35% of the genes are present as triads, with one homoeologous copy in each of the subgenomes (1:1:1 in A:B:D)[34]. A total of 18,474 triads were identified using the Refseq v1.0 annotation. However, after taking into account the revisions made in Refseq v2.1 annotation, a set of 18,357 triads, comprised of gene models annotated with high confidence[33,34], were used for this study of comparative transcriptomes. An expression matrix was generated from the gene count data, and the reads per kilobase per million mapped reads (RPKM) values were estimated from the raw count data using the *rpkm* function in edgeR[92]. Homoeologue expression bias (HEB), representing the number of fold bias and its direction, was estimated using the following formula[35]:

$$HEB = \log_2\left(\frac{RPKM_D}{RPKM_{AB}}\right)$$

Where, $RPKM_{AB}$ and $RPKM_D$ represent the expression values from the AB and D subgenomes, respectively. For the RPKM expression values from the AB subgenome, the average between the expression levels of the A and B homoeologues of a triad gene set was utilized, considering the balanced expression levels of the A, B and D homoeologues, and the difference in gene length is accounted for by the normalization.

Further, the likelihood ratio test developed on MATLAB by Smith et al.[35] and first used in Edger et al.[93], was applied to identify the triads showing significant bias

towards the tetraploid (AABB) or diploid (DD) genomes. Likelihood ratio tests were performed using the expression data from the SHW lines as well as the data from their corresponding tetraploid and diploid parents, creating an *in-silico* SHW without inter-subgenome interactions, in order to determine the shift in expression bias upon polyploidization. The upstream input preparation and downstream summarization steps were performed with custom scripts in bash and R. In order to investigate the bias patterns of tissue-specific and broadly expressed genes, the Tau method[94] was used for all the genotypes with a cut-off of 0.8. The strength of association between the tissue specificity of the homoeologues and the expression bias of the triads was estimated using Cramer's-V statistic and the significance of association was determined using the Chi-square test of independence. The gene ontology analysis of the homoeologues in triads of root and mature anther tissues with A, B and D homoeologues showing same tissue specificity across all four SHW lines, was performed using the Triticeae-Gene Tribe tool kit[95]. The R scripts used for input file preparation for LRT and tissue-specificity analyses are available at https://github.com/akshaya-v/SHW-Expression-bias.

To unravel the HEB in domesticated wheat lines, the RNA-seq data from Ramírez-Gonzalez et al.[36] for the landrace Chinese Spring and the cultivar Azhurnaya were downloaded for the following tissues: Chines Spring: leaf (at 14 days), root (at 14 days), ovary (early anthesis), ovary (late anthesis) and spike (at booting); Azhurnaya: stigma and ovary (anthesis), glumes (milk grain stage), anther (anthesis), root (seedling), spike (30% spike out). The downstream read processing, alignment and HEB analysis was carried out as described above for the SHWs.

Differential expression analysis. The read count data from the ten tissues were grouped into two pools: vegetative (shoot, root, hypocotyl, glume and palea +lemma) and reproductive (head, pistil—when anthers are green, pistil—when anthers are yellow, pistil—one day after anthesis and yellow anther just prior to dehiscence). The logic was to group the tissues comprised of only vegetative cells into one pool and the tissues comprised of both vegetative and reproductive cells into another pool. The differential expression analysis in the two pools was performed using the edgeR statistical package[92], and a false discovery rate (FDR) of <0.05 and log fold change (LogFC) cut-off of ±2.00 were applied as thresholds for inferring differentially expressed genes.

Alternative splicing analysis. The alternative splicing analysis was carried out using rMATS[96]. The BAM files generated using STAR were utilized. The allow-clipping option was enabled to accept alignments with soft clipping. The detection of novel splice sites not present in the annotation files were allowed using the novelSS option. The alternative splicing analysis was performed for all eight possible SHW vs diploid/tetraploid parent comparisons in the ten tissues. The sashimi plots were made using rmats2sashimiplot[97]. The number and types of differentially spliced isoforms detected in parents or SHW lines were summarized for comparison. The alternative splice types were alternative 3' splice site, alternative 5' splice site, mutually exclusive exons, retained intron and skipped exon.

Statistics and reproducibility. Three biological replicates per tissue per genotype were utilized. There were ten tissues and eight genotypes, and hence, there were 240 tissue samples in total. The likelihood ratio test to detect significant HEB was performed with the null hypothesis that the homoeologues show equal levels of expression in AB and D subgenomes, and the alternate hypothesis that the homoeologues show unequal expression levels between subgenomes, as described in Smith et al.[35]. The differential expression analysis was performed using edgeR[92] v3.38.4 with FDR < 0.05 and logFC cut-off ±2.00. The alternative splicing analysis was carried out with rMATS[96] using the default splicing difference cut-off of 0.01%.

