

Homoeolog gene expression analysis reveals novel expression biases in upland hybrid cotton under intraspecific hybridization

Kashif Shahzad (✉ kashifshahzad85@yahoo.com)

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute

Xuexian Zhang

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute

Meng Zhang

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute

Liping Guo

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute

Tingxiang Qi

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute

Huini Tang

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute

Hailin Wang

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute

Iqra Mubeen

University of Agriculture Faisalabad

Xiuqin Qiao

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute

Jianyong Wu

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute

Chaozhu Xing

CAAS CRI: Chinese Academy of Agricultural Sciences Cotton Research Institute <https://orcid.org/0000-0001-5099-6094>

Research Article

Keywords: Cotton, Heterosis, Homoeolog expression, RNA-Seq, Secondary metabolism

Posted Date: February 11th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1325696/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Hybridization is useful to enhance yield potential of agronomic crops in the world. Cotton has genome doubling due to allotetraploidy process and hybridization process in coordinate with duplicated genome can produce more yield and adaptability. Therefore, expression of homoeolog gene pairs between hybrids and inbred parents are vital to characterize genetic source of heterosis in cotton. Investigation results of homoeolog gene pairs between two contrasting hybrids and their respective inbred parents identified 36853 homoeolog genes in hybrids. It was observed both high and low hybrids had similar trends in homoeolog gene expression patterns in each tissue under study. An average of 96% of homoeolog genes had no biased expression and their expressions were derived from the equal contribution of both parents. Besides, very few homoeolog genes (An average of 1%) showed no biased or novel expression in both hybrids. The functional analysis described secondary metabolic pathways had a majority of novel biased homoeolog genes in hybrids. These results contribute preliminary knowledge about how hybridization affects expression patterns of homoeolog gene pairs in upland cotton hybrids. Our study also highlights the functional genomics of metabolic genes to explore the genetic mechanism of heterosis in cotton.

Background

Hybridization or polyploidy is a vital process, produces exceptional phenotypes through the interaction of two distinct genomes, and has a substantial influence on the plant genome (1, 2). The novel phenotypes in plants can be consequent of many modifications not only at genomic and epigenetic levels but also in gene expression patterns (3, 4). Several examples in cotton (5) *Brassica napus* (6), and *Spartina* (7) stated that hybridization (genome merger) modified more transcriptome than polyploidy. About 15% of flowering plants are polyploidy (8) and divided into categories of autopolyploid and allopolyploid. Autopolyploid is evolved from a genome merger of the same or similar species but allopolyploids contain a genomic combination of two different species (3). Allopolyploids have enhanced growth and wider adaptations relative to autopolyploid. Many important crops are allopolyploids together with upland cotton (*Gossypium hirsutum*). In allopolyploids, hybridization or polyploidy is allied with various kinds of exchanges and interactions between homoeolog genes (duplicate genes) (9–11). Further modifications are consisting of epigenetic (3, 12), gene activation (4, 13), and sequence removal (14, 15). The result of different genomic interactions by hybridization or polyploidy produced several expression patterns of duplicate or genes with similar functions. For instance, these may be gene silencing, gene activation, novel expression, and altered parental contribution (2, 16). However, the direction and magnitude of duplicate gene expression are varied with crop, tissue, and organ (17–20).

Previous studies have identified four different expression forms of duplicated genes in plants e.g., one duplicate gene copy may be lost its expression pattern as compared to other (21), duplicate gene copies may maintain original expression (22), one duplicate gene copy may adopt new expression (23, 24), and sometimes original expression is the equal contribution of duplicates (25). Importantly, the duplicate gene pairs may show homoeolog expression bias in allopolyploid (26). Where, one homoeolog display preferential expression to the transcriptome as compared to the other. The phenomena of homoeolog

expression have already been stated for allopolyploids like cotton (27–29), wheat (18, 30), Arabidopsis (31, 32), and brassica (33, 34). The analysis results in cotton showed that hybridization caused more homoeolog expression changes and fewer changes in gene expressions (2). In contrast, polyploidy formation is linked with more gene expression modifications and fewer changes in homoeolog expression. A study in cyprinid fish also reported hybridization brought more homoeolog expression changes than polyploidy (35). Homoeolog expression can be either balanced or unbalanced (26). When duplicate gene pairs display biased expression towards one parental genome is almost equal to the biased expression towards other parental genomes. It shows balanced homoeolog expression. On the other hand, an unbalanced homoeolog expression change states to a situation where more duplicate gene pairs demonstrate expression twisted toward one parental genome. An unbalanced form of homoeolog expression is mostly reported in allopolyploids (36–38).

Upland cotton is a prime fiber, an important source of edible oil, and cash crops in the world (39). The demand for cotton fiber in the textile industry is increasing. While supply is reducing due to low area of cultivation and staggered yield potential of cultivars. Heterosis is and will continue to be a basic breeding tool to produce superior genotypes (40). Many researchers reported meaningful heterosis in seed cotton yield and fiber quality traits in upland cotton (41, 42). Further, cotton genotypes produced through the utilization of heterosis exhibit more stability and adaptability to changing the environment (43). The genetic mechanism of heterosis is a perturbed and so-called genetic model of dominance (44), over dominance (45), and epistasis (46) does not reveal it. However, progress has already been achieved in the recent few years with advanced molecular research. For instance, high-throughput transcriptome sequencing is a useful technology to get accurate knowledge about global gene expressions (47, 48). This also creates a possibility to discriminate the expression of homologous genes (49). To date, many previous studies in agronomic crops analyzed patterns of homoeolog expression in allotetraploids and their diploid progenitors e.g., cotton (2) and brassica (50). However, information regarding how hybridization effect homoeolog expression changes in newly formed upland cotton hybrids is limited. Thus, we conducted a study to investigate the effect of hybridization on homoeolog expression modifications on a genome-wide scale. The direction and extent of homoeologue expressions between contrasting F_1 hybrids and their reciprocal parents were analyzed by RNA-seq of leaf, root, and flower buds. The result will provide new insights into how homoeologue expressions contribute to growth heterosis in upland cotton and enhance our genetic knowledge for hybridization process.

Material And Methods

Plant material and RNA sequencing

The plant material used in this study included two seed cotton yield contrasting F_1 hybrids of upland cotton and their three reciprocal inbred parents named as A, B, and D. The high yielding hybrid (H) was cross of inbred A and B. whereas, a low hybrid was a combination of cross A and D. The seedling was raised in the lab for all these materials, root and leaf samples were picked at true leaf stage, and flower

buds were picked at squaring stage from the field in three biological repeats. Total RNA was isolated using the Sigma Spectrum Plant Total RNA kit (Sigma-Aldrich, USA) as described by the manufacturing company. After quality checks and preparations, paired-end sequencing (300 ± 50 bp) on a second-generation Illumina Hiseq 4000 was performed for hybrid H and L. These sequencing details can be seen at NCBI with accession number GSE144676 (45) and GSE150052 (47). The full-length Pacbio transcriptome sequencing (Unpublished) for parental A, B, and D were used for a better comparison of homoeolog expression patterns between parents and progeny.

