Parental legacy versus regulatory innovation in salt stress responsiveness of allopolyploid cotton (*Gossypium*) species

Yating Dong[1,2] (iD), Guanjing Hu[3,4,*] (iD), Corrinne E. Grover[2] (iD), Emma R. Miller[2] (iD), Shuijin Zhu[1,*] (iD) and Jonathan F. Wendel[2,*] (iD)

[1]*Department of Agronomy, Zhejiang University, Hangzhou, Zhejiang 310 053, China,*

[2]*Department of Ecology, Evolution, and Organismal Biology (EEOB), Bessey Hall, Iowa State University, Ames, IA 50011, USA,*

[3]*State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang 455 000, China, and*

[4]*Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518 120, China*

SUMMARY

Polyploidy provides an opportunity for evolutionary innovation and species diversification, especially under stressful conditions. In allopolyploids, the conditional dynamics of homoeologous gene expression can be either inherited from ancestral states pre-existing in the parental diploids or novel upon polyploidization, the latter potentially permitting a wider range of phenotypic responses to stresses. To gain insight into regulatory mechanisms underlying the diversity of salt resistance in *Gossypium* species, we compared global transcriptomic responses to modest salinity stress in two allotetraploid (AD-genome) cotton species, *Gossypium hirsutum* and *G. mustelinum*, relative to their model diploid progenitors (A-genome and D-genome). Multivariate and pairwise analyses of salt-responsive changes revealed a profound alteration of gene expression for about one third of the transcriptome. Transcriptional responses and associated functional implications of salt acclimation varied across species, as did species-specific coexpression modules among species and ploidy levels. Salt responsiveness in both allopolyploids was strongly biased toward the D-genome progenitor. A much lower level of transgressive downregulation was observed in the more salt-tolerant *G. mustelinum* than in the less tolerant *G. hirsutum*. By disentangling inherited effects from evolved responses, we show that expression biases that are not conditional upon salt stress approximately equally reflect parental legacy and regulatory novelty upon allopolyploidization, whereas stress-responsive biases are predominantly novel, or evolved, in allopolyploids. Overall, our work suggests that allopolyploid cottons acquired a wide range of stress response flexibility relative to their diploid ancestors, most likely mediated by complex suites of duplicated genes and regulatory factors.

Keywords: allopolyploidy, cotton, *Gossypium*, salt stress, transcriptomics.

INTRODUCTION

Polyploidy is a prevalent phenomenon among eukaryotes, especially in plants, and is recognized as a driving force in evolution and diversification (Blanc & Wolfe, 2004; Bowman et al., 2007; Wood et al., 2009; Jiao et al., 2011; Wendel, 2015; Van de Peer et al., 2017; Fox et al., 2020; Van de Peer et al., 2020). Polyploidy is also considered a key factor in domestication of crops such as wheat (*Triticum aestivum*), rapeseed (*Brassica napus*), cotton (*Gossypium* spp.), and soybean (*Glycine max*) (among others; Udall & Wendel, 2006; Salman-Minkov et al., 2016), as genome

doubling is thought to promote diversification and adaptation via various genetic and epigenetic processes (Udall & Wendel, 2006; Renny-Byfield & Wendel, 2014; Salman-Minkov et al., 2016; Ding & Chen, 2018). Allopolyploids, which combine two divergent parental genomes (Adams & Wendel, 2005; Doyle et al., 2008; Doyle & Sherman-Broyles, 2017), often show enhanced characteristics, e.g., better growth vigor (Solhaug et al., 2016), broader ecological adaptation (Shimizu-Inatsugi et al., 2017), and better stress tolerance (Liu et al., 2012; Yang et al., 2014), compared to their diploid ancestors (Comai, 2005; Fawcett &

Maere, 2009; Ni et al., 2009; Chen, 2010; Chen & Birchler, 2013). The increased competitiveness of allopolyploid species in environments that are challenging or unsuitable for the parental species is thought to be attributable to processes including fixed heterozygosity, intergenomic interactions, and expression dosage effects that accompany hybridization and genome doubling (Solhaug et al., 2016).

Allopolyploid success involves accommodating divergent genomes and transcriptomes within the same nucleus (Comai, 2005; Chen, 2007; Leitch & Leitch, 2008; Jackson & Chen, 2010), generating enormous combinatorial complexity and thus the potential for evolutionary novelty. Although additive expression of duplicated genes (homoeologs) may represent an expected outcome for polyploid species (where joint expression of homoeologs is equal to the arithmetic average or summed expression of parental species, i.e., the mid-parent value), non-additive expression has been widely documented in a variety of polyploid plants (Flagel & Wendel, 2010; Bardil et al., 2011; Grover et al., 2012; Chelaifa et al., 2013; Yoo et al., 2014; Wendel, 2015; Powell & Doyle, 2017), representing a deviation from vertical transmission of the pre-existing parental state, also known as parental legacy (Buggs et al., 2014). Examination of non-additive expression patterns at the whole transcriptome level (Pumphrey et al., 2009; Yoo et al., 2013; Powell et al., 2017; Takahagi et al., 2018; Yu et al., 2018; Peng et al., 2020) led to the recognition of the widely observed phenomena of homoeolog expression bias (HEB), expression level dominance, and transgressive expression of homoeologs (Grover et al., 2012; Yoo et al., 2014; Wu et al., 2018; Sigel et al., 2019; Li et al., 2020). These phenomena collectively represent regulatory novelty that either accompanied or evolved following allopolyploidization, often leading to phenotypic innovation in polyploids relative to their diploid parents. For example, conditional HEB during development and in response to environmental stress, respectively, have been implicated in causing delayed leaf maturation and impaired chlorophyll accumulation in *Cucumis* × *hytivus* (Yu et al., 2018) and influencing the response to *Fusarium* infection (Powell et al., 2017) in allohexaploid wheat. These and other studies support the notion that polyploid success is intimately related to stress responses (Van de Peer et al., 2020), although at present, the patterns and mechanisms by which polyploids alter homoeolog expression to accommodate stressful environments are not yet understood.

Cotton is a major crop with global economic significance as a natural source of textile fiber. Notably, two of the major cotton crop species, *Gossypium hirsutum* and *Gossypium barbadense*, are descendants of a single, natural allopolyploidization event that occurred approximately 1–2 million years ago (mya), leading to the evolution of seven modern polyploid species. The evolutionary history of *Gossypium* L. (the cotton genus) is well known (Wendel & Grover, 2015; Hu et al., 2021) and model diploid progenitors for the allopolyploid clade have been identified (*Gossypium arboreum* and *Gossypium herbaceum*, maternal; *Gossypium raimondii*, paternal; Wendel & Cronn, 2003). *Gossypium* contains more than 50 recognized species that are extensively distributed worldwide in arid and semi-arid regions of the tropics and subtropics (Endrizzi et al., 1985; Wendel et al., 2010; Wendel & Grover, 2015; Wang et al., 2018). In general, there exists a geographic distinction in natural distribution of diploid cotton species (mostly found inland) versus polyploids (largely species of littoral environments), which suggests an association between polyploidization and ecological adaptation. However, the underlying genetic basis and molecular processes still remain to be discovered. In terms of the consequences of polyploidy for gene expression, HEB and other phenomena have been globally characterized in various tissues and/or developmental stages of allopolyploid cotton (Rapp et al., 2009; Flagel & Wendel, 2010; Hu et al., 2013; Yoo et al., 2013), but the conditional changes under environmental stress have been minimally surveyed (Dong & Adams, 2011).

As a major abiotic stress, soil salinity limits plant growth and causes substantial yield loss worldwide (Munns & Tester, 2008). Although cotton is classified as a moderately salt-tolerant crop, soil salinity can still inhibit cotton growth and fiber production, depending upon the genotype and growth stage (Abdelraheem et al., 2019). In a recent survey of salt tolerance among *Gossypium* species/accessions, we found highly variable responses to a 2-week treatment of hydroponic salinity (50 and 100 mM NaCl) in allopolyploids relative to the diploid progenitors (Dong et al., 2020). Although the ranking of allopolyploids with respect to salt tolerance is generally intermediate between the more tolerant A-genome and the less tolerant D-genome diploids, the allopolyploid species *Gossypium mustelinum* was most tolerant among all genotypes surveyed. Another interesting finding was that under control conditions, the morphological and physiological traits of allopolyploids surveyed were more like those of the A-genome diploid progenitor, whereas under salt stress the opposite pattern (D-like traits) was observed. To further understand the underlying mechanisms of diversity and divergence in salt tolerance following polyploidy, here we investigated the transcriptomic response to salt stress for two allopolyploid cotton species (*G. hirsutum* and *G. mustelinum*) relative to their model diploid progenitors, *G. arboreum* (A-genome) and *G. raimondii* (D-genome). Based on the *total* expression of homoeologous gene pairs, we conducted differential expression and coexpression analyses in conjunction with our previously generated morpho-physiological data (Dong et al., 2020). We also studied *partitioned* homoeolog expression, investigating stress responsiveness in allopolyploid cottons. By comparing both the *total* and *partitioned*

expression patterns between the salt-sensitive *G. hirsutum* and salt-tolerant *G. mustelinum*, our study provides new evidence regarding how polyploids gain better tolerance to salt stress.

RESULTS

RNA-seq analysis of salt tolerance in four *Gossypium* species

To gain insight into the underlying mechanisms of divergence in salt tolerance among *Gossypium* species, we implemented RNA-seq to survey the global transcriptomic response to modest salinity stress (50 mM NaCl) in four cotton species, including two allopolyploids, i.e., *G. hirsutum* (AD$_1$) and *G. mustelinum* (AD$_4$), as well as representatives of their diploid progenitors, A-genome *G. arboreum* (A$_2$) and D-genome *G. raimondii* (D$_5$). All surveyed species were able to survive the salinity treatment while displaying different levels of stress symptoms (e.g., inhibited growth and smaller, thicker leaves) in treated versus control plants. Consistent with our previous findings (Dong et al., 2020), *G. mustelinum* was most tolerant, followed sequentially by *G. arboreum*, *G. raimondii*, and *G. hirsutum*.