Reporting summary. Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw RNA-seq data has been deposited in the Short Read Archive (SRA) at NCBI under Bioproject PRJNA905376. The base data for the HEB figures, alternative splicing figures and tissue-specificity analyses results are available in Figshare (https://doi.org/10.6084/m9.figshare.c.6443621) (ref. [98]).

Code availability

The source code of the R scripts used for preparing input files and all other analyses are provided within the Methods section.

References

1. De Storme, N. & Mason, A. Plant speciation through chromosome instability and ploidy change: cellular mechanisms, molecular factors and evolutionary relevance. *Curr. Plant Biol.* **1**, 10–33 (2014).
2. Cui, L. et al. Widespread genome duplications throughout the history of flowering plants. *Genome Res.* **16**, 738–749 (2006).
3. Ren, R. et al. Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Mol. Plant* **11**, 414–428 (2018).
4. Salman-Minkov, A., Sabath, N. & Mayrose, I. Whole-genome duplication as a key factor in crop domestication. *Nat. Plants* **2**, 16115 (2016).
5. Jiao, Y., Li, J., Tang, H. & Paterson, A. H. Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. *Plant Cell* **26**, 2792–2802 (2014).
6. Alix, K., Gérard, P. R., Schwarzacher, T. & Heslop-Harrison, J. S. Polyploidy and interspecific hybridization: partners for adaptation, speciation and evolution in plants. *Ann. Bot.* **120**, 183–194 (2017).
7. Fawcett, J. A., Maere, S. & Van de Peer, Y. Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. *Proc. Natl Acad. Sci. USA* **106**, 5737 (2009).
8. Doyle, J. J. et al. Evolutionary genetics of genome merger and doubling in plants. *Annu. Rev. Genet.* **42**, 443–461 (2008).
9. Van de Peer, Y., Ashman, T.-L., Soltis, P. S. & Soltis, D. E. Polyploidy: an evolutionary and ecological force in stressful times. *Plant Cell* **33**, 11–26 (2021).
10. Huang, S. et al. 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 (2002).
11. Feuillet, C., Langridge, P. & Waugh, R. Cereal breeding takes a walk on the wild side. *Trends Genet.* **24**, 24–32 (2008).
12. Cox, T. S. Deepening the wheat gene pool. *J. Crop Prod.* **1**, 1–25 (1997).
13. He, F. et al. Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nat. Genet.* **51**, 896–904 (2019).
14. Akhunov, E. D. et al. Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC Genomics* **11**, 702 (2010).
15. Dempewolf, H. et al. Past and future use of wild relatives in crop breeding. *Crop Sci.* **57**, 1070–1082 (2017).
16. Ogbonnaya, F. C. et al. Synthetic hexaploids: harnessing species of the primary gene pool for wheat improvement. *Plant Breed. Rev.* **37**, 35–122 (2013).
17. Mujeeb-Kazi, A., Gul, A., Farooq, M., Rizwan, S. & Ahmad, I. Rebirth of synthetic hexaploids with global implications for wheat improvement. *Aust. J. Agric. Res.* **59**, 391–398 (2008).
18. Mirzaghaderi, G. & Mason, A. S. Broadening the bread wheat D genome. *Theor. Appl Genet.* **132**, 1295–1307 (2019).
19. Dreisigacker, S. & Kishii, M. The use of synthetic hexaploid wheat in the CIMMYT Global Wheat Program. *CIMMYT* https://repository.cimmyt.org/handle/10883/20692 (2019).
20. Hiebert, C. W. et al. Stem rust resistance in wheat is suppressed by a subunit of the mediator complex. *Nat. Commun.* **11**, 1123 (2020).
21. Bai, D. & Knott, D. Suppression of rust resistance in bread wheat (*Triticum aestivum* L.) by D-genome chromosomes. *Genome* **35**, 276–282 (1992).
22. Yang, C. et al. Evolution of physiological responses to salt stress in hexaploid wheat. *Proc. Natl Acad. Sci. USA* **111**, 11882–11887 (2014).
23. Hong, Y. et al. Transcript suppression of TaGW2 increased grain width and weight in bread wheat. *Func. Integr. Genomics* **14**, 341–349 (2014).
24. Dubcovsky, J. & Dvorak, J. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* **316**, 1862–1866 (2007).
25. Nieto Feliner, G., Casacuberta, J. & Wendel, J. F. Genomics of evolutionary novelty in hybrids and polyploids. *Front. Genet.* **11**, 792 (2020).
26. Levy, A. A. & Feldman, M. Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization. *Biol. J. Linn. Soc.* **82**, 607–613 (2004).
27. Wicker, T. et al. Impact of transposable elements on genome structure and evolution in bread wheat. *Genome Biol.* **19**, 103 (2018).
28. Guo, X. & Han, F. Asymmetric epigenetic modification and elimination of rDNA sequences by polyploidization in wheat. *Plant Cell* **26**, 4311–4327 (2014).
29. Lv, Z. et al. Conservation and trans-regulation of histone modification in the A and B subgenomes of polyploid wheat during domestication and ploidy transition. *BMC Biol.* **19**, 42 (2021).
30. Feldman, M. & Levy, A. A. Genome evolution due to allopolyploidization in wheat. *Genetics* **192**, 763–774 (2012).
31. Jiao, W. et al. Asymmetrical changes of gene expression, small RNAs and chromatin in two resynthesized wheat allotetraploids. *Plant J.* **93**, 828–842 (2018).