Analysis of homoeolog expression changes and functional enrichment

To find homologous sequences between reciprocal parents of hybrid H and L, the full-length Pacbio sequences of the two parents were homologously aligned using the reciprocal BLAST (BLASTN) hit method with screening index of e^{-20} (35). Two sequences were homologous, if each of them was the best hit of the other, and if the sequences were aligned over 300 bp. The total number of homologous gene pairs identified in parent A and B were 45216. While the parents A and D had 44920 homologous gene pairs. These homologous sequences were used as a reference for the already published second-generation Illumina sequencing H and L hybrid samples (45, 47). Each sample was subjected to mutation detection and specific SNPs. The hybrids offspring samples may be homozygous or heterozygous at the specific SNP sites of their respective parents. Later, we compared the SNPs of two parents to select the homoeologue-specific SNPs bins using custom python scripts (50). Further, genes with homoeologue-specific SNPs were compared on a case by case basis between the parents and then progeny to perform synchronous classification of homoeolog expression bias. The index to determine the significant difference was *p-value* < 0.05 . Compared with the diploid progenitors, the homologous gene pairs between both parents of hybrids were divided into 12 main categories with three subcategories, including the parental condition, no bias, and novel bias in progeny as already described (2). Functional analyses including gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) was performed for all homoeolog expression genes that showed novel bias in different tissues of hybrids. GO terms and enriched pathways were retrieved from the cotton functional genomics database (<https://cottonfgd.org/analyz>) with *p-value* < 0.05 .

Results

Homoeolog expression changes in different tissues of high hybrid

To observe the pattern of homoeolog expression in different tissues of high hybrid, a case by case gene comparison was performed for high hybrid (H) and its parents (maternal=A, paternal=B). The results are detailed in Table 1. Notably, a pattern of homoeolog expression was conserved in this hybrid for the

greater number of genes in all tissues under investigation. Fewer homoeolog gene copies reflected novel biased pattern. For example, 36853 total number homoeolog gene pairs were identified in each tissue. We observed that 96% in the root, 97% in leaf, and 97% in flower buds had maintained parental expression patterns in hybrid. The total expression level of a homologous gene pair had no change or equal to that of both parents. An average of 1% homoeolog gene pairs showed no bias in each hybrid tissue. According to the results of novel bias expression, 2.27% homologous gene pairs had novel bias in the root. Only 2.1% homologous gene pairs in leaf exhibited novel bias expression. It was observed that 2.02% homologous gene pairs had novel bias in flower buds. Majority of novel biased homologous gene pairs exhibited maternal like expression patterns in root and leaf (Figure 1a-b). Wherein, parental like novel expression were observed in flower buds (Figure 1c). Further analysis indicated that 1.2% homoeolog gene pairs in the root, 1.06% in leaf, and 1.1% in flower buds demonstrated overall maternal parent biased expression. Whereas, overall homoeolog gene pairs with paternal parent biased expression was 1.1%, 1.04%, and 0.95% in the root, leaf, and flower buds, respectively. The overall statistics of homoeolog expression changes in different tissues of high hybrid seem to show a highly balanced biased expression of homoeolog gene pairs in each tissue. These results suggested homoeolog gene pairs that displayed fewer novel patterns of homoeolog expression may be responsible for the high level of heterosis in upland cotton.

Homoeolog expression changes in different tissues of low hybrid

The results of homoeolog expression analysis in different tissues of low hybrid (L) and its reciprocal parents (A= maternal, B= paternal) are represented in Table 2. Interestingly, the homoeolog expression patterns in hybrid were in the same direction for the majority of homoeolog gene pairs as observed in parents. For instance, 97.3% in the root, 97.2% in leaf, and 97.6% in flower buds out of all homoeolog gene pairs (36853) reflected equal parental expression patterns in this hybrid. These result predicted that expression biases pre-existing in parents was simply conserved in these tissues of hybrid. The group of homoeolog gene pairs with bias expression in parents and reverted to no bias expression in hybrid had an average of 1% in root and leaf. On the other hand, only 0.60% of homologous gene pairs in flower buds had this reversion. According to the results of novel bias expression in hybrid, 1.75, 1.78, and 1.76% homologous gene pairs adopted novel bias in the root, leaf, and flower buds, respectively. Majority of homologous gene pairs exhibited novel biased expression similar to maternal parents in root and leaf (Figure 2a-b). In contrast, flower buds had most of homologous gene pairs that had shown novel biased expression with higher paternal contribution in progeny (Figure 2c). Overall, a group of maternal or paternal biased homologous gene pairs had very similar trends in each tissue. There was 0.83% maternal and 0.84% paternal biased homologous gene expression in the root, whereas leaf had 0.98 and 0.81% maternal and paternal biased expression, respectively. Flower buds also showed a similar percentage of maternal and parental biased expression. Altogether, these results suggest that the majority of the homoeolog gene exhibited balanced or non-differential homoeolog expression patterns in low hybrid.

Further, homologous genes with few novel patterns of homoeolog expression probably contribute to the low level of heterosis in upland cotton.

Distribution of novel expressed genes among different hybrids

This study further investigated the unique and common novel expressed genes among each tissue of hybrids and among different hybrids. In high hybrid, there were 374, 325, and 301 unique expressed novel genes in root, leaf, and flower buds, respectively (Figure 3). Moreover, a major portion of genes were overlapped among comparison of these tissues. For example, 94 genes were overlapped between comparison of root and leaf whereas 100 genes were overlapped between root and flower buds. The distribution results had shown that 298 novel genes were unique in root of low hybrid (Figure 4). The unique novel genes in leaf were 324 and 308 in flower buds. There were 57 overlapped novel genes between root and leaf while flower bud and leaf had 76 overlapped novel genes. The novel gene distribution between high and low hybrids revealed majority of novel genes were unique and less were overlapped among both hybrids (Figure S1). Only 122 genes were overlapped among hybrids wherein 877 genes were unique in high hybrid and 808 in low hybrids. These results anticipated that novel genes with unique expression most likely caused contrasting level of heterosis in hybrids.

Functional analysis of homoeolog genes exhibiting novel expression changes in hybrids

As to understand possible functions of homoeolog genes that had novel biased expression under hybridization in cotton, we performed GO and KEGG enrichment analysis in each hybrid. GO enrichment analysis ($p\text{-value} < 0.05$) for high hybrid showed a majority of novel biased expressed homoeolog genes had functional annotation related to catalytic activity, nucleus, metabolic process, transmembrane transport, binding of nucleotide, nucleic acid, protein, ATP, and RNA (Fig. 5a). Genes associated with binding of protein and ATP had the highest portion of novel biased expressed homoeolog genes in this hybrid. Most enriched GO terms for novel biased expressed homoeolog genes in low hybrid were involved in functional processes related to the nucleus, catalytic activity, binding of ATP, protein, and nucleotide (Fig. 5b). Interestingly, genes that performed the binding function of protein and ATP also had the majority of novel biased expressed homoeolog genes in this hybrid.