A total of 24 RNA-seq libraries were obtained with an average of 48 million paired-end reads sequenced per library, of which 95.9–98.8% were mapped onto the reference genome of *G. raimondii* (Table S1). For the allopolyploid species, approximately equal proportions of A- and D-homoeolog-specific reads were assigned (Table S1); in the following analyses, the *total* expression for each homoeologous pair is referred to as gene expression unless further specified. Principal component analysis (PCA) of the *total* expression revealed clear clustering of replicated samples, except for one replicate from the *G. raimondii* control group, which instead clustered with the allopolyploids (Figure 1a); a consistent pattern was also observed for the subgenome *partitioned* expression (Figure S1b). Hence, this mislabeled *G. raimondii* sample was
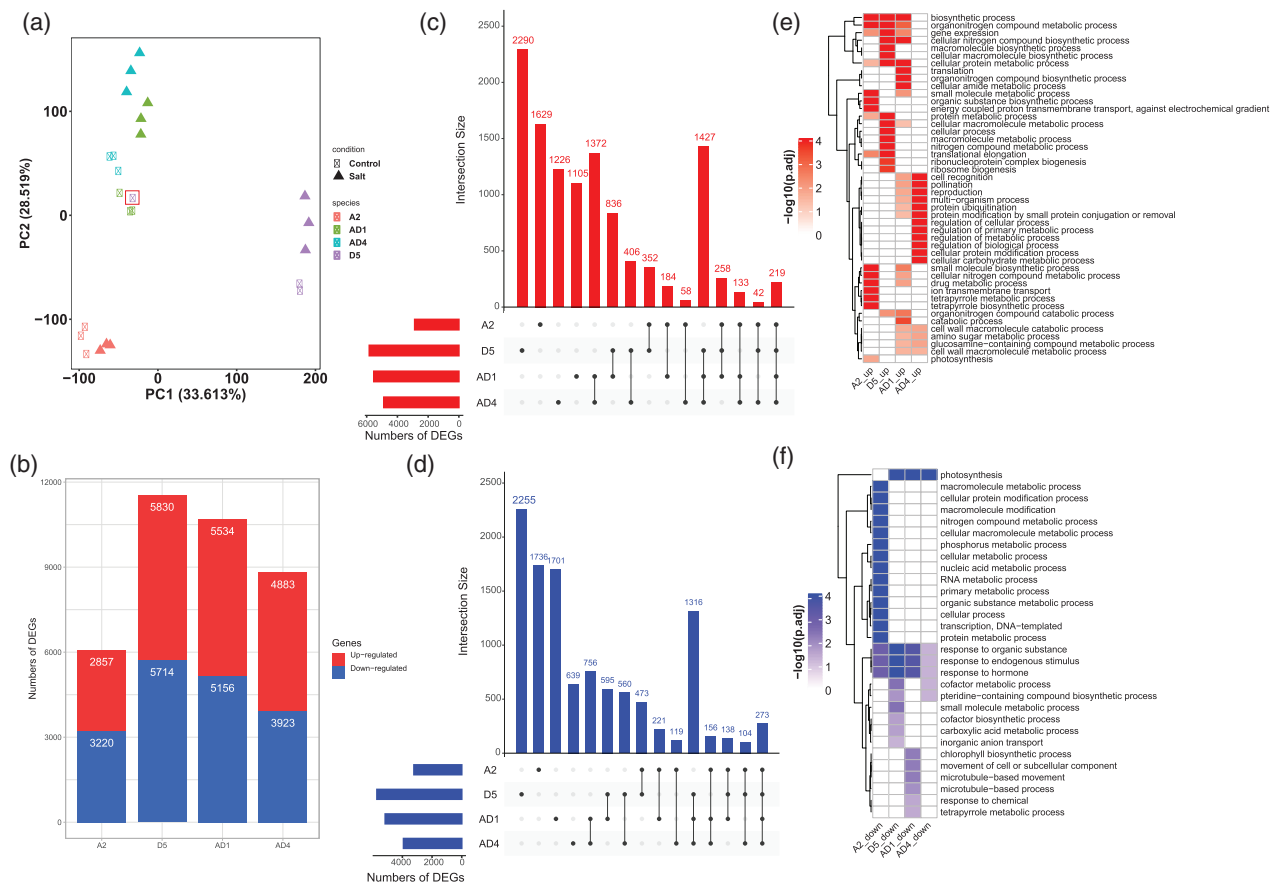
Figure 1. Transcriptomic diversity and salt-responsive dynamics in cotton. (a) Principal component analysis (PCA) of 24 RNA-seq samples across diploid and allopolyploid *Gossypium* species. The mislabeled D$_5$ sample is circled in red. (b) Barplot of differentially expressed genes (DEGs) in response to salt stress treatment in each species. (c, d) UpSet plots of upregulated (c) and downregulated (d) DEGs and their overlap between species. The bottom left horizontal barplot shows the total number of DEGs identified in each species. The top barplot summarizes the number of DEGs for each unique or overlapping combination as indicated by the connecting circles below. (e, f) Gene Ontology (GO) enrichment analysis of upregulated and downregulated DEGs in each species. The top 15 significant, non-redundant GO terms of each species are shown. The color intensity in both panels denotes $-\log_{10}$ transformed 'adjusted *P*-values', where darker red/blue indicates a higher degree of enrichment.

excluded from further analysis. The first two principal components (PCs) accounted for around 62% of the total variance: PC1 mainly separated *G. raimondii* samples from the others, while PC2 was able to separate the stress-treated and control samples per species except for *G. arboreum*.

Transcriptomic changes in response to salt treatment in both diploid and allopolyploid cottons

A primary goal of this study was to explore the transcriptomic diversity of salt stress responses among the four *Gossypium* taxa. To characterize the interspecific divergence of salt stress responses, we first applied a multifactor design (*species* + *treatment* + *species*:*treatment*) to all samples (Table 1). A significant *species* effect was identified for 29 054 genes, or approximately 80% of the transcriptome, reflecting massive interspecific differences regardless of the salt stress treatment. A significant *treatment* effect common to all four species was identified for 13 378 (35.9%) genes. Gene Ontology (GO) enrichment analysis revealed that the upregulated genes were primarily involved in oxidoreductase, peptidase, ATPase, and GTPase activity, peptidase activity, and various cellular biosynthetic and catabolic processes, whereas the downregulated genes were primarily enriched in photosynthesis, ion binding, and response to endogenous stimulus (Table S2 and Figure S2). The *species*:*treatment* interaction effect reflects species-specific responses to salt treatment, which was significant for 12 979 (34.9%) genes enriched in GO terms related to ubiquitin-protein transferase activity, the oxidoreductase complex, and photosynthesis (Table S2).

We next conducted pairwise comparisons between control and treated samples for each species to further examine the effect of salt treatment. The numbers of differentially expressed genes (DEGs) in response to salt stress per species ranged from 6077 to 11 544 genes, corresponding to 16.2–30.8% of the leaf transcriptome (Figure 1b). The diploid species *G. arboreum* and *G. raimondii* exhibited the lowest and highest number of DEGs, respectively, while the allotetraploid species *G. hirsutum* and *G. mustelinum* showed intermediate amounts of change between their diploid parents. Intersections between these DEG sets showed that *G. arboreum* and *G. raimondii* exhibited the largest number of species-specific changes for both up- and downregulation (Figure 1c,d, respectively). The common

set of DEGs shared by the two allopolyploids (1372 upregulated and 756 downregulated) was much larger than the common set shared by two diploids (352 upregulated and 473 downregulated), perhaps reflecting the approximately fivefold greater time since divergence of the diploids than since divergence of the two polyploid species (Wendel & Grover, 2015; Hu et al., 2021). Interestingly, the number of shared DEGs between both allopolyploid cottons and *G. raimondii* (1427 upregulated and 1316 downregulated) was remarkably higher than the number of DEGs shared with *G. arboreum* (133 upregulated and 156 downregulated). Only a small set of DEGs (219 upregulated and 273 downregulated) overlapped in all four species. Enriched GO terms of pairwise DEG sets were cataloged (Table S3; Figures S3 and S4) and compared to assess whether the associated biological functions are common or species-specific (Figure 1e,f). For example, photosynthesis was upregulated in *G. arboreum* yet downregulated in the other species, while the response to hormones was uniformly enriched in the downregulated DEGs. Notably, no GO term was enriched in upregulated DEGs across species. These pairwise comparisons, together with the multivariate analysis results, suggest that the transcriptional response to salt stress and the involved cellular processes differed significantly among cotton species.

Coexpression gene network analysis uncovers divergent responsiveness to salt stress

To further investigate global transcriptional organization underlying salt responses and interspecific divergence, we applied weighted gene coexpression network analysis (WGCNA) to all samples across the different cotton species. A total of 35 network modules were identified, containing from 102 to 4946 coexpressed member genes (Figure S5). By relating the representative coexpression pattern (i.e., module eigengene) of each module to the 29 phenotypic traits previously measured (Dong et al., 2020), we found that nearly all modules (except ME23 and ME32) were significantly correlated with at least one trait (Figure S5d). More than half of these modules showed significant correlations with either the categorical 'species' (i.e., samples from the species of interest were labeled 1; other sample were labeled 0) or 'condition' traits (i.e., the treated samples were labeled 1; other samples were labeled 0) (Figure 2a). Eight modules exhibiting a significant 'condition' correlation, including ME4, ME6, ME10, ME16, ME17, ME20, ME22, and ME30 (Figure 2b), were further examined to characterize the salt-responsive coexpression dynamics and functional association.

According to the salt tolerance ranking ($AD_4 > A_2 > D_5 > AD_1$) reported previously, we were particularly interested in modules exhibiting expression profiles correlated with the ranking; however, no such modules were identified. Among the eight salt-responsive modules, ME6 and ME17 were

Table 1 Multivariate analysis of differential gene expression between species and under salt stress treatment