32. Li, Z. et al. The bread wheat epigenomic map reveals distinct chromatin architectural and evolutionary features of functional genetic elements. *Genome Biol.* **20**, 139 (2019).
33. Zhu, T. et al. Optical maps refine the bread wheat Triticum aestivum cv. Chinese Spring genome assembly. *Plant J.* **107**, 303–314 (2021).
34. IWGSC. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**, eaar7191 (2018).
35. Smith, R. D., Kinser, T. J., Smith, G. D. C. & Puzey, J. R. A likelihood ratio test for changes in homeolog expression bias. *BMC Bioinform.* **20**, 1–11 (2019).
36. Ramírez-González, R. H. et al. The transcriptional landscape of polyploid wheat. *Science* **361**, eaar6089 (2018).
37. Zhang, L.-Q. et al. Frequent occurrence of unreduced gametes in *Triticum turgidum–Aegilops tauschii* hybrids. *Euphytica* **172**, 285–294 (2010).
38. Liu, M. et al. Stripe rust resistance in *Aegilops tauschii* germplasm. *Crop Sci.* **53**, 2014–2020 (2013).
39. Shaked, H., Kashkush, K., Ozkan, H., Feldman, M. & Levy, A. A. 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 (2001).
40. Juery, C. et al. New insights into homoeologous copy number variations in the hexaploid wheat genome. *Plant Genome* **14**, e20069 (2021).
41. Li, A. et al. mRNA and Small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell* **26**, 1878–1900 (2014).
42. Wu, J. et al. Homoeolog expression bias and expression level dominance in resynthesized allopolyploid *Brassica napus*. *BMC Genomics* **19**, 586 (2018).
43. Yoo, M., Szadkowski, E. & Wendel, J. Homoeolog expression bias and expression level dominance in allopolyploid cotton. *Heredity* **110**, 171–180 (2013).
44. Pfeifer, M. et al. Genome interplay in the grain transcriptome of hexaploid bread wheat. *Science* **345**, 1250091 (2014).
45. Powell, J. J. et al. The defence-associated transcriptome of hexaploid wheat displays homoeolog expression and induction bias. *Plant Biotechnol. J.* **15**, 533–543 (2017).
46. Veitia, R. A. & Birchler, J. A. Gene-dosage issues: a recurrent theme in whole genome duplication events. *Trends Genet.* **38**, 1–3 (2022).
47. Birchler, J. A. Insights from paleogenomic and population studies into the consequences of dosage sensitive gene expression in plants. *Curr. Opin. Plant Biol.* **15**, 544–548 (2012).
48. Conant, G. C., Birchler, J. A. & Pires, J. C. Dosage, duplication, and diploidization: clarifying the interplay of multiple models for duplicate gene evolution over time. *Curr. Opin. Plant Biol.* **19**, 91–98 (2014).
49. Birchler, J. A. & Veitia, R. One hundred years of gene balance: How stoichiometric issues affect gene expression, genome evolution, and quantitative traits. *Cytogenet Genome Res.* **161**, 529–550 (2021).
50. Hu, G. & Wendel, J. F. Cis–trans controls and regulatory novelty accompanying allopolyploidization. *N. Phytol.* **221**, 1691–1700 (2019).
51. Margarido, G.R.A., Correr, F.H., Furtado, A., Botha, F.C. & Henry, R.J. Limited allele-specific gene expression in highly polyploid sugarcane. *Genome Res.* https://doi.org/10.1101/gr.275904.121 (2021).
52. Peng, Z. et al. Natural polymorphisms in a pair of NSP2 homoeologs can cause loss of nodulation in peanut. *J. Exp. Bot.* **72**, 1104–1118 (2020).
53. Correr, F.H., Furtado, A., Garcia, A.A.F., Henry, R.J. & Margarido, G.R.A. Allele expression biases in mixed-ploid sugarcane accessions. *bioRxiv* https://doi.org/10.1101/2021.08.26.457296 (2021).
54. Dong, Y. et al. Parental legacy versus regulatory innovation in salt stress responsiveness of allopolyploid cotton (Gossypium) species. *Plant J.* **111**, 872–887 (2022).
55. Pei, L. et al. Dynamic 3D genome architecture of cotton fiber reveals subgenome-coordinated chromatin topology for 4-staged single-cell differentiation. *Genome Biol.* **23**, 45 (2022).
56. Khan, D. et al. Gene expression profiling reveals transcription factor networks and subgenome bias during Brassica napus seed development. *Plant J.* **109**, 477–489 (2022).
57. Fang, L., Guan, X. & Zhang, T. Asymmetric evolution and domestication in allotetraploid cotton (*Gossypium hirsutum* L.). *Crop J.* **5**, 159–165 (2017).
58. Combes, M.-C., Cenci, A., Baraille, H., Bertrand, B. & Lashermes, P. Homeologous gene expression in response to growing temperature in a recent allopolyploid (*Coffea arabica* L.). *J. Heredity* **103**, 36–46 (2011).
59. Griffiths, A. G. et al. Breaking free: The genomics of allopolyploidy-facilitated niche expansion in white clover. *Plant Cell* **31**, 1466–1487 (2019).
60. Coombes, B. et al. Whole genome sequencing uncovers the structural and transcriptomic landscape of hexaploid wheat/*Am. muticum* introgression lines. *bioRxiv* https://doi.org/10.1101/2021.11.16.468825 (2021).
61. Rey, E. et al. Transcriptome reprogramming due to the introduction of a barley telosome into bread wheat affects more barley genes than wheat. *Plant Biotechnol. J.* **16**, 1767–1777 (2018).