KEGG analysis revealed that the majority of novel biased expressed homoeolog genes in high hybrid were enriched in pathways such as biosynthesis of amino acids, mRNA surveillance pathway, Glycolysis / Gluconeogenesis, biosynthesis of antibiotics, metabolism of carbon, glutathione, cysteine/ methionine, biosynthesis of secondary metabolites, and microbial metabolism in diverse environments (Fig. 6a). Biosynthesis of secondary metabolites had more significant and gene enrichment as compared to other pathways. It was observed that a major portion of novel biased expressed homoeolog genes in low hybrid

had enrichment in mRNA surveillance pathway, RNA degradation, Oocyte meiosis, metabolism of pyruvate, carbon, sugars, and biosynthesis of secondary metabolites (Fig. 6b). In a low hybrid, biosynthesis of secondary metabolites also had a maximum number of novel biased expressed homoeolog genes as compared to other pathways.

Discussion

Cotton is the backbone of the textile industry because of its unique fiber characteristics. However, the area of cotton cultivation, as well as production, is decreasing in a recent decade all over the world. Hybridization has effectively resolved the problem of decreased crop productivity. For example, maize and rice cultivars developed with exploitation of heterosis have more harvest potential than conservative genotypes (51). Cotton belongs to the genus *Gossypium* and upland cotton (*G. hirsutum*) is stated as new world cotton. The hybridization in upland cotton whether it is between species (interspecific) or within species (intraspecific) produced meaningful heterosis in yield-related traits (42). It was observed that intraspecific crosses and improvement in fiber quality traits require more breeding efforts in cotton. How hybridization increased yield potential in upland cotton needs plenty of investigation at the genomic, transcriptomic, and epigenetic levels. Interestingly, the previous study has shown hybridization itself had sound effects on duplicate genes (homoeolog genes) in allopolyploids (2, 35). This study analyzed the effects of intraspecific hybridization on naturally occurring homoeolog genes in upland cotton through RNA-seq data of contrasting hybrids and full-length Pacbio sequences of their respective parents. In particular, gene expression data from three tissues (root, leaf, and flower buds) were used to explore changes in homoeolog expression and functional annotation of novel biased genes of hybrids.

Hybridization produced parent legacy of homoeolog expressions in upland cotton

Hybridization is a vital phenomenon to produce genetic improvement in crops. It is a common method in which two genetically different genotypes are crossed to obtain vigor progeny in desired traits. More than two sets of chromosomes in allotetraploid upland cotton caused genetic, gene expression, and epigenetic changes. These effects may result in more vigorous growth or heterosis (52). Previous studies in maize (53) and alfalfa (54) suggested that allopolyploids exhibit more heterosis as compared to autopolyploid and tetraploid display increased heterosis as compared with diploids. It was found that intraspecific crossing in upland cotton produced progeny that often displays heterosis in early biomass growth and seed cotton yield (45, 47). However, there has been no study to analyze changes in homoeolog gene expression of inbred parents due to hybridization. The result of our study revealed that intraspecific hybridization in upland cotton produced highly balanced changes in the pattern of homoeolog expression. The pattern was conserved in hybrids for the greater number of genes. For example, both higher and lower hybrids had around 95% homoeolog gene pairs with similar expression pre-existing in parents in all tissues under investigation. Fewer homoeolog gene copies imitated novel homoeolog expression patterns in hybrids. Further, fewest homoeolog gene pairs exhibited maternal or paternal

overall biases expression. A previous study in allopolyploid cotton showed that most homoeolog gene pairs reflected parental expression conditions in hybrid and synthetic allopolyploids (2). Similar results were reported in hybrid fish and its progenitors (35, 55). Assessment of homoeolog expressions in *Brassica napus* described that majority of homoeolog gene pairs maintained parental expression patterns in synthetic hybrids and natural allotetraploids (56). The effect of polyploidization on homoeolog expression in hexaploid wheat revealed equal homoeolog expression of parental lines is often conserved in progeny (16). Altogether, these results indicated that the effect of hybridization or genome doubling on the expression of homoeolog gene pairs in derived hybrids is often inherited from diploid parents. Though the relationship between heterosis and homoeolog gene pairs is still not fully understood, these findings suggest that fewer changes at the transcriptomic level could be associated with heterosis in upland cotton.

Novel homoeolog gene expressions may be regulated heterosis in cotton through secondary metabolism

As already described in the majority of the result of homoeolog gene pairs had a parental legacy in hybrids. Theoretically, these gene pairs do not contribute to the superior or inferior performance of cotton hybrids. Homoeolog gene pairs with a novel or differential expression in hybrids could regulate heterosis. Functional analysis of these novel homoeolog gene pairs revealed that the binding function of protein and ATP through the secondary metabolism pathway plays the role to increase growth potential in offspring relative to their parents. Metabolites are an essential component for adaptation, stress management, and regulation of different biological processes in plants (57, 58). Modification of secondary metabolism is useful to growth heterosis in oilseed crop *Brassica juncea* (59). Further, these researchers reported an additive mode of inheritance of metabolites in hybrids. Previous metabolic investigation in hybrids of maize (60), rice (61), and Arabidopsis (62) provided preliminary perceptions and enhanced knowledge about the molecular mechanism of plant heterosis. There is plenty of emphasis on genetic, gene expression, and epigenetic related studies to explore the biological basis of heterosis in cotton. Yet, the relationship between heterosis and metabolic changes in hybrids has been poorly known in cotton. Our results underpin the importance of genomic inheritance of secondary metabolites between hybrids and inbred parents in cotton.

Conclusion

In this study, global homoeolog gene expression patterns between intraspecific contrasting cotton hybrids and their reciprocal inbred parents were analyzed using RNA sequencing. The results indicated that hybridization caused no expression biases of homoeolog gene pairs in hybrids. Hybrids showed no biased expression preference towards any parents. Further, functional annotation of novel expressed homoeolog gene pairs explained changed secondary metabolic gene expression between parents and hybrids may regulate heterosis in upland cotton. Together, our results provide a general understanding of homoeolog gene expression patterns and may help to explore the genetic basis of heterosis in cotton.

Declarations

Acknowledgments

The authors thank Yasir Muhammad for his valuable suggestions.

Authors' contributions

Xing CZ and Wu JY designed this research. Gou LP, Qi TX, Tang HN, Qiao XQ, and Wang HL prepared plant population. Zhang XX, Zhang M, and Zhang BB carried out RNA extraction and performed data analysis. Shahzad K, Zhang XX, Xing CZ, and Wu JY drafted the manuscript. All authors read and approved the final manuscript.

Funding

This work was funded by the Zhongyuan Academician Foundation (212101510001), the Fundamental Research Funds for State Key Laboratory of Cotton Biology (CB2021C08), and the General Program of the National Natural Science Foundation of China (31871679). The funding organization was not involved in the design of the study, analysis, and writing the manuscript.

