

RESEARCH

Open Access



Natural variations of heterosis-related allele-specific expression genes in promoter regions lead to allele-specific expression in maize

Weimin Zhan^{1,2†}, Lianhua Cui^{1†}, Shuling Yang¹, Kangni Zhang¹, Yanpei Zhang^{1*} and Jianping Yang^{1*}

Abstract

Background Heterosis has successfully enhanced maize productivity and quality. Although significant progress has been made in delineating the genetic basis of heterosis, the molecular mechanisms underlying its genetic components remain less explored. Allele-specific expression (ASE), the imbalanced expression between two parental alleles in hybrids, is increasingly being recognized as a factor contributing to heterosis. ASE is a complex process regulated by both epigenetic and genetic variations in response to developmental and environmental conditions.

Results In this study, we explored the differential characteristics of ASE by analyzing the transcriptome data of two maize hybrids and their parents under four light conditions. On the basis of allele expression patterns in different hybrids under various conditions, ASE genes were divided into three categories: bias-consistent genes involved in basal metabolic processes in a functionally complementary manner, bias-reversal genes adapting to the light environment, and bias-specific genes maintaining cell homeostasis. We observed that 758 ASE genes (ASEGs) were significantly overlapped with heterosis quantitative trait loci (QTLs), and high-frequency variations in the promoter regions of heterosis-related ASEGs were identified between parents. In addition, 10 heterosis-related ASEGs participating in yield heterosis were selected during domestication.

Conclusions The comprehensive analysis of ASEGs offers a distinctive perspective on how light quality influences gene expression patterns and gene-environment interactions, with implications for the identification of heterosis-related ASEGs to enhance maize yield.

Keywords Maize, Heterosis, Allele-specific expression, Promoter variation, Domestication

[†]Weimin Zhan and Lianhua Cui contributed equally to this work.

*Correspondence:

Yanpei Zhang
zhangyanpei@henau.edu.cn
Jianping Yang
jpyang@henau.edu.cn

¹College of Agronomy, Henan Agricultural University, Zhengzhou 450002, China

²Guangdong Provincial Key Laboratory of Plant Adaptation and Molecular Design, Guangzhou Key Laboratory of Crop Gene Editing, Innovative Center of Molecular Genetics and Evolution, School of Life Sciences, Guangzhou University, Guangzhou 510006, China



Background

Hybrid vigor, or heterosis, is a phenomenon where hybrid offspring exhibit superior vitality and growth potential than their parents. This concept has been used as a theoretical basis for crop breeding for over 150 years to enhance agricultural yield. Classical theoretical models, including dominance, overdominance, and epistasis, have been used to explain heterosis [1–5]. However, because heterosis is closely associated with agronomic traits and growth environments, these models do not fully capture the complexity of heterosis [6–8].

Advancements in sequencing technology in the last two decades have shed light on allele-specific expression (ASE), where the imbalance in allele expression between parents in a hybrid offers insights into heterosis [9–13]. The formation of ASE is a complex process regulated by both epigenetic and genetic variations in response to developmental and environmental conditions [14, 15]. A notable instance is genomic imprinting, an epigenetically driven phenomenon where expression depends on the allele's parental origin in hybrids [16]. The genetic mechanisms of ASE formation involve transcriptional regulation, post-transcriptional regulation, and translational regulation [14]. For instance, the whole-genome analysis of apples revealed that transposable element insertions in the upstream region of genes affect ASE gene (ASEG) transcription [13]. Similarly, in hybrid rice, nonsense-mediated mRNA decay led to biased consistent expression of alleles [17]. Furthermore, studies in rice and potato have identified a negative correlation between ASE and differences in CHG (H=A, C, or T) methylation [12, 18].

Light, as a pivotal environmental cue, not only drives photosynthesis but also acts as a signaling factor guiding plant growth and reproduction. Plants primarily monitor and absorb far-red (700–750 nm), red (600–700 nm), and blue light (400–500 nm) [19–21], which are crucial for carbohydrate transport, photosynthesis, and terpenoid biosynthesis in maize hybrids, respectively [22]. Continuous light irradiation significantly enhances biomass heterosis [22]. However, the genetic and molecular mechanisms underlying the effect of different light spectra on maize hybrid ASEG remain to be elucidated.