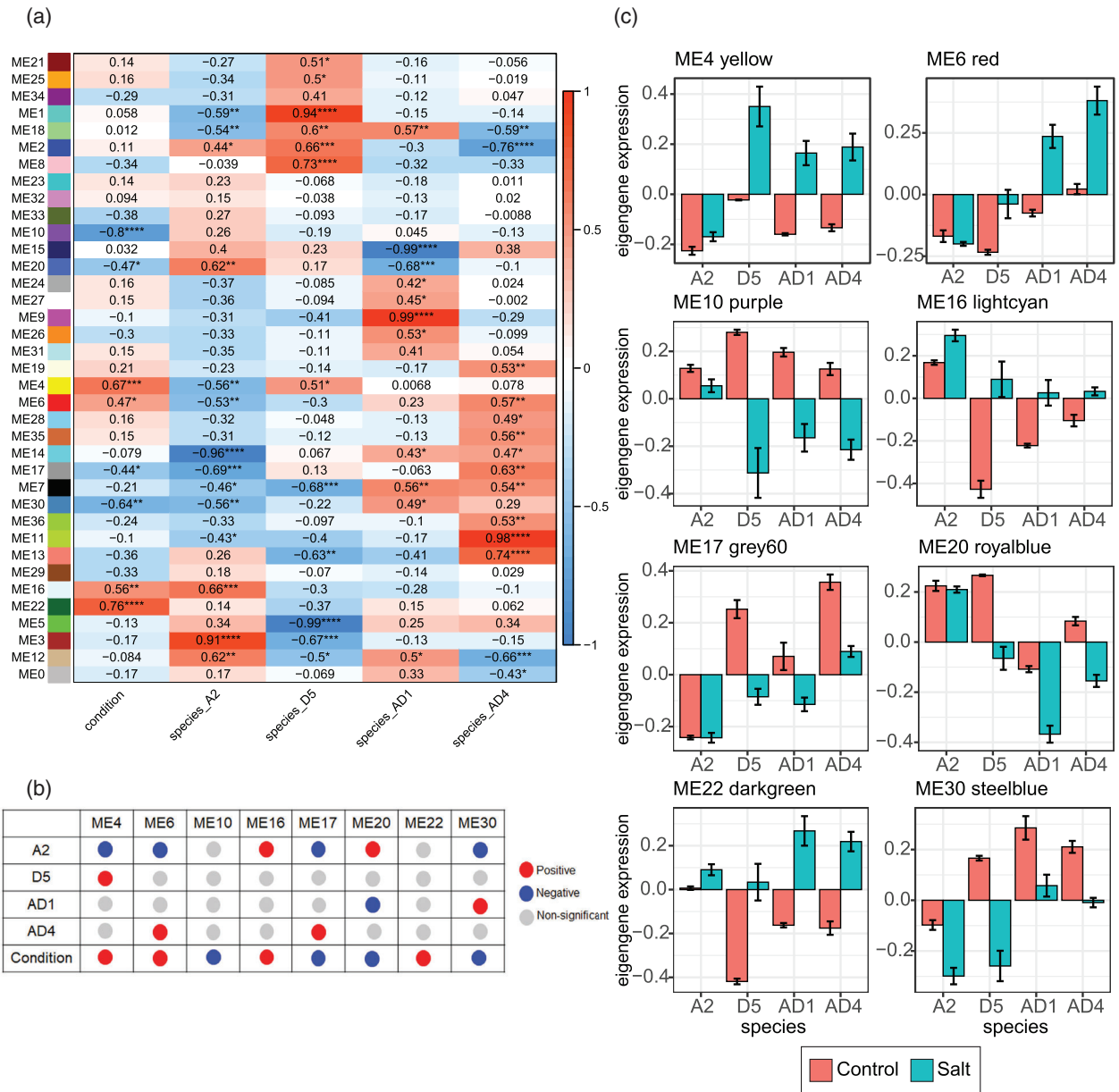| Effect	Differentially expressed	Upregulated	Downregulated
Species	29 054	–	–
Treatment	13 378	6852	6526
Species × treatment	12 979	–	–

Figure 2. WGCNA of four cotton species in response to salt stress. (a) Heatmap showing Pearson correlation between module eigengenes (by row) and categorical traits of 'condition' and 'species' (by column). The 'condition' trait is labeled 1 for the salt stress-treated samples and 0 for the untreated samples. The 'species_A_2' trait, for example, is labeled 1 for A_2 and 0 for the other samples. Each cell shows the correlation coefficient with significance as follows: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$. (b) Eight modules of interest with significant correlations. Red circles indicate positive correlations, blue circles indicate negative correlations, and gray circles indicate non-significant correlations. (c) Barplots of module eigengenes, with error bars denoting the standard deviation calculated from replicates.

significantly correlated with the two more tolerant species *G. arboreum* and *G. mustelinum* (Figure 2b). Interestingly, both modules were positively correlated with *G. mustelinum* but negatively correlated with *G. arboreum*; that is, the highest and the lowest expression levels were observed in *G. mustelinum* and *G. arboreum*, respectively. No GO terms were significantly enriched for ME17, but the functional enrichment of ME6 pointed to biological regulation, oxidation–reduction, reproduction, and protein phosphorylation

(Table S5). ME4 showed the lowest expression in *G. arboreum*, whereas ME16 showed the highest expression in *G. arboreum* (Figure 2c and Table S5), both exhibiting increased expression by salt treatment and enriched for 'translation'. ME20 and ME30 showed opposite AD_1 and A_2 'species' correlation patterns, with both exhibiting decreased expression by salt treatment; no GO terms were significantly enriched, probably due to their small module sizes (281 and 134 member genes, respectively).

Without significant 'species' correlations, ME10 exhibited reduced expression by salt treatment in all four cotton species (condition-specific); this module was significantly enriched with GO terms related to photosynthesis (Figure 2b). These results are also consistent with the positive correlation between ME10 and photosynthetic-related traits (Figure S5d). ME22 exhibited the opposite condition-specific pattern to that of ME10, where the genes whose expression increased during salt treatment were strongly enriched for GO terms associated with 'endoplasmic reticulum membrane'. Together, our module-specific GO analysis further indicated that salt resistance is mediated by divergent strategies in the four closely related *Gossypium* species.

Gossypium mustelinum shows lower level of transgressive downregulation upon salt stress than does *G. hirsutum*

Despite their shared ancestry, the two allopolyploid cottons represent the most (*G. mustelinum*) and least (*G. hirsutum*) salt-tolerant species surveyed here, while their model progenitors were intermediate to the two, prompting the question of how these two allotetraploid cottons respond to salt treatment relative to each other and to their parental *G. arboreum* and *G. raimondii* diploids. For each trio of allopolyploid + diploids (i.e., AD_1 versus A_2 versus D_5 or AD_4 versus A_2 versus D_5), DEGs with significant *species:treatment* interaction effects from our initial analyses were extracted to compare their expression changes by salt treatment (i.e., $\log_2$(fold change) values). Heatmap and clustering analysis revealed that gene regulation patterns in both polyploids were first clustered together with *G. raimondii* rather than *G. arboreum* (Figure 3), consistent with the higher number of shared pairwise DEGs between both allopolyploids and *G. raimondii* (Figure 1c,d).

Additionally, both comparisons exhibited allopolyploid-specific expression changes outside the range of the parental diploids (Figure 3), either upregulated (red box, termed as 'transgressive upregulation') or downregulated (blue box, termed as 'transgressive downregulation') in the
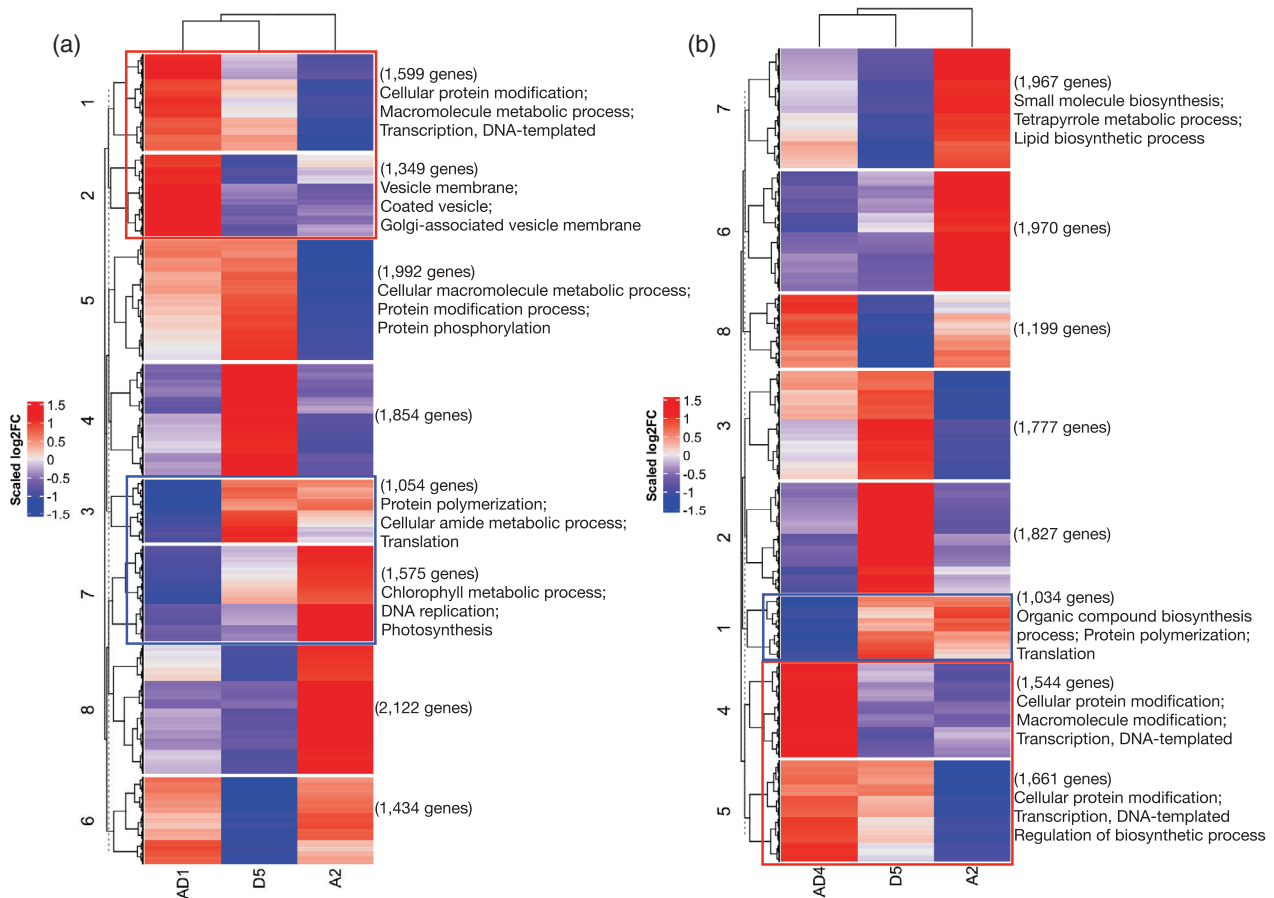
Figure 3. Heatmap of expression changes by salt treatment in both allopolyploids relative to their parental parents. (a) The trio of AD_1, A_2, and D_5. (b) The trio of AD_4, A_2, and D_5. For 12 979 genes with significant *species:treatment* interaction effects, the $\log_2$ treated/control fold change values were normalized and scaled by *z*-score in each trio. Beside each gene cluster, the cluster size and representative enriched functional classification terms are noted. Clusters displaying transgressive upregulation in allopolyploids are boxed in red, and clusters displaying transgressive downregulation are boxed in blue.

allopolyploid. Interestingly, the number of genes that showed transgressive upregulation in *G. mustelinum* (3205 genes, clusters 4 and 5) was higher than that in *G. hirsutum* (2948 genes, clusters 1 and 2). In contrast, there were 2629 genes (clusters 3 and 7) transgressively downregulated in *G. hirsutum*, and only 1034 genes were transgressively downregulated in *G. mustelinum* (cluster 1). These results suggest that salt exposure provokes a less dramatic transcriptional response in *G. mustelinum* than in *G. hirsutum* and that the lower level of transgressive downregulation in *G. mustelinum* may account for its higher resistance to stress. GO enrichment results of the transgressively upregulated genes were generally consistent between *G. hirsutum* and *G. mustelinum*; i.e., GO terms linked to 'protein modification', 'gene expression regulation', and 'macromolecule metabolism' were significantly overrepresented in both. On the contrary, the transgressively downregulated genes in *G. hirsutum* were associated with categories such as photosynthesis, DNA replication, and amide metabolism, whereas transgressively downregulated genes in *G. mustelinum* were specifically involved in protein polymerization (full GO lists are provided in Table S6). In addition to the transgressive regulatory patterns, clusters displaying expression changes intermediate to those of diploids were also observed, with *G. mustelinum* (clusters 3 and 7, 3744 genes) containing more genes than those of *G. hirsutum* (cluster 5, 1992 genes).