Availability of data and materials

All data generated or analyzed in this study included in published article

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Dufresne F and Hebert PD (1994) Hybridization and origins of polyploidy. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 258, 141-146
2. Yoo M, Szadkowski E and Wendel J (2013) Homoeolog expression bias and expression level dominance in allopolyploid cotton. *Heredity* 110, 171-180
3. Doyle JJ, Flagel LE, Paterson AH et al (2008) Evolutionary genetics of genome merger and doubling in plants. *Annual review of genetics* 42, 443-461
4. Kashkush K, Feldman M and Levy AA (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nature genetics* 33, 102-106
5. Flagel LE and Wendel JF (2010) Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation. *New Phytologist* 186, 184-193
6. Gaeta RT, Yoo S-Y, Pires J, Doerge RW, Chen ZJ and Osborn TC (2009) Analysis of gene expression in resynthesized *Brassica napus* allopolyploids using *Arabidopsis* 70mer oligo microarrays. *PLoS One* 4, e4760
7. Chelaifa H, Monnier A and Ainouche M (2010) Transcriptomic changes following recent natural hybridization and allopolyploidy in the salt marsh species *Spartina × townsendii* and *Spartina anglica* (Poaceae). *New Phytologist* 186, 161-174
8. Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB and Rieseberg LH (2009) The frequency of polyploid speciation in vascular plants. *Proceedings of the national Academy of sciences* 106, 13875-13879
9. Kovarik A, Dadejova M, Lim YK et al (2008) Evolution of rDNA in *Nicotiana* allopolyploids: a potential link between rDNA homogenization and epigenetics. *Annals of Botany* 101, 815-823
10. Salmon A, Flagel L, Ying B, Udall JA and Wendel JF (2010) Homoeologous nonreciprocal recombination in polyploid cotton. *New Phytologist* 186, 123-134
11. Szadkowski E, Eber F, Huteau V et al (2010) The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phytologist* 186, 102-112
12. Madlung A, Masuelli RW, Watson B, Reynolds SH, Davison J and Comai L (2002) Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. *Plant physiology* 129, 733-746
13. Kraitshtein Z, Yaakov B, Khasdan V and Kashkush K (2010) Genetic and epigenetic dynamics of a retrotransposon after allopolyploidization of wheat. *Genetics Molecular Research* 186, 801-812
14. Anssour S, Krügel T, Sharbel T, Saluz H, Bonaventure G and Baldwin I (2009) Phenotypic, genetic and genomic consequences of natural and synthetic polyploidization of *Nicotiana attenuata* and *Nicotiana obtusifolia*. *Annals of Botany* 103, 1207-1217
15. Han F, Fedak G, Guo W and Liu B (2005) Rapid and repeatable elimination of a parental genome-specific DNA repeat (pGc1R-1a) in newly synthesized wheat allopolyploids. *Genetics Molecular Research* 170, 1239-1245

16. Li A, Liu D, Wu J et al (2014) mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *The Plant Cell* 26, 1878-1900
17. Adams KL, Cronn R, Percifield R and Wendel JF (2003) Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of sciences* 100, 4649-4654
18. Bottley A, Xia G and Koebner R (2006) Homoeologous gene silencing in hexaploid wheat. *The Plant Journal* 47, 897-906
19. Buggs RJ, Elliott NM, Zhang L et al (2010) Tissue-specific silencing of homoeologs in natural populations of the recent allopolyploid *Tragopogon mirus*. *New Phytologist* 186, 175-183
20. Chaudhary B, Flagel L, Stupar RM et al (2009) Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (*Gossypium*). *Genetics Molecular Research* 182, 503-517
21. De Smet R, Adams KL, Vandepoele K, Van Montagu MC, Maere S and Van de Peer Y (2013) Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants. *Proceedings of the National Academy of Sciences* 110, 2898-2903
22. McGrath CL, Gout J-F, Johri P, Doak TG and Lynch M (2014) Differential retention and divergent resolution of duplicate genes following whole-genome duplication. *Genome research* 24, 1665-1675
23. Hughes TE, Langdale JA and Kelly S (2014) The impact of widespread regulatory neofunctionalization on homeolog gene evolution following whole-genome duplication in maize. *Genome research* 24, 1348-1355
24. Liu S-L and Adams KL (2010) Dramatic change in function and expression pattern of a gene duplicated by polyploidy created a paternal effect gene in the Brassicaceae. *Molecular biology evolution* 27, 2817-2828
25. Kleinjan DA, Bancewicz RM, Gautier P et al (2008) Subfunctionalization of duplicated zebrafish *pax6* genes by cis-regulatory divergence. *PLoS Genet* 4, e29
26. Grover C, Gallagher J, Szadkowski E, Yoo M, Flagel L and Wendel J (2012) Homoeolog expression bias and expression level dominance in allopolyploids. *New Phytologist* 196, 966-971
27. Flagel L, Udall J, Nettleton D and Wendel J (2008) Duplicate gene expression in allopolyploid *Gossypium* reveals two temporally distinct phases of expression evolution. *BMC biology* 6, 16
28. Hovav R, Udall JA, Chaudhary B, Rapp R, Flagel L and Wendel JF (2008) Partitioned expression of duplicated genes during development and evolution of a single cell in a polyploid plant. *Proceedings of the National Academy of Sciences* 105, 6191-6195
29. Samuel Yang S, Cheung F, Lee JJ et al (2006) Accumulation of genome-specific transcripts, transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. *The Plant Journal* 47, 761-775
30. Mochida K, Yamazaki Y and Ogiwara Y (2004) Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence

tags. *Molecular Genetics Genomics* 270, 371-377

31. Chang PL, Dilkes BP, McMahon M, Comai L and Nuzhdin SV (2010) Homoeolog-specific retention and use in allotetraploid *Arabidopsis suecica* depends on parent of origin and network partners. *Genome biology* 11, R125
32. Wang J, Tian L, Madlung A et al (2004) Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics* 167, 1961-1973
33. Auger B, Baron C, Lucas M-O et al (2009) Brassica orthologs from BANYULS belong to a small multigene family, which is involved in procyanidin accumulation in the seed. *Planta* 230, 1167
34. Gaeta RT, Pires JC, Iniguez-Luy F, Leon E and Osborn TC (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. *The Plant Cell* 19, 3403-3417
35. Ren L, Li W, Tao M et al (2016) Homoeologue expression insights into the basis of growth heterosis at the intersection of ploidy and hybridity in Cyprinidae. *Scientific reports* 6, 1-12
36. Akhunova AR, Matniyazov RT, Liang H and Akhunov ED (2010) Homoeolog-specific transcriptional bias in allopolyploid wheat. *BMC genomics* 11, 505
37. Schnable JC and Freeling M (2011) Genes identified by visible mutant phenotypes show increased bias toward one of two subgenomes of maize. *PloS one* 6, e17855
38. Wang J, Tian L, Lee H-S et al (2006) Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics* 172, 507-517
39. SHAHZAD K, Xue L, Tingxiang Q et al (2019) Genetic analysis of yield and fiber quality traits in upland cotton (*Gossypium hirsutum* L.) cultivated in different ecological regions of China. *Journal of Cotton Research* 2, 14
40. Birchler JA, Auger DL and Riddle NC (2003) In search of the molecular basis of heterosis. *The Plant Cell* 15, 2236-2239
41. Ali I, Shakeel A, Saeed A et al (2016) Combining ability analysis and heterotic studies for within-boll yield components and fibre quality in cotton. *J Animal Plant Sci* 26, 156-162
42. Li X, Shahzad K, Guo L et al (2019) Using yield quantitative trait locus targeted SSR markers to study the relationship between genetic distance and yield heterosis in upland cotton (*Gossypium hirsutum*). *Plant Breeding* 138, 105-113
43. Shahzad K, Qi T, Guo L et al (2019) Adaptability and stability comparisons of inbred and hybrid cotton in yield and fiber quality traits. *Agronomy* 9, 516
44. Xiao J, Li J, Yuan L and Tanksley SD (1995) Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. *Genetics* 140, 745-754
45. Shahzad K, Zhang X, Guo L et al (2020) Comparative transcriptome analysis of inbred lines and contrasting hybrids reveals overdominance mediate early biomass vigor in hybrid cotton. *BMC genomics* 21, 1-16
46. Yu S, Li J, Xu C et al (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences* 94, 9226-9231