62. Dong, Z. et al. Genome-wide impacts of alien chromatin introgression on wheat gene transcriptions. *Sci. Rep.* **10**, 4801 (2020).
63. Li, H. et al. A spontaneous wheat-Aegilops longissima translocation carrying *Pm66* confers resistance to powdery mildew. *Theor. Appl Genet.* **133**, 1149–1159 (2020).
64. Bin, Z. et al. Transcriptional aneuploidy responses of *Brassica rapa-oleracea* monosomic alien addition lines (MAALs) derived from natural allopolyploid *B. napus. Front. Genet.* **10**, 67 (2019).
65. Shao, Y., Pan, Q., Zhang, D., Kang, L. & Li, Z. Global gene expression perturbations in rapeseed due to the introduction of alien radish chromosomes. *J. Genet.* **100**, 25 (2021).
66. Gallois, J.-L., Moury, B. & German-Retana, S. Role of the genetic background in resistance to plant viruses. *Int J. Mol. Sci.* **19**, 2856 (2018).
67. Sharma, S. et al. Introducing beneficial alleles from plant genetic resources into the wheat germplasm. *Biology* **10**, 982 (2021).
68. Szabo-Hever, A. et al. Genetic diversity and resistance to Fusarium head blight in synthetic hexaploid wheat derived from *Aegilops tauschii* and diverse *Triticum turgidum* subspecies. *Front. Plant Sci.* **9**, 1829 (2018).
69. Moskal, K., Kowalik, S., Podyma, W., Łapiński, B. & Boczkowska, M. The pros and cons of rye chromatin introgression into wheat genome. *Agronomy* **11**, 456 (2021).
70. Zhao, L. et al. Chromosome stability of synthetic-natural wheat hybrids. *Front. Plant Sci.* **12**, 654382 (2021).
71. Ozkan, H., Levy, A. A. & Feldman, M. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops–Triticum*) group. *Plant Cell* **13**, 1735–1747 (2001).
72. Yu, M., Guan, L.-L., Chen, G.-Y., Pu, Z.-E. & Hou, D.-B. Allopolyploidy-induced rapid genomic changes in newly generated synthetic hexaploid wheat. *Biotechnol. Biotechnol. Equip.* **31**, 236–242 (2017).
73. Gutbrod, M. J. & Martienssen, R. A. Conserved chromosomal functions of RNA interference. *Nat. Rev. Genet.* **21**, 311–331 (2020).
74. Yaakov, B. & Kashkush, K. Massive alterations of the methylation patterns around DNA transposons in the first four generations of a newly formed wheat allohexaploid. *Genome* **54**, 42–49 (2011).
75. Yuan, J. et al. Dynamic and reversible DNA methylation changes induced by genome separation and merger of polyploid wheat. *BMC Biol.* **18**, 171 (2020).
76. Liu, Y. et al. Histone H3K27 dimethylation landscapes contribute to genome stability and genetic recombination during wheat polyploidization. *Plant J.* **105**, 678–690 (2021).
77. Wang, M. et al. An atlas of wheat epigenetic regulatory elements reveals subgenome divergence in the regulation of development and stress responses. *Plant Cell* **33**, 865–881 (2021).
78. Lu, F.-H. et al. Reduced chromatin accessibility underlies gene expression differences in homologous chromosome arms of diploid *Aegilops tauschii* and hexaploid wheat. *GigaScience* **9**, giaa070 (2020).
79. Garcia-Lozano, M. et al. Altered chromatin conformation and transcriptional regulation in watermelon following genome doubling. *Plant J.* **106**, 588–600 (2021).
80. Wang, M. et al. Evolutionary dynamics of 3D genome architecture following polyploidization in cotton. *Nat. Plants* **4**, 90–97 (2018).