Stress-responsive homoeolog expression bias is novel in allopolyploids

In addition to the above analyses of *total* homoeolog expression in allopolyploids, we next asked whether salt stress treatment also affects relative *homoeologous* contributions to the total expression. In other words, we tested whether patterns of HEB were different between treated and control samples. PCA of subgenomic expression profiles (Figure S1b) showed that the subgenomic expression in allopolyploid homoeologs is, in general, closer to that of their corresponding diploids than to each other. This observation suggested that expression divergence between homoeologs in the allopolyploids mostly reflects divergence between parental diploids, or 'parental legacy' (Buggs et al., 2014). To further characterize the homoeolog expression divergence that is novel in each allopolyploid and/or in response to salt stress treatment, we applied a multifactor design (*homoeolog* + *ploidy* + *treatment* + *homoeolog*:*ploidy* + *homoeolog*:*treatment* + *ploidy*:*treatment* + *homoeolog*:*ploidy*:*treatment*) to each quartet of (sub)genome-specific gene expression profiles (e.g., AD_1:At versus AD_1:Dt versus A_2 versus D_5). Four types of HEB patterns were categorized to dissect the static and conditional HEB for each allopolyploid relative to the parental diploids. In *G. hirsutum* for example (Figure 4), a total of 4738

homoeolog pairs exhibited type I HEB characterized by a significant *homoeolog* effect, which represent A versus D expression divergence at both the diploid (A_2 versus D_5) and allopolyploid (At versus Dt) levels regardless of salt treatment; in other words, this category of HEB was inherited from the diploid parental state and is constitutive. Type II HEB, characterized by a significant *homoeolog*:*polyploid* interaction effect, was detected for 4787 homoeolog pairs and represents novel HEB patterns not previously detected in diploids. In contrast to type I and type II patterns, which are *not* condition-specific, type III and type IV HEB are responsive to salt treatment, as characterized by the significant *homoeolog*:*treatment* and *homoeolog*:*treatment*:*polyploidy* interaction effects, respectively. Like type I, type III effects represent those inherited from the parental state and type IV are salt-responsive effects novel in allopolyploids. Whereas inherited and novel categories for both polyploid species are nearly equivalent under normal conditions (types I and II, Figure 4a), both polyploids exhibit a similar excess of novel HEB during salt stress (average 152 versus 1732 homoeologous pairs; Figure 4b), suggesting that the salt-responsive HEB was predominantly novel upon allopolyploidization rather than inherited from diploids.

For genes exhibiting stress-responsive HEB (types III and IV), we further examined the magnitude and direction of changes in bias between control and treated conditions (Table 2). Under each condition, pairwise At versus Dt comparisons revealed more biases under salt treatment than under control conditions in both allopolyploid species (*G. hirsutum*, 24.2% versus 10.5%; *G. mustelinum*, 18.3% versus 12.9%). In response to the salt stress treatment, 387 and 282 homoeologous pairs gained new HEB patterns in *G. hirsutum* and *G. mustelinum*, respectively, while an additional 125 and 179 pairs lost the HEB exhibited under control conditions; fewer pairs switched direction of bias. GO analyses revealed that genes with conserved patterns of HEB were enriched in photosynthesis-related terms in both species (Table S7). Genes that gained bias in *G. hirsutum* under salt conditions were associated with catalytic activity and transcription factors while no statistically significant GO terms were enriched for those that gained bias in *G. mustelinum*, which might suggest diverse functions. Because no significant difference was detected between the numbers of A- and D-biased genes, the overall HEB was considered balanced, as previously defined (Grover et al., 2012), in all scenarios (Table 2). For example, of the 191 homoeolog pairs exhibiting HEB in *G. hirsutum* under control conditions, 89 homoeologous pairs exhibited At bias while 102 exhibited Dt bias (Chi-squared test, $P > 0.05$).

Stress-responsive HEB can alter expression in one or both homoeologs

Although all species studied generally exhibited balanced HEB, we consider the possibility that salt stress may

(a)

(b)
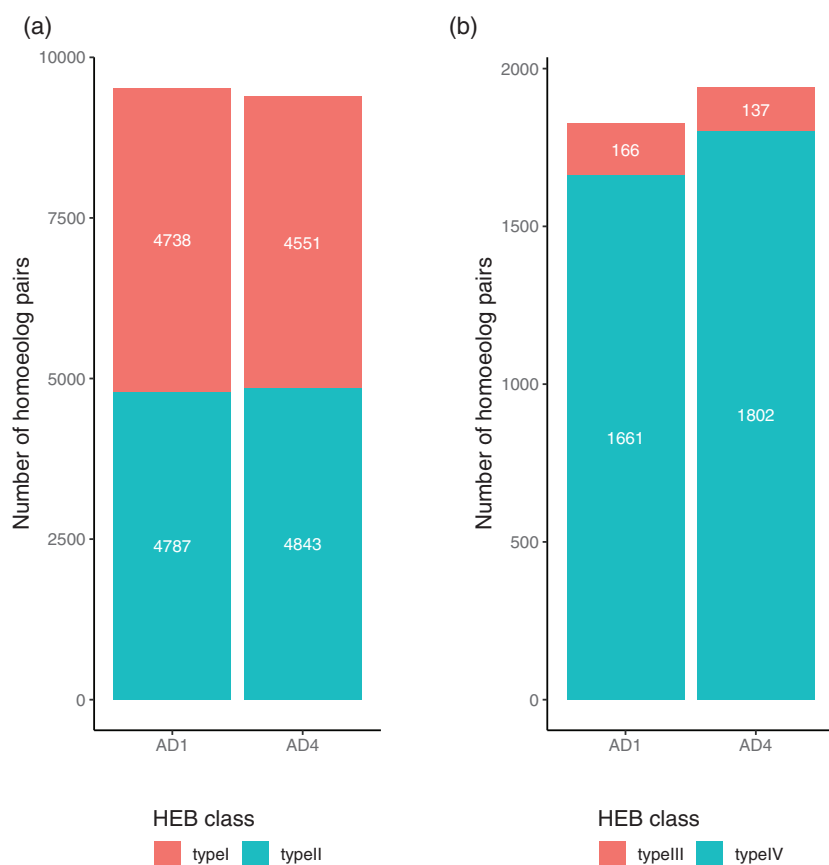
Figure 4. Homoeolog expression bias (HEB) categorization in *G. hirsutum* (AD$_1$) and *G. mustelinum* (AD$_4$). (a) Constitutive, not affected by salt stress. (b) Interaction, affected by salt treatment.

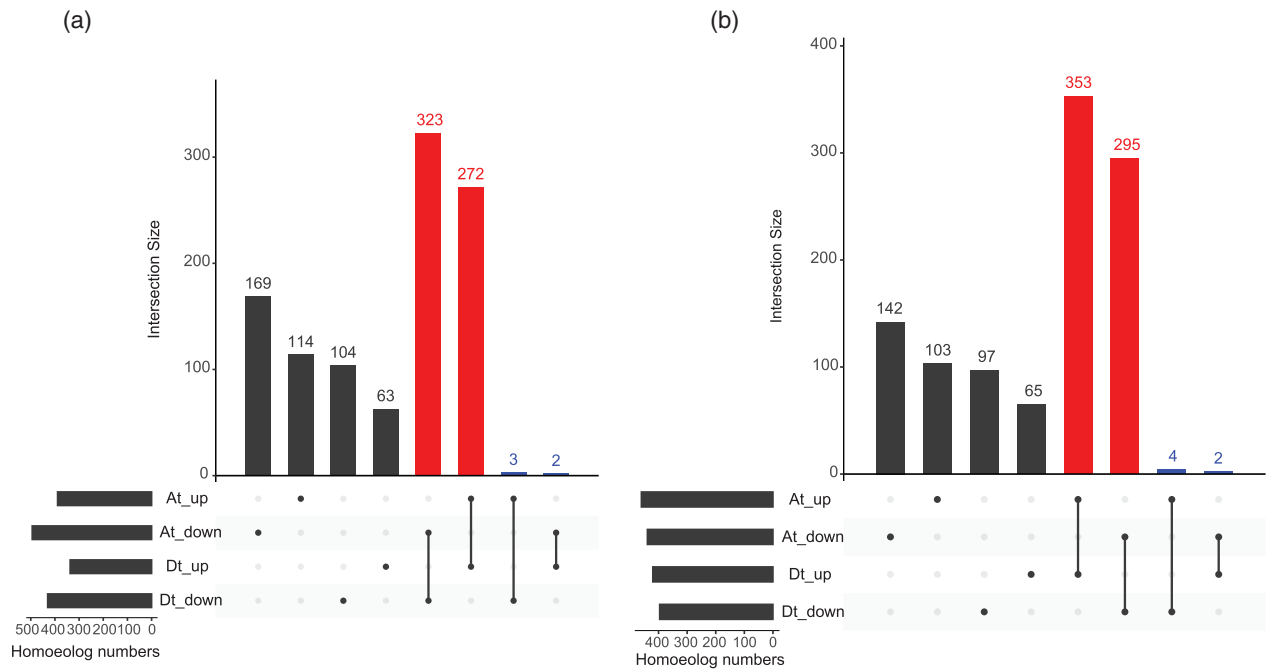**Table 2** Stress-responsive homoeolog expression bias (HEB) in allopolyploid cottons

Bias pattern	Control	Salt	AD1 Number (%)	AD4 Number (%)
Direction conserved	At = Dt	At = Dt	1259 (68.9%)	1397 (72.4%)
	At > Dt	At > Dt	28 (1.5%)	28 (1.5%)
	At < Dt	At < Dt	28 (1.5%)	35 (1.8%)
Bias gain by stress	At = Dt	At > Dt	176 (9.6%)	150 (7.8%)
		At < Dt	201 (11.0%)	132 (6.8%)
Direction switched	At > Dt	At < Dt	5 (0.3%)	3 (0.2%)
	At < Dt	At > Dt	5 (0.3%)	6 (0.3%)
Bias loss by stress	At > Dt	At = Dt	56 (3.1%)	96 (5.0%)
	At < Dt		69 (3.1%)	83 (4.3%)
Total At bias under control conditions			89 (4.9%)	127 (6.5%)
Total Dt bias under control conditions			102 (5.6%)	124 (6.4%)
Total At bias under salt conditions			209 (11.4%)	184 (9.5%)
Total Dt bias under salt conditions			234 (12.8%)	170 (8.8%)