47. Shahzad K, Zhang X, Guo L et al (2020) Comparative transcriptome analysis between inbred and hybrids reveals molecular insights into yield heterosis of upland cotton. *BMC Plant Biology* 20, 1-18
48. Wang Z, Gerstein M and Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics* 10, 57-63
49. Wu J, Lin L, Xu M et al (2018) Homoeolog expression bias and expression level dominance in resynthesized allopolyploid *Brassica napus*. *BMC genomics* 19, 1-13
50. Li M, Wang R, Wu X and Wang J (2020) Homoeolog expression bias and expression level dominance (ELD) in four tissues of natural allotetraploid *Brassica napus*. *BMC genomics* 21, 1-15
51. Schnable PS and Springer NM (2013) Progress toward understanding heterosis in crop plants. *Annual review of plant biology* 64
52. Chen ZJ (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu. Rev. Plant Biol.* 58, 377-406
53. Yao H, Gray AD, Auger DL and Birchler JA (2013) Genomic dosage effects on heterosis in triploid maize. *Proceedings of the National Academy of Sciences* 110, 2665-2669
54. Bingham E, Groose R, Woodfield D and Kidwell KK (1994) Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. *Crop Science* 34, 823-829
55. Ren L, Cui J, Wang J et al (2017) Analyzing homoeolog expression provides insights into the rediploidization event in gynogenetic hybrids of *Carassius auratus* red var. × *Cyprinus carpio*. *Scientific reports* 7, 1-11
56. Zhang D, Pan Q, Tan C et al (2016) Genome-wide gene expressions respond differently to A-subgenome origins in *Brassica napus* synthetic hybrids and natural allotetraploid. *Frontiers in plant science* 7, 1508
57. Kleessen S, Laitinen R, Fusari CM et al (2014) Metabolic efficiency underpins performance trade-offs in growth of *Arabidopsis thaliana*. *Nature Communications* 5, 1-10
58. Miller M, Song Q, Shi X, Juenger TE and Chen ZJ (2015) Natural variation in timing of stress-responsive gene expression predicts heterosis in intraspecific hybrids of *Arabidopsis*. *Nature communications* 6, 1-13
59. Bajpai PK, Reichelt M, Augustine R, Gershenzon J and Bisht NC (2019) Heterotic patterns of primary and secondary metabolites in the oilseed crop *Brassica juncea*. *Heredity* 123, 318-336
60. Wang Z, Xue Z and Wang T (2014) Differential analysis of proteomes and metabolomes reveals additively balanced networking for metabolism in maize heterosis. *Journal of proteome research* 13, 3987-4001
61. Ma Q, Hedden P and Zhang Q (2011) Heterosis in rice seedlings: its relationship to gibberellin content and expression of gibberellin metabolism and signaling genes. *Plant physiology* 156, 1905-1920
62. Korn M, Gärtner T, Erban A, Kopka J, Selbig J and Hinch DK (2010) Predicting *Arabidopsis* freezing tolerance and heterosis in freezing tolerance from metabolite composition. *Molecular plant* 3, 224-

Tables

Table 1

Homoeolog expression bias in different tissues of the high F₁ upland cotton hybrid

Expression in parents	Expression in progeny	Root (%)	Leaf (%)	Flower buds (%)
A=B	A=B Parental condition	35471 (96.2)	35734 (97.1)	35779 (97.1)
A<B	A<B Parental condition	4 (0.01)	0 (0.0)	1 (0.0)
A>B	A>B Parental condition	5 (0.01)	7 (0.02)	4 (0.01)
A>B	A=B No bias in progeny	190 (0.52)	279 (0.76)	210 (0.57)
A<B	A=B No bias in progeny	345 (0.94)	66 (0.18)	107 (0.29)
A=B	A>B Novel bias in progeny	416 (1.13)	375 (1.02)	391 (1.06)
A<B	A>B Novel bias in progeny	20 (0.05)	10 (0.03)	12 (0.03)
A>B	A<B Novel bias in progeny	10 (0.03)	18 (0.05)	10 (0.03)
A=B	A<B Novel bias in progeny	392 (1.06)	364 (1.0)	339 (0.92)
Overall A-biased in progeny		441 (1.20)	392 (1.06)	407 (1.10)
Overall B-biased in progeny		406 (1.10)	382 (1.04)	350 (0.95)
Total number of genes		36853	36853	36853
Here A= Maternal parent and B= Paternal parent				

Table 2

Homoeolog expression bias in different tissues of the low F₁ upland cotton hybrid

Expression in parents	Expression in progeny	Root (%)	Leaf (%)	Flower buds (%)
A=D	A=D Parental condition	35844 (97.3)	35827 (97.2)	35981 (97.6)
A<D	A<D Parental condition	0 (0.0)	0 (0.0)	1 (0.0)
A>D	A>D Parental condition	2 (0.01)	1 (0.0)	1 (0.0)
A>D	A=D No bias in progeny	202 (0.55)	233 (0.63)	107 (0.29)
A<D	A=D No bias in progeny	197 (0.53)	135 (0.37)	116 (0.31)
A=D	A>D Novel bias in progeny	290 (0.89)	352 (0.96)	294 (0.80)
A<D	A>D Novel bias in progeny	12 (0.03)	8 (0.02)	14 (0.04)
A>D	A<D Novel bias in progeny	6 (0.02)	8 (0.02)	8 (0.02)
A=D	A<D Novel bias in progeny	300 (0.81)	288 (0.78)	331 (0.90)
Overall A-biased in progeny		304 (0.83)	361 (0.98)	309 (0.84)
Overall D-biased in progeny		306 (0.84)	297 (0.81)	340 (0.92)
Total number of genes		36853	36853	36853
Here A= Maternal parent and D= Paternal parent				