81. Kazan, K. Alternative splicing and proteome diversity in plants: the tip of the iceberg has just emerged. *Trends Plant Sci.* **8**, 468–471 (2003).
82. Laloum, T., Martín, G. & Duque, P. Alternative splicing control of abiotic stress responses. *Trends Plant Sci.* **23**, 140–150 (2018).
83. Huang, J. et al. Phytophthora effectors modulate genome-wide alternative splicing of host mRNAs to reprogram plant immunity. *Mol. Plant* **13**, 1470–1484 (2020).
84. Yu, K. et al. Changes in alternative splicing in response to domestication and polyploidization in wheat. *Plant Physiol.* **184**, 1955–1968 (2020).
85. Martín, G., Márquez, Y., Mantica, F., Duque, P. & Irimia, M. Alternative splicing landscapes in Arabidopsis thaliana across tissues and stress conditions highlight major functional differences with animals. *Genome Biol.* **22**, 35 (2021).
86. Min, X. J. et al. Genome-wide cataloging and analysis of alternatively spliced genes in cereal crops. *BMC Genomics* **16**, 721 (2015).
87. Wang, M. et al. A global survey of alternative splicing in allopolyploid cotton: landscape, complexity and regulation. *N. Phytol.* **217**, 163–178 (2018).
88. Zhou, R., Moshgabadi, N. & Adams, K. L. Extensive changes to alternative splicing patterns following allopolyploidy in natural and resynthesized polyploids. *Proc. Natl Acad. Sci. USA* **108**, 16122–16127 (2011).
89. Gao, P. et al. Alternative splicing dynamics and evolutionary divergence during embryogenesis in wheat species. *Plant Biotechnol. J.* **19**, 1624–1643 (2021).
90. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
91. Dobin, A. et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2012).
92. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).
93. Edger, P. P. et al. Subgenome Dominance in an Interspecific Hybrid, Synthetic Allopolyploid, and a 140-Year-Old Naturally Established Neo-Allopolyploid Monkeyflower. *Plant Cell* **29**, 2150–2167 (2017).
94. Kryuchkova-Mostacci, N. & Robinson-Rechavi, M. A benchmark of gene expression tissue-specificity metrics. *Brief. Bioinform.* **18**, 205–214 (2017).
95. Chen, Y. et al. A Collinearity-Incorporating Homology Inference Strategy for Connecting Emerging Assemblies in the Triticeae Tribe as a Pilot Practice in the Plant Pangenomic Era. *Mol. Plant* **13**, 1694–1708 (2020).
96. Shen, S. et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl Acad. Sci. USA* **111**, E5593–E5601 (2014).
97. rmats2sashimiplot. https://github.com/Xinglab/rmats2sashimiplot. (2015).
98. Vasudevan, A., Lévesque-Lemay, M., Edwards, T. & Cloutier, S. Global transcriptome analysis of allopolyploidization reveals large-scale repression of the D-subgenome in synthetic hexaploid wheat. *Figshare Repository* https://doi.org/10.6084/m9.figshare.c.6443621 (2023).

Acknowledgements

Author contributions

S.C. designed the study. S.C., M.L.-L. and T.E. performed the experiments and collected data. A.V. and S.C. finalized the data analysis methods, carried out data analysis, interpreted the results and visualized them. A.V. and S.C. wrote the paper. All the authors read and approved the paper.

Competing interests

The authors declare no competing interests.

Additional information