At = A-homoeolog; Dt = D-homoeolog; At = Dt, equal expression of homoeologs; At > Dt, biased expression of the At homoeolog; At < Dt, biased expression of the Dt homoeolog; AD1 = *Gossypium hirsutum*; AD4 = *Gossypium mustelinum*.

disproportionately alter homoeolog usage from one or the other of the two divergent and coresident genomes ('At' and 'Dt') in allopolyploid cotton, a phenomenon known as 'homoeolog induction bias', previously proposed in hexaploid wheat (Powell et al., 2017). For homoeologous pairs exhibiting stress-responsive HEB in *G. hirsutum*, 283 pairs

exhibited significant expression alteration of only the At homoeolog and 167 pairs exhibited significant expression alteration of only Dt (Figure 5a, first four black sets in top barplot); 595 pairs exhibited significant changes of both At and Dt in the same direction (middle two red sets); and only five pairs exhibited changes in opposite directions

(a)

(b)

Figure 5. Identification of the homoeolog regulation patterns in response to salt stress in *G. hirsutum* (AD$_1$) (a) and *G. mustelinum* (AD$_4$) (b). For homoeologous pairs exhibiting stress-responsive HEBs, further differential expression analyses were applied to identify regulation patterns after salt treatment. For example, this analysis identified 389 At homoeologs that were upregulated and 494 that were downregulated in response to salinity stress in AD$_1$. Also shown are UpSet plots with counts of differentially expressed homoeologs in each pattern. Counts represent homoeologous genes where one (no intersection) or both (intersection of two circles) homoeologs were differentially expressed.

(last two blue sets). A similar pattern was observed in *G. mustelinum*, where 245 and 162 pairs exhibited significant expression alteration of At and Dt, respectively; 648 pairs exhibited significant changes in both At and Dt in the same direction; and only six pairs exhibited changes in opposite directions (Figure 5b). These results indicate that the salt treatment was more likely to affect both homoeologs than only one homoeolog (42.9% versus 57.1% in *G. hirsutum*; 38.4% versus 61.6% in *G. mustelinum*); however, when only one homoeolog was affected, it was more frequently an expression change in the At homoeolog, possibly indicating that the A-subgenome is more responsive to salt stress than the D-subgenome in allopolyploid cottons. Notably, *G. hirsutum* showed more down- than upregulation of both homoeologs, whereas *G. mustelinum* showed more up- than downregulation of both homoeologs, congruent with the aforementioned results.

DISCUSSION

Gossypium species exhibit divergent global transcriptional responses to salt stress

Salt tolerance is a complex quantitative trait that involves many genes and pathways. Global analysis of the stress-responsive gene expression profiles have been widely applied to understanding the molecular mechanisms and evolution of salt tolerance in plants (Edelist et al., 2009;

Peng et al., 2014; Guo et al., 2015; Zhang et al., 2016; Lei et al., 2018; Yuan et al., 2019; Wang et al., 2021). In spite of many previous RNA-seq studies in cotton, the transcriptional landscape of salt tolerance diversity between allopolyploid and diploid cottons has not been explored. Based on our previous understanding of the diverse morphological and physiological responses of 17 *Gossypium* species and genotypes (Dong et al., 2020), we selected four genotypes to examine their transcriptomic responses to relatively long-term (14 days treatment) and mild salinity (50 mM NaCl) stress. The two tetraploid species *G. hirsutum* (AD$_1$) and *G. mustelinum* (AD$_4$) represent the least and the most salt-tolerant allopolyploid genotypes, respectively, outside the range of two parental diploid species *G. arboreum* (A$_2$) and *G. raimondii* (D$_5$). Aiming to dissect the transcriptional changes underlying the diversity of salt resistance in *Gossypium*, we specifically asked how the tetraploid species acquired a wider range of tolerance than their parental diploids and why the two allopolyploid species exhibit such high differences in salt tolerance.

Our transcriptomic results demonstrated that salt stress has a profound global effect that alters gene expression for about one third of the entire transcriptome across four species. Phenotypically, we previously observed that nearly all of the phenotypic and physiological traits related to the salt response were altered after the same treatment (Dong et al., 2020). These observations together lend support to

the notion that plant salt tolerance is a complex emergent phenotype with a broad genomic underpinning (Gugger et al., 2017). Across the four species, multivariate and pairwise analyses of salt-responsive changes revealed diversified transcriptional programs and functional implications of salt acclimation (Figure 1), and subsequent coexpression network analysis further identified species-specific coexpression profiles that distinguish the functional responsiveness between species and ploidy levels (Figure 2). For example, the relatively low number of DEGs found in *G. arboreum* suggests that this species requires less transcriptional adjustment responding to salt stress compared to the other cottons; this pattern was also evident in the coexpression network profiles (Figure 2c). One potential explanation is that *G. arboreum* presents a more 'stress-ready' state that is pre-adapted for potential salt stress, in a manner similar to many halophytes or extremophytes (Kazachkova et al., 2013). Unfortunately, no wild form of *G. arboreum* has ever been discovered, so nothing is known about the conditions under which it evolved.

Interestingly, we found a D-dominant salt-responsive pattern in both allopolyploids, evidenced by the high number of salt stress DEGs shared between both tetraploids and *G. raimondii* (Figures 1c,d), as well as by the observation that the salt-induced expression changes in both tetraploids were more similar to those of *G. raimondii* than to those of *G. arboreum* (Figure 3). These results provide a transcriptional basis for the phenotypic observation that the allopolyploid traits were more like those of *G. raimondii* under salt stress; interestingly, the allopolyploid traits under control conditions were more like those of *G. arboreum* instead (Dong et al., 2020). Hence, we conclude that the D-genome progenitor, modeled here by *G. raimondii*, acts as the 'dominant' parent of allopolyploid cotton with respect to salt responsiveness. This conditional asymmetry in transcriptional profiles between allopolyploids and their diploid progenitors has previously been reported in allopolyploid coffee (*Coffea arabica*) under two temperature conditions (Bardil et al., 2011; Combes et al., 2013).

Two allopolyploid species, *G. hirsutum* and *G. mustelinum*, exhibit the lowest and highest tolerance among the four surveyed species, respectively, outside the range of parental diploids. From the perspective of functional genomics, it is of particular interest to characterize the molecular basis of higher tolerance in *G. mustelinum*, which can be the key for improving salt tolerance in cultivated *G. hirsutum*. Coexpression network analysis revealed that the module ME6 specifically showed the highest level of upregulation by salt treatment in *G. mustelinum*, which links its salt tolerance phenotype to effective scavenging of ROS, since genes in this module were significantly enriched in the GO term 'oxidation–reduction process' (Table S5), as reported in some halophytes (Flowers &

Colmer, 2008). This is in agreement with our previous physiological results that *G. mustelinum* showed highest peroxidase (POD) activities under mild salinity stress (Dong et al., 2020). Sodium can enter plant cells through non-selective cation channels that are regulated by salinity-induced signals such as ROS and Ca^{2+} (Deinlein et al., 2014; van Zelm et al., 2020). In addition, the enrichment of phosphorylation-related processes in this module (Table S5) implied that protein kinases involved in calcium flux mediation might have been induced during sodium signaling in *G. mustelinum*. For instance, the calcineurin B-like proteins (CBLs) bind calcium and cause protein phosphorylation through their interaction with CBL-interacting protein kinases (CIPKs), and such CBL–CIPK pathways have been widely studied (reviewed in Manishankar et al., 2018). As the best characterized CBL–CIPK pathway, the salt overly sensitive (SOS) pathway has been proposed to mediate cellular signaling during salt stress (Ji et al., 2013). By phosphorylating the H^+/cation antiporters, SOS complexes can transport sodium out of the cell, thus maintaining ion homoeostasis (Halfter et al., 2000; van Zelm et al., 2020). To some extent, this can be reflected by our previous physiological data that *G. mustelinum* showed a higher K^+/Na^+ ratio under salt treatment (Dong et al., 2020).

Balanced homoeolog expression bias in response to salt stress in allopolyploid cottons

In addition to providing insights into the genetic underpinnings of salt responsiveness among *Gossypium* species, our transcriptome analysis also has relevance with respect to long-standing questions regarding the functional and adaptive significance of polyploidy with respect to stress (Van de Peer et al., 2017; Van de Peer et al., 2020). Central to the connection between polyploidy and stress is the increased genomic complexity following whole genome duplication, which has been hypothesized to enable the rapid evolution of stress adaptation, hence enabling the establishing of polyploids under stressful environments. Proposed mechanisms responsible for increased biotic or abiotic stress tolerance include heterosis, expression dosage due to increased gene copy number, and functional plasticity of duplicate genes in polyploid species (Chen, 2007; Liu & Adams, 2007; Dong & Adams, 2011; Madlung, 2013; Liu et al., 2015; Takahagi et al., 2018).

At the transcriptome level, HEB is a widely observed phenomenon in plant allopolyploids. Earlier empirical examples suggest that allopolyploids selectively express homoeologs relevant for an appropriate response to environmental stresses. For instance, the allotetraploid grass *Brachypodium hybridum* shows long-term heat stress tolerance similar to its diploid progenitor *Brachypodium stacei*, unlike the other progenitor *Brachypodium distachyon*, and this physiological trait can be attributed to the

preferential usage of homoeologs originating from the progenitor *B. stacei* to maintain cellular homeostasis under stress (Takahagi et al., 2018). Powell et al. (2017) demonstrated that allohexaploid wheat displayed HEB in response to pathogen infection and that such biases were mainly contributed by the B- and D-homoeologs. In a more recent study of *B. napus*, more A_T- than C_T-biases were found across control and three abiotic stress conditions (i.e., cold, drought, and heat) (Lee and Adams, 2020). Here, in tetraploid cottons, the salt-responsive HEB displayed an equivalent level of differentially altered At and Dt homoeologs (Figure 5; 52–54% versus 46–48%, respectively). In addition, the A- and D-biases were statistically balanced under both control and treatment conditions (Table 2). This is consistent with previous studies in other polyploids such as coffee (Combes et al., 2012; Combes et al., 2013) and wheat (Bhanbhro et al., 2020), where the contribution of different subgenomes appeared to be balanced during varied growing conditions. These results collectively imply that allopolyploids might respond to environmental stresses through differential or subgenome-specific usage of homoeologs, but that these responses will vary across species and conditions. We note that our study only investigated the cotton leaf transcriptional response to salinity stress at the seedling stage and that there is a clear need for further research using additional tissues and including different developmental stages.