Figures

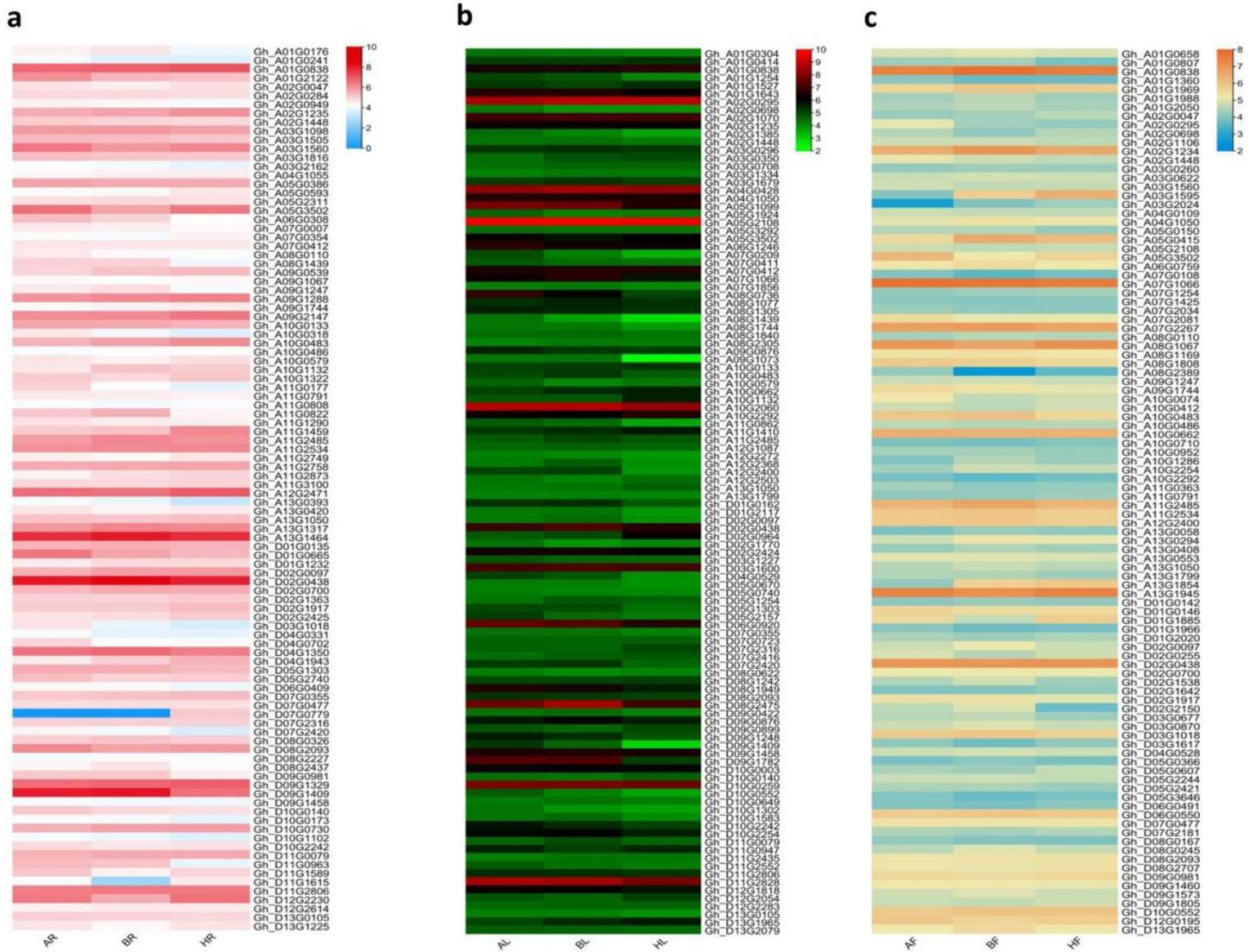


Figure 1

Heatmaps for novel biased homoeolog expressed genes in each tissue of high hybrid. **a** Novel biased expressed genes in root. **b** Novel biased expressed genes in leaf. **c** Novel biased expressed genes in flower buds.

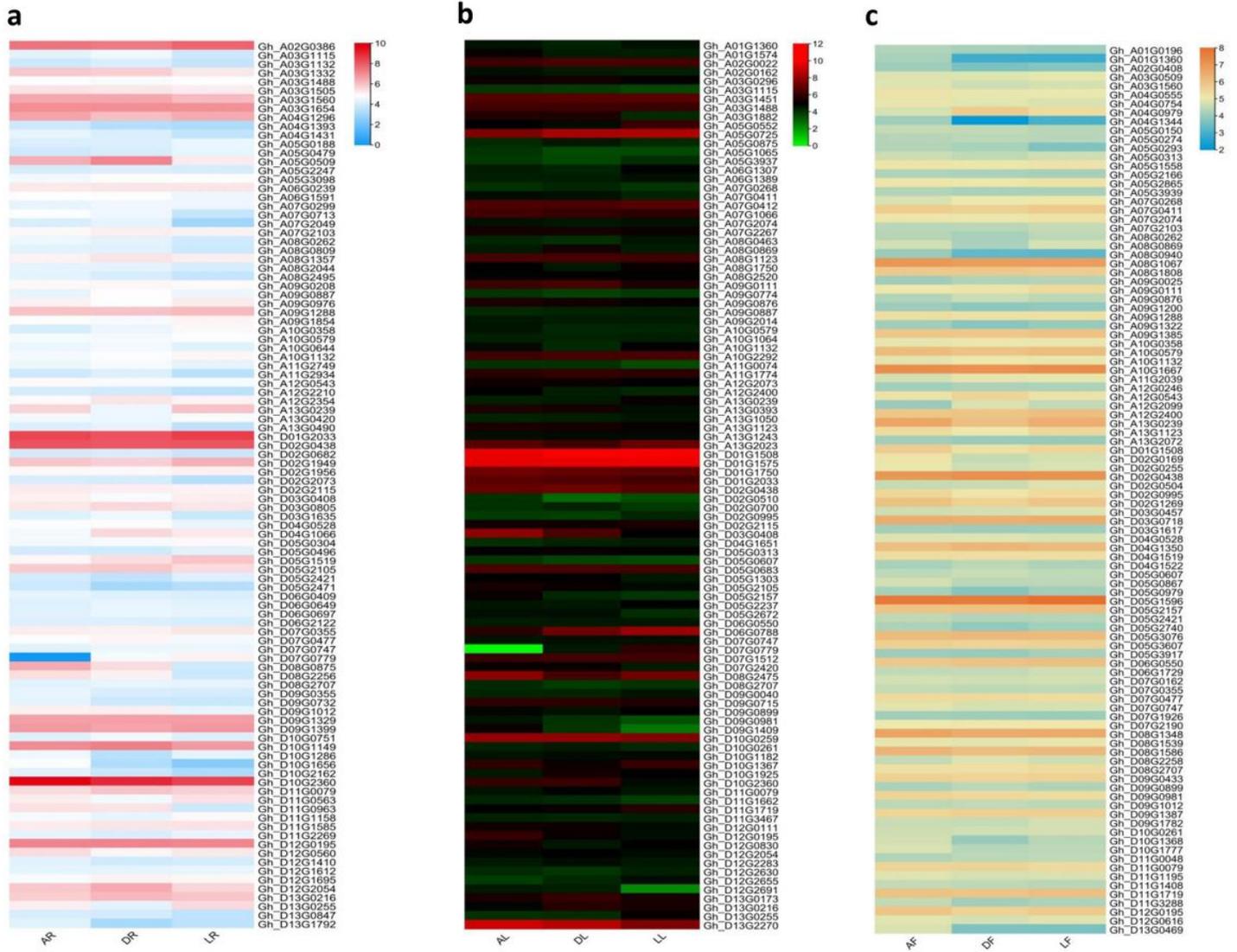


Figure 2

Heatmaps for novel biased homoeolog expressed genes in each tissue of low hybrid. **a** Novel biased expressed genes in root. **b** Novel biased expressed genes in leaf. **c** Novel biased expressed genes in flower buds.