Disentangling parental legacy from evolutionary responses in allopolyploid cottons

The analytical approach used in the present study was expressly designed to elucidate and distinguish evolutionary changes in stress responsiveness that arose subsequent to polyploidization from those that pre-existed in the diploid progenitors and thus were vertically inherited. This is important in that a leading idea about the adaptive relevance of polyploidy is that polyploids might be able to respond to changing or stressful conditions better than do their diploid relatives (Van de Peer et al., 2017; Van de Peer et al., 2020). To gain insight into this we jointly modeled stress responsiveness and ploidy level, with respect to a partitioned analysis of homoeologs. This multivariate HEB analysis identified more than 9000 homoeologous gene pairs that showed biased expression patterns in both polyploids, irrespective of stress treatment (Figure 4a). Notably, about half of these HEB patterns were inherited from the diploid parental state (type I reflecting 'parental legacy'), whereas the other half were novel in allopolyploids (type II reflecting 'regulatory novelty'). Perhaps the most important result, though, from the general standpoint of understanding the significance of polyploidy in evolution, is that among the stress-responsive HEB patterns (Figure 4b), over 90% were novel in allopolyploids (type IV). This sharp contrast between control and stress-responsive HEB

patterns suggests that allopolyploid-specific regulatory novelty likely plays a prominent role in salt stress responsiveness. We speculate that this reflects the numerous complex and novel interactions among both homoeologous regulatory factors and their target genes in allopolyploids, relative to their diploid progenitors, that arise immediately with polyploidization (Hu and Wendel, 2019). This regulatory robustness of gene regulatory networks conferred by polyploidy in adaptations to stresses has been testified in yeast through evolution experiments (Keane et al., 2014; Selmecki et al., 2015), as well as by recent computational simulation work (Yao et al., 2019). That is, deleterious mutations can often be buffered by the more complex polyploid gene networks or make the polyploids slightly less fit under stable environments, whereas the newly formed interactions or disturbances of the network structure may actually allow a wider range of evolvability of polyploids (reviewed in Carretero-Paulet and Van de Peer, 2020). In this light, we further note that many homoeologs displayed novel expression patterns under stress in polyploid cottons. Thus, it is reasonable to assume that the involved *cis*- and *trans*-regulatory elements act as interaction partners to coevolve the new phenotypic space that is specific to polyploids. It is worth mentioning that the diploids studied here are not the *actual* parents of polyploid cottons but are instead the best living models of diploid donors. Thus, the expression divergence between allopolyploids and their model diploids parents can also result from the independent evolutionary trajectories of diploid sister genomes and their corresponding subgenomes in allopolyploids. Nevertheless, this study still allows us to evaluate the role of homoeolog transcriptional regulation in salt responsiveness of allopolyploids. Further analyses carried out on artificial, synthetic allopolyploids systems would help shed light on diploids parental legacy versus polyploids regulatory novelty in responding to environmental stresses.

One additional result relevant to this polyploid-specific stress responsiveness becomes evident when one considers *total* homoeolog expression changes. As shown in Figure 1b, 10 690 and 8806 genes in *G. hirsutum* and in *G. mustelinum*, respectively, exhibited significant transcriptional alteration in response to modest salt stress. By examining *total* expression changes in allopolyploids relative to the parental diploids, we found that transgressively up- and downregulatory changes accounted for about one third of these changes (Figure 3). This large stress-responsive alteration in polyploids also likely reflects the massive combinatorial increase in regulatory complexity that is inherent in an allopolyploid relative to a diploid nucleus. Furthermore, these transgressive regulatory changes distinguished the two allopolyploid species studied, *G. hirsutum* and *G. mustelinum*, in terms of global level of transgressive downregulation. That is, fewer salt-responsive genes were

transgressively downregulated in the more salt-tolerant *G. mustelinum* than in the less tolerant *G. hirsutum*, which in turn attenuated the salt responses. Because the two allopolyploid species originate from a single allopolyploidy event (Grover et al., 2015), this divergence may lead to new understanding of the post-polyploidization evolution of diversity in adaptation to salinity stress.

In conclusion, our study revealed divergent responses to salt stress among four closely related diploid and polyploid cottons. We (i) show that the transcriptomic landscape associated with the salinity response reflects both parental legacy and regulatory novelty in polyploids and (ii) disentangle these effects through statistical categorization of four distinct HEB patterns. Our findings highlight the transcriptional regulatory consequences of duplicated genes in response to salt stress of polyploids and provide new insights into the putative adaptive role of allopolyploidization during evolution under stressful environments.

EXPERIMENTAL PROCEDURES

Plant materials and salinity stress treatment

Four cotton species were selected based on our previous study (Dong et al., 2020), including two natural allopolyploid cottons, *G. hirsutum* (AD_1) cv. TM1 and *G. mustelinum* (AD_4), and their model diploid progenitors *G. arboreum* (A_2) cv. 101 and *G. raimondii* (D_5). Mature seeds were delinted with sulfuric acid and germinated at 30°C on moistened filter paper placed in Petri dishes. Germinated seeds were next transferred to a hydroponic system containing half-strength Murashige and Skoog (MS) medium and grown in the Pohl Conservatory greenhouse at Iowa State University (Ames, IA). Uniformly developed 2-week-old seedlings were collected for each species and randomly divided into control and treatment groups, each with three biological replicates (three to four plants per replicate). Salt treatment of 50 mM NaCl for 2 weeks was applied as previously described (Dong et al., 2020).

RNA extraction, sequencing, and data processing

Fully expanded true leaves (seventh to ninth) were collected and pooled from three to four individual plants per replicate. Harvested tissues were flash-frozen in liquid nitrogen and stored at −80°C until extraction. Total RNA was isolated using the Qiagen Plant RNeasy Kit (Qiagen, Stanford, CA, USA) following the manufacturer's instructions. The concentration and integrity of extracted RNA were determined using an Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA). RNA samples were sent to the Beijing Genomics Institute (BGI, Hong Kong) for library preparation and 150-bp paired-end sequencing on a BGISEQ-500 platform. Raw reads were deposited at the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) under project number PRJNA601953.

RNA-seq reads were mapped against the *G. raimondii* reference genome (D_5; 37 505 protein-coding genes) (Paterson et al., 2012) using the single nucleotide polymorphism (SNP)-tolerant mode of the Genomic Short-Read Nucleotide Alignment Program (GSNAP) (Wu and Nacu, 2010). A set of pre-built, species-specific SNP indices (Page and Udall, 2015) were used to distinguish A- and D-genome orthologs and homoeologs. For allopolyploids, mapped reads were partitioned by polyCat (Page et al., 2013; Page et al., 2014) as homoeolog A-specific, D-specific, unclassified, or chimeric reads; because the latter two lack homoeologous SNPs or contain both A- and D-SNPs, respectively, reads in these categories were excluded from subsequent analyses as per (Hu et al., 2020). Read counts were generated using featureCounts from the subread package v.1.6.0 (Liao et al., 2014). The *total* expression of homoeologous genes in the allopolyploids was calculated as the sum of the homoeolog A- and D-specific expressions.

Differential gene expression analysis

Raw read counts were used for differential expression analysis with the DESeq2 package (Love et al., 2014) in R v. 3.6.3 (R Core Team, 2017). First, a multifactor design (*species* + *treatment* + *species*:*treatment*) was used to differentiate gene expression changes due to species, salt treatment, or the interaction between the two, using a likelihood ratio test-based approach to compare the full model against the corresponding reduced models. Next, pairwise comparisons were conducted with Wald tests to determine DEG sets between control and salt-treated conditions, between any two genotypes, and due to species-specific responses to salinity (interaction). The resulting *P*-values were corrected using a false discovery rate (FDR) of $\alpha = 0.05$ by the Benjamini–Hochberg method (Benjamini and Hochberg, 1995). The numbers of DEGs were divided by the total number of genes in the reference genome (i.e., 37 505) to obtain DEG percentages for direct comparisons (Rapp et al., 2009; Bardil et al., 2011). Clustering analysis of DEGs with interaction effect was performed by the Complexheatmap package (Gu et al., 2016) with the Spearman distance matrix of *z*-score scaled $\log_2$(fold change) values.

Homoeolog expression bias

HEB refers to differential expression between homoeologous genes in allopolyploid species (Grover et al., 2012); in tetraploid cotton, an A-bias indicates higher expression of the At than the Dt homoeolog, and a D-bias indicates higher Dt than At expression. In addition to this classic definition of HEB under static conditions (e.g., in either treated or control plants), we are also interested in the conditional HEB in response to stress treatment. Therefore, we applied a multifactor design to each allopolyploid species relative to the parental diploids: expression *homoeolog* + *ploidy* + *treatment* + *homoeolog*:*ploidy* + *homoeolog*:*treatment* + *ploidy*:*treatment* + *homoeolog*:*ploidy*:*treatment*, where (sub)genome-specific gene expression is subject to the effects of genome origin (*homoeolog*: A_2/At versus D_5/Dt), ploidy level (*ploidy*: allopolyploid versus diploid), and experimental condition (*treatment*: control versus treated). Based on the significant effects of *homoeolog*-related terms, four HEB patterns were inferred as follows: *homoeolog*, type I HEB that is inherited from the parental state and not affected by treatment; *homoeolog*:*ploidy*, type II HEB that is novel in allopolyploids and not affected by treatment; *homoeolog*:*treatment*, type III stress-responsive HEB that is inherited from the parental state; and *homoeolog*:*ploidy*:*treatment*, type IV stress-responsive HEB that is novel in allopolyploids. An FDR-adjusted *P*-value of 5% was used to identify significant effects.

Functional enrichment analysis

GO enrichment analysis was performed using the R package topGO (Alexa and Rahnenführer, 2016). Briefly, Fisher's exact test was applied to the gene sets of interest to detect the

overrepresented GO terms against the universal set of genes. The *P*-values were adjusted for multiple testing using the Benjamini–Hochberg method (Benjamini and Hochberg, 1995) to retain significant GO terms with an adjusted *P* < 0.05. GO annotations of the *G. raimondii* reference genes were downloaded from CottonGen (Paterson et al., 2012; Yu et al., 2014). GO enrichment results were summarized using the REVIGO web server (Supek et al., 2011).

Weighted gene coexpression network analysis

Coexpression network analysis was conducted using the R package WGCNA (Langfelder and Horvath, 2008). The raw read count data were normalized using the regularized log transformation method built in DESeq2 (Love et al., 2014). After removing genes with zero expression or without variation across samples, coexpression modules were obtained using the automatic network construction function *blockwiseModules*. Briefly, a similarity matrix was calculated based on Pearson correlation indices between each pair of genes, which was then applied to generate a 'signed' adjacency matrix with an optimized soft threshold of $\beta = 26$. Network interconnectedness was next measured by calculating the 'signed' topological overlaps using the *TOMdist* function. By performing average linkage hierarchical clustering on the topological overlap dissimilarity matrix (DistTOM = 1 − TOM), genes that were highly connected were grouped together into coexpression modules with a dynamic tree cutting algorithm. The parameter *mergeCutHeight* = 0.25 was used to merge highly similar modules.

For each coexpression module, a module eigengene that characterizes the representative gene expression pattern was calculated as the first principal component of the scaled module member gene expression profiles. Using the 29 phenotypic and physiological traits data obtained from our previous study (Dong et al., 2020) (the same accessions as those used in the present study), module–trait relationships were examined by Pearson correlation indices between module eigengenes and trait measurements using the *cor* function in WGCNA. The resulting *P*-values were corrected for multiple testing by the Benjamini–Hochberg method (Benjamini and Hochberg, 1995).

AUTHOR CONTRIBUTIONS

GH, YD, and JFW conceived the research; YD and ERM carried out the experiments; YD and CEG performed data analysis; YD wrote the manuscript; GH, JFW, SZ, and CEG contributed substantially to revisions.

ACKNOWLEDGMENTS

CONFLICT OF INTERESTS

The authors have declared that no competing interests exist.

DATA AVAILABILITY STATEMENT

RNA-seq data were deposited under NCBI Bioproject PRJNA601953; custom R scripts are available at https://github.com/Wendellab/SaltStressTranscriptome.

SUPPORTING INFORMATION

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

Figure S1. (a) Correlation heatmap of RNA-seq samples. CK and Salt refer to control and salt treatment samples, respectively; numbers 1, 2, and 3 represent three biological replicates. (b) Principal component analysis of A- and D-subgenomic expression in allopolyploid cottons (*G. hirsutum* AD_1 and *G. mustelinum* AD_4) and their diploid progenitors *G. arboreum* (A_2) and *G. raimondii* (D_5). Diploid samples are circled in black, and the one mislabeled D_5 sample is circled in red.

Figure S2. GO enrichment of differentially expressed genes with significant *species*, *treatment*, and *species:treatment* effects. Enriched GO terms of biological processes were summarized and visualized by REVIGO, which uses multidimensional scaling (MDS) to reduce the dimensionality of a matrix of the GO terms pairwise semantic similarities. Circle color indicates the $\log_{10}$ adjusted *P*-value and circle size indicates the GO term frequency.

Figure S3. GO enrichment of up- and downregulated genes in response to salt stress of each species.

Figure S4. GO enrichment of species-specific up- and downregulated genes in response to salt stress, corresponding to the first four columns in Figure 1c,d. Results are shown for A_2, D_5, and AD_4 genes, because no significant GO terms were enriched in AD_1-specific differentially expressed genes.

Figure S5. WGCNA network construction. (a) Selection of the soft-thresholding powers. The scale-free fit index and the mean connectivity on the *y*-axis are plotted against soft-thresholding power on the *x*-axis. (b) Coexpression modules were identified by cutting the hierarchical clustering tree. (c) Eigengene dendrogram and correlation heatmap showing the relationships among identified modules. (d) Correlation heatmap between module eigengenes and traits. In each cell, color and text indicate correlation and *P*-value.

Figure S6. GO enrichment of module member genes.

Table S1. Summary of RNA-seq data generated in this study.

Table S2. GO enrichment analysis of differentially expressed genes exhibiting species, treatment, and interaction effects.

Table S3. GO enrichment analysis of differentially expressed genes in response to salt stress in four cotton species.

Table S4. GO enrichment analysis of species-specific salt stress-responsive genes.

Table S5. GO enrichment analysis of WGCNA coexpression modules.

Table S6. GO enrichment analysis of genes from each cluster.

Table S7. GO enrichment analysis of genes showing responsiveness HEB patterns in polyploids.

OPEN RESEARCH BADGES

REFERENCES

Abdelraheem, A., Esmaeili, N., O'Connell, M. & Zhang, J. (2019) Progress and perspective on drought and salt stress tolerance in cotton. *Industrial Crops and Products*, **130**, 118–129.

Adams, K.L. & Wendel, J.F. (2005) Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology*, **8**, 135–141.

Alexa, A. & Rahnenfuhrer, J. (2016) Gene set enrichment analysis with topGO. *Bioconductor*, **27**, 1–26.

Bardil, A., de Almeida, J.D., Combes, M.C., Lashermes, P. & Bertrand, B. (2011) Genomic expression dominance in the natural allopolyploid *Coffea arabica* is massively affected by growth temperature. *The 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 (Statistical Methodology)*, **57**, 289–300.

Bhanbhro, N., Xiao, B., Han, L., Lu, H., Wang, H. & Yang, C. (2020) Adaptive strategy of allohexaploid wheat to long-term salinity stress. *BMC Plant Biology*, **20**, 210.

Blanc, G. & Wolfe, K.H. (2004) Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell*, **16**, 1667–1678.

Bowman, J.L., Floyd, S.K. & Sakakibara, K. (2007) Green genes–comparative genomics of the green branch of life. *Cell*, **129**, 229–234.

Buggs, R.J.A., Wendel, J.F., Doyle, J.J., Soltis, D.E., Soltis, P.S. & Coate, J.E. (2014) The legacy of diploid progenitors in allopolyploid gene expression patterns. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **369**, 20130354.

Carretero-Paulet, L. & Van de Peer, Y. (2020) The evolutionary conundrum of whole-genome duplication. *American Journal of Botany*, **107**, 1101–1105.

Chelaifa, H., Chagué, V., Chalabi, S., Mestiri, I., Arnaud, D., Deffains, D. et al. (2013) Prevalence of gene expression additivity in genetically stable wheat allohexaploids. *The New Phytologist*, **197**, 730–736.

Chen, Z.J. (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annual Review of Plant Biology*, **58**, 377–406.

Chen, Z.J. (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends in Plant Science*, **15**, 57–71.

Chen, Z.J. & Birchler, J.A. (2013) *Polyploid and hybrid genomics*. New York, NY: John Wiley & Sons.

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

Combes, M.-C., Cenci, A., Baraille, H., Bertrand, B. & Lashermes, P. (2012) Homeologous gene expression in response to growing temperature in a recent allopolyploid (*Coffea arabica* L.). *The Journal of Heredity*, **103**, 36–46.

Combes, M.-C., Dereeper, A., Severac, D., Bertrand, B. & Lashermes, P. (2013) Contribution of subgenomes to the transcriptome and their intertwined regulation in the allopolyploid *Coffea arabica* grown at contrasted temperatures. *The New Phytologist*, **200**, 251–260.

Deinlein, U., Stephan, A.B., Horie, T., Luo, W., Xu, G. & Schroeder, J.I. (2014) Plant salt-tolerance mechanisms. *Trends in Plant Science*, **19**, 371–379.

Ding, M. & Chen, Z.J. (2018) Epigenetic perspectives on the evolution and domestication of polyploid plant and crops. *Current Opinion in Plant Biology*, **42**, 37–48.

Dong, S. & Adams, K.L. (2011) Differential contributions to the transcriptome of duplicated genes in response to abiotic stresses in natural and synthetic polyploids. *The New Phytologist*, **190**, 1045–1057.

Dong, Y., Hu, G., Yu, J., Thu, S.W., Grover, C.E., Zhu, S. et al. (2020) Salt-tolerance diversity in diploid and polyploid cotton (*Gossypium*) species. *The Plant Journal*, **101**, 1135–1151.

Doyle, J.J., Flagel, L.E. & Paterson, A.H. (2008) Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics*, **42**, 443–461.

Doyle, J.J. & Sherman-Broyles, S. (2017) Double trouble: taxonomy and definitions of polyploidy. *The New Phytologist*, **213**, 487–493.

Edelist, C., Raffoux, X., Falque, M., Dillmann, C., Sicard, D., Rieseberg, L.H. et al. (2009) Differential expression of candidate salt-tolerance genes in the halophyte *Helianthus paradoxus* and its glycophyte progenitors *H. annuus* and *H. petiolaris* (Asteraceae). *American Journal of Botany*, **96**, 1830–1838.

Endrizzi, J.E., Turcotte, E.L. & Kohel, R.J. (1985) Genetics, Cytology, and Evolution of *Gossypium*. In: Caspari, E.W. & Scandalios, J.G. (Eds.) *Advances in genetics*. New York, NY: Academic Press, pp. 271–375.

Fawcett, J.A. & Maere, S. (2009) Plants with double genomes might have had a better chance to survive the cretaceous–tertiary extinction event. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 5737–5742.

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

Flowers, T.J. & Colmer, T.D. (2008) Salinity tolerance in halophytes. *The New Phytologist*, **179**, 945–963.

Fox, D.T., Soltis, D.E., Soltis, P.S., Ashman, T.-L. & Van de Peer, Y. (2020) Polyploidy: a biological force from cells to ecosystems. *Trends in Cell Biology*, **30**, 688–694.

Grover, C.E., Gallagher, J.P., Jareczek, J.J., Page, J.T., Udall, J.A., Gore, M.A. et al. (2015) Re-evaluating the phylogeny of allopolyploid *Gossypium* L. *Molecular Phylogenetics and Evolution*, **92**, 45–52.

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

Gu, Z., Eils, R. & Schlesner, M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, **32**, 2847–2849.

Gugger, P.F., Peñaloza-Ramírez, J.M., Wright, J.W. & Sork, V.L. (2017) Whole-transcriptome response to water stress in a California endemic oak, *Quercus lobata*. *Tree Physiology*, **37**, 632–644.

Guo, J., Shi, G., Guo, X., Zhang, L., Xu, W., Wang, Y. et al. (2015) Transcriptome analysis reveals that distinct metabolic pathways operate in salt-tolerant and salt-sensitive upland cotton varieties subjected to salinity stress. *Plant Science*, **238**, 33–45.

Halfter, U., Ishitani, M. & Zhu, J.K. (2000) The Arabidopsis SOS2 protein kinase physically interacts with and is activated by the calcium-binding protein SOS3. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 3735–3740.

Hu, G., Grover, C.E., Arick, M.A., Liu, M., Peterson, D.G. & Wendel, J.F. (2020) Homoeologous gene expression and co-expression network analyses and evolutionary inference in allopolyploids. *Briefings in Bioinformatics*, **22**, 1819–1835.

Hu, G., Grover, C.E., Yuan, D., Dong, Y., Miller, E., Conover, J.L. et al. (2021) Evolution and diversity of the cotton genome. In: Rahman, M., Zafar, Y. & Zhang, T. (Eds.) *Cotton precision breeding*. Switzerland AG: Springer Nature, pp. 25–78.

Hu, G., Koh, J., Yoo, M.-J., Grupp, K., Chen, S. & Wendel, J.F. (2013) Proteomic profiling of developing cotton fibers from wild and domesticated *Gossypium barbadense*. *The New Phytologist*, **200**, 570–582.

Hu, G. & Wendel, J.F. (2019) *Cis–trans* controls and regulatory novelty accompanying allopolyploidization. *The New Phytologist*, **221**, 1691–1700.