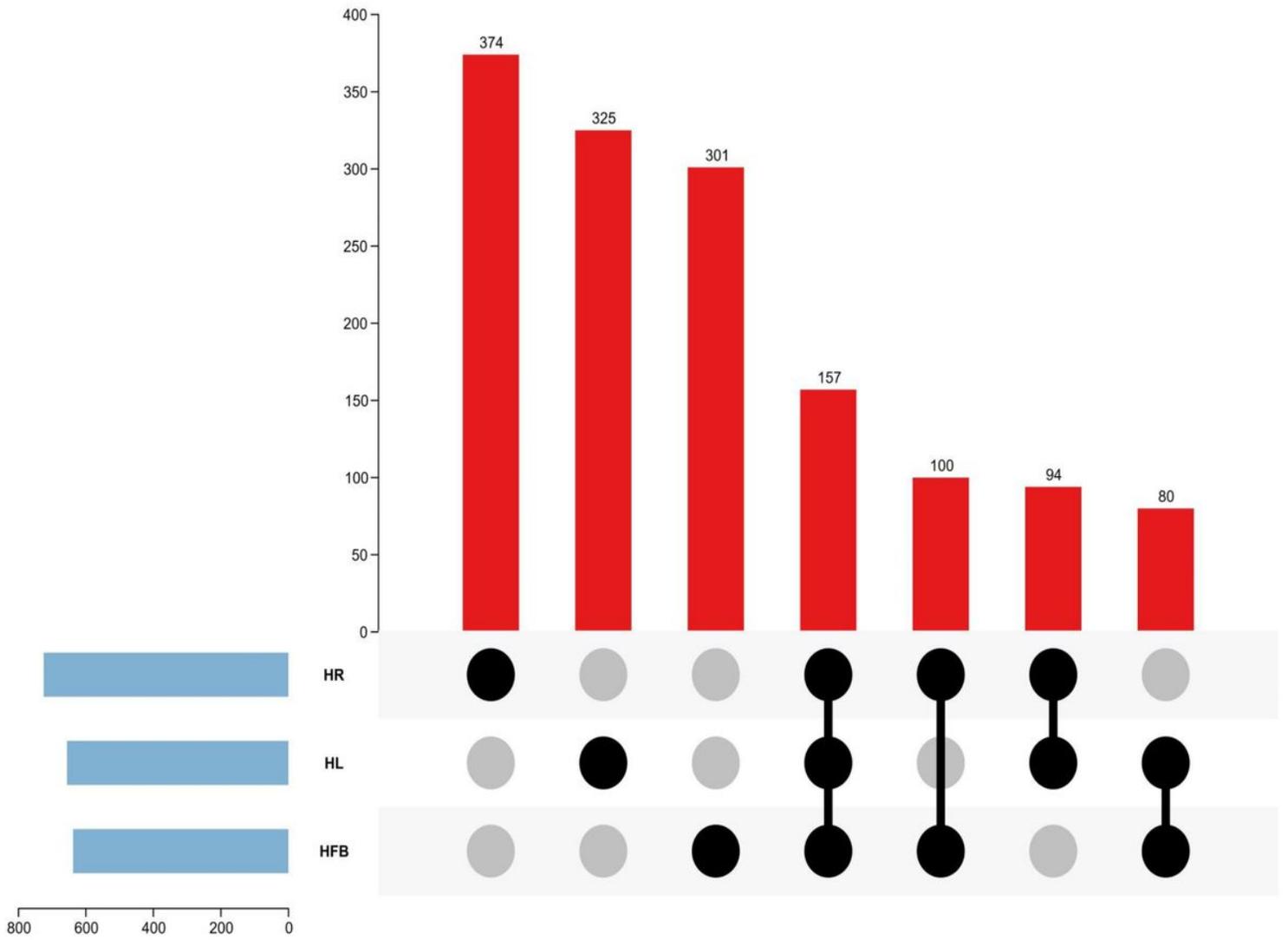


Figure 3

Distribution of novel biased homoeolog expressed genes in high hybrid. Here R, L, FB stands for root, leaf, flower buds, respectively.

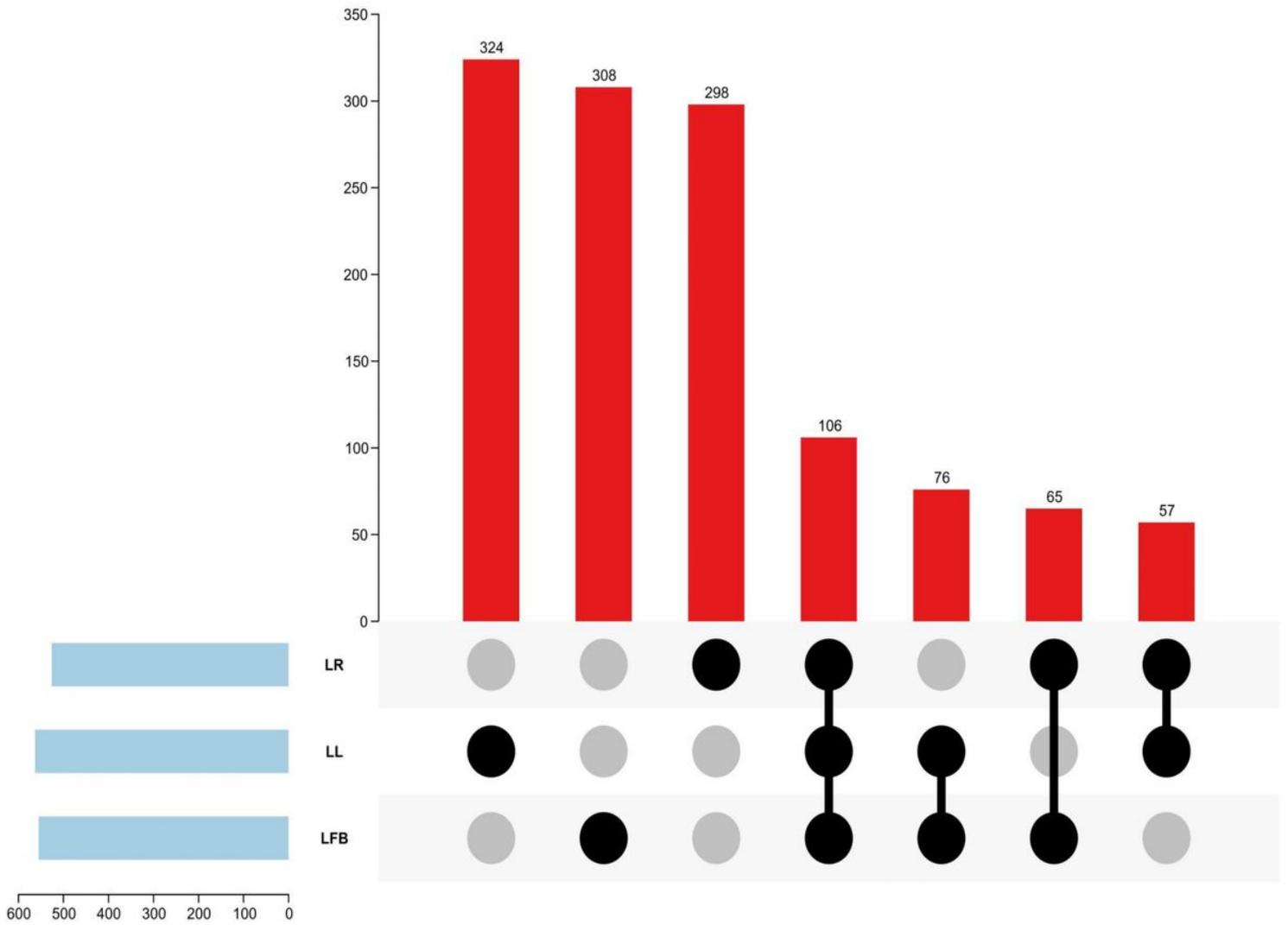


Figure 4

Distribution of novel biased homoeolog expressed genes in low hybrid. Here R, L, FB stands for root, leaf, flower buds, respectively.

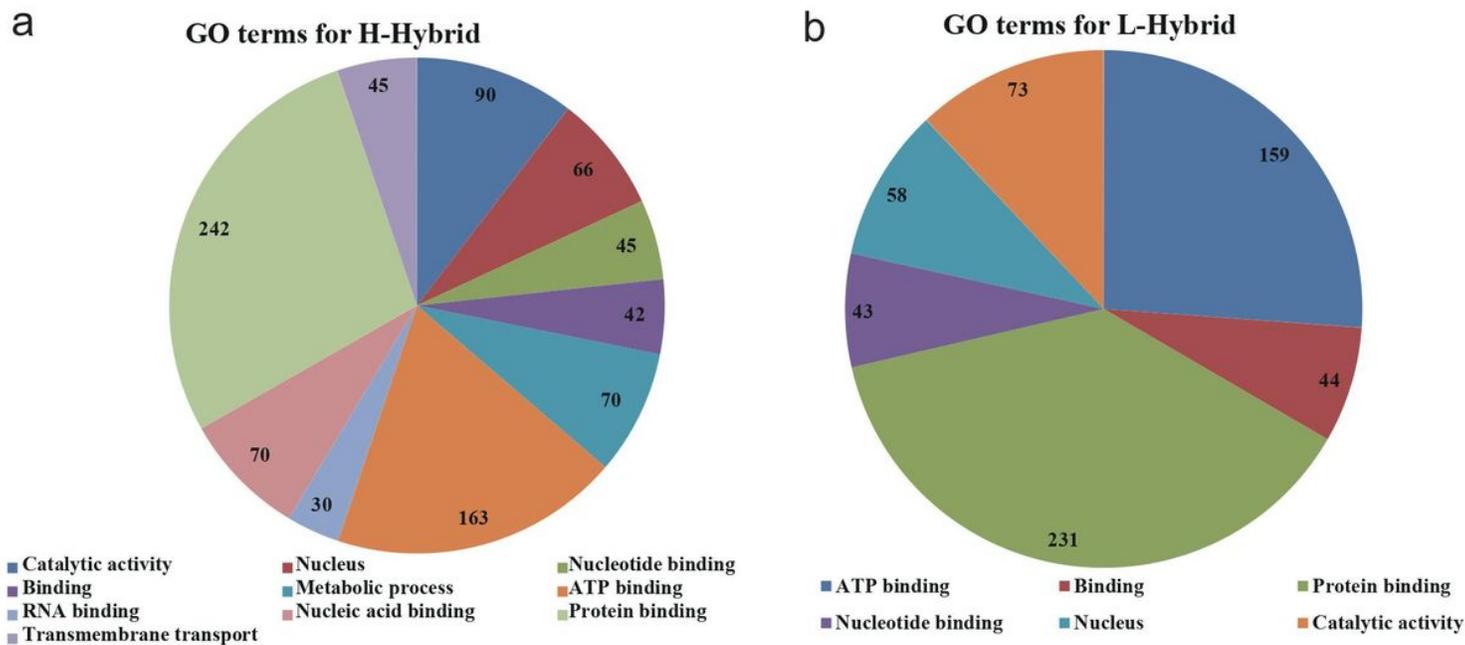


Figure 5

Most enriched GO terms for novel biased homoeolog expressed genes in each hybrid. **a** GO terms for high hybrid. Black numbers represent total genes for each term at significance of 0.05. **b** GO terms for low hybrid. Black numbers represent total genes for each term at significance of 0.05.

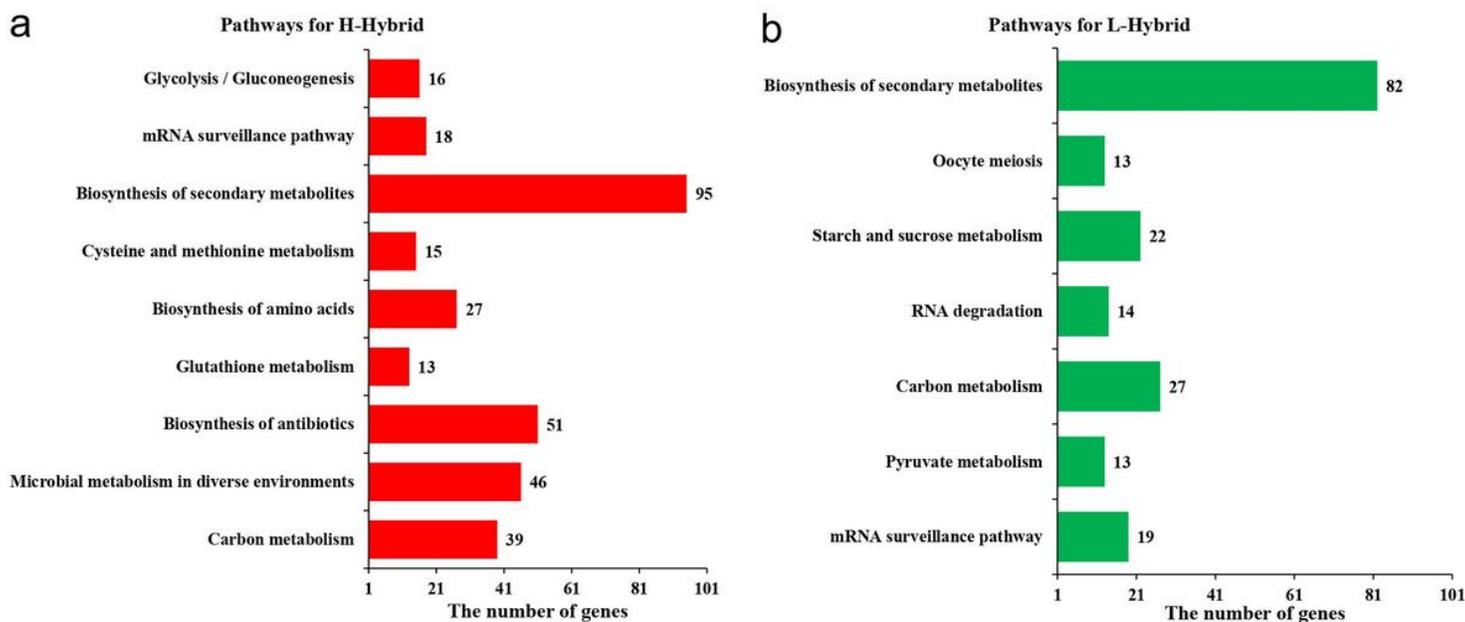


Figure 6

Most enriched Pathways for novel biased homoeolog expressed genes in each hybrid. **a** Pathways for high hybrid. **b** Pathways for low hybrid.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.png](#)