Jackson, S. & Chen, Z.J. (2010) Genomic and expression plasticity of polyploidy. *Current Opinion in Plant Biology*, **13**, 153–159.

Ji, H., Pardo, J.M., Batelli, G., Van Oosten, M.J., Bressan, R.A. & Li, X. (2013) The salt overly sensitive (SOS) pathway: established and emerging roles. *Molecular Plant*, **6**, 275–286.

Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E. et al. (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature*, **473**, 97–100.

Kazachkova, Y., Batushansky, A., Cisneros, A., Tel-Zur, N., Fait, A. & Barak, S. (2013) Growth platform-dependent and -independent phenotypic and metabolic responses of Arabidopsis and its halophytic relative, *Eutrema salsugineum*, to salt stress. *Plant Physiology*, **162**, 1583–1598.

Keane, O.M., Toft, C., Carretero-Paulet, L., Jones, G.W. & Fares, M.A. (2014) Preservation of genetic and regulatory robustness in ancient gene duplicates of *Saccharomyces cerevisiae*. *Genome Research*, **24**, 1830–1841.

Langfelder, P. & Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 559.

Lee, J.S. & Adams, K.L. (2020) Global insights into duplicated gene expression and alternative splicing in polyploid *Brassica napus* under heat, cold, and drought stress. *Plant Genome*, **13**, e20057.

Lei, Y., Xu, Y., Hettenhausen, C., Lu, C., Shen, G., Zhang, C. *et al.* (2018) Comparative analysis of alfalfa (*Medicago sativa* L.) leaf transcriptomes reveals genotype-specific salt tolerance mechanisms. *BMC Plant Biology*, **18**, 35.

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

Li, M., Wang, R., Wu, X. & Wang, J. (2020) Homoeolog expression bias and expression level dominance (ELD) in four tissues of natural allotetraploid *Brassica napus. BMC Genomics*, **21**, 330.

Liao, Y., Smyth, G.K. & Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.

Liu, Y., Ma, J., Yan, W., Yan, G. & Zhou, M. (2012) Different tolerance in bread wheat, durum wheat and barley to *fusarium* crown rot disease caused by *fusarium pseudograminearum. Journal of Phytopathology*, **160**, 412–417.

Liu, Z. & Adams, K.L. (2007) Expression partitioning between genes duplicated by polyploidy under abiotic stress and during organ development. *Current Biology*, **17**, 1669–1674.

Liu, Z., Xin, M., Qin, J., Peng, H., Ni, Z., Yao, Y. *et al.* (2015) Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (*Triticum aestivum* L.). *BMC Plant Biology*, **15**, 152.

Love, M.I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.

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

Manishankar, P., Wang, N., Köster, P., Alatar, A.A. & Kudla, J. (2018) Calcium signaling during salt stress and in the regulation of ion homeostasis. *Journal of Experimental Botany*, **69**, 4215–4226.

Munns, R. & Tester, M. (2008) Mechanisms of salinity tolerance. *Annual Review of Plant Biology*, **59**, 651–681.

Ni, Z., Kim, E.-D., Ha, M., Lackey, E., Liu, J., Zhang, Y. *et al.* (2009) Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature*, **457**, 327–331.

Page, J.T., Gingle, A.R. & Udall, J.A. (2013) PolyCat: a resource for genome categorization of sequencing reads from allopolyploid organisms. *G3: Genes, Genomes, Genetics*, **3**, 517–525.

Page, J.T., Liechty, Z.S., Huynh, M.D. & Udall, J.A. (2014) BamBam: genome sequence analysis tools for biologists. *BMC Research Notes*, **7**, 829.

Page, J.T. & Udall, J.A. (2015) Methods for mapping and categorization of DNA sequence reads from allopolyploid organisms. *BMC Genetics*, **16** (Suppl 2), S4.

Paterson, A.H., Wendel, J.F., Gundlach, H., Guo, H., Jenkins, J., Jin, D. *et al.* (2012) Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. *Nature*, **492**, 423–427.

Peng, Z., Cheng, H., Sun, G., Pan, Z. & Wang, X. (2020) Expression patterns and functional divergence of homologous genes accompanied by polyploidization in cotton (*Gossypium hirsutum* L.). *Science China. Life Sciences*, **63**, 1565–1579.

Peng, Z., He, S., Gong, W., Sun, J., Pan, Z., Xu, F. *et al.* (2014) Comprehensive analysis of differentially expressed genes and transcriptional regulation induced by salt stress in two contrasting cotton genotypes. *BMC Genomics*, **15**, 760.

Powell, A.F. & Doyle, J.J. (2017) Non-additive transcriptomic responses to inoculation with rhizobia in a young allopolyploid compared with its diploid progenitors. *Genes*, **8**, 357.

Powell, J.J., Fitzgerald, T.L., Stiller, J., Berkman, P.J., Gardiner, D.M., Manners, J.M. *et al.* (2017) The defence-associated transcriptome of hexaploid wheat displays homoeolog expression and induction bias. *Plant Biotechnology Journal*, **15**, 533–543.

Pumphrey, M., Bai, J. & Laudencia-Chingcuanco, D. (2009) Nonadditive expression of homoeologous genes is established upon polyploidization in hexaploid wheat. *Genetics*, **181**, 1147–1157.

R Core Team. (2017, 2016) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

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

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

Salman-Minkov, A., Sabath, N. & Mayrose, I. (2016) Whole-genome duplication as a key factor in crop domestication. *Nature Plants*, **2**, 16115.

Selmecki, A.M., Maruvka, Y.E., Richmond, P.A., Guillet, M., Shoresh, N., Sorenson, A.L. *et al.* (2015) Polyploidy can drive rapid adaptation in yeast. *Nature*, **519**, 349–352.

Shimizu-Inatsugi, R., Terada, A., Hirose, K., Kudoh, H., Sese, J. & Shimizu, K.K. (2017) Plant adaptive radiation mediated by polyploid plasticity in transcriptomes. *Molecular Ecology*, **26**, 193–207.

Sigel, E.M., Der, J.P., Windham, M.D. & Pryer, K.M. (2019) Expression level dominance and homeolog expression bias in recurrent origins of the allopolyploid fern *Polypodium hesperium. American Fern Journal*, **109**, 224.

Solhaug, E.M., Ihinger, J., Jost, M., Gamboa, V., Marchant, B., Bradford, D. *et al.* (2016) Environmental regulation of heterosis in the allopolyploid *Arabidopsis suecica. Plant Physiology*, **170**, 2251–2263.

Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*, **6**, e21800.

Takahagi, K., Inoue, K. & Shimizu, M. (2018) Homoeolog-specific activation of genes for heat acclimation in the allopolyploid grass *Brachypodium hybridum. Gigascience*, **7**, giy020.

Udall, J.A. & Wendel, J.F. (2006) Polyploidy and crop improvement. *Crop Science*, **46**, S-3–S-14.

Van de Peer, Y., Ashman, T.-L., Soltis, P.S. & Soltis, D.E. (2020) Polyploidy: an evolutionary and ecological force in stressful times. *The Plant Cell*, **33**, 11–26.

Van de Peer, Y., Mizrachi, E. & Marchal, K. (2017) The evolutionary significance of polyploidy. *Nature Reviews. Genetics*, **18**, 411–424.

van Zelm, E., Zhang, Y. & Testerink, C. (2020) Salt tolerance mechanisms of plants. *Annual Review of Plant Biology*, **71**, 403–433.

Wang, K., Wendel, J.F. & Jinping, H.U.A. (2018) Designations for individual genomes and chromosomes in *Gossypium. Journal of Cotton Research*, **1**, 3.

Wang, L., Cao, S., Wang, P., Lu, K., Song, Q., Zhao, F.-J. *et al.* (2021) DNA hypomethylation in tetraploid rice potentiates stress-responsive gene expression for salt tolerance. *Proceedings of the National Academy of Sciences of the United States of America*, **118**, e2023981118.

Wendel, J.F. (2015) The wondrous cycles of polyploidy in plants. *American Journal of Botany*, **102**, 1753–1756.

Wendel, J.F., Brubaker, C.L. & Seelanan, T. (2010) The origin and evolution of *Gossypium*. In: Stewart, J.M., Oosterhuis, D.M., Heitholt, J.J. & Mauney, J.R. (Eds.) *Physiology of cotton*. Switzerland: Springer Nature, pp. 1–18.

Wendel, J.F. & Cronn, R.C. (2003) Polyploidy and the evolutionary history of cotton. *Advances in Agronomy*, **78**, 139.

Wendel, J.F. & Grover, C.E. (2015) Taxonomy and evolution of the cotton genus, *Gossypium*. In: Fang, D.D. & Percy, R.G. (Eds.) *Cotton*, 2nd edition. Madison, WI: American Society of Agronomy Inc., pp. 25–44.

Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B. & Rieseberg, L.H. (2009) The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 13875–13879.

Wu, J., Lin, L., Xu, M., Chen, P., Liu, D., Sun, Q. *et al.* (2018) Homoeolog expression bias and expression level dominance in resynthesized allopolyploid *Brassica napus. BMC Genomics*, **19**, 586.

Wu, T.D. & Nacu, S. (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*, **26**, 873–881.

Yang, C., Zhao, L., Zhang, H., Yang, Z., Wang, H., Wen, S. *et al.* (2014) Evolution of physiological responses to salt stress in hexaploid wheat. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 11882–11887.

Yao, Y., Carretero-Paulet, L. & Van de Peer, Y. (2019) Using digital organisms to study the evolutionary consequences of whole genome duplication and polyploidy. *PLoS One*, **14**, e0220257.

Yoo, M.-J., Liu, X., Pires, J.C., Soltis, P.S. & Soltis, D.E. (2014) Nonadditive gene expression in polyploids. *Annual Review of Genetics*, **48**, 485–517.

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

Yu, J., Jung, S., Cheng, C.-H., Ficklin, S.P., Lee, T., Zheng, P. *et al.* (2014) CottonGen: a genomics, genetics and breeding database for cotton research. *Nucleic Acids Research*, **42**, D1229–D1236.

Yu, X., Wang, X., Hyldgaard, B., Zhu, Z., Zhou, R., Kjaer, K.H. *et al.* (2018) Allopolyploidization in *Cucumis* contributes to delayed leaf maturation with repression of redundant homoeologous genes. *The Plant Journal*, **94**, 393–404.

Yuan, Y., Xing, H., Zeng, W., Xu, J., Mao, L., Wang, L. *et al.* (2019) Genome-wide association and differential expression analysis of salt tolerance in *Gossypium hirsutum* L at the germination stage. *BMC Plant Biology*, **19**, 394.

Zhang, Y., Liu, Z., Khan, A.A., Lin, Q., Han, Y., Mu, P. *et al.* (2016) Expression partitioning of homeologs and tandem duplications contribute to salt tolerance in wheat (*Triticum aestivum* L.). *Scientific Reports*, **6**, 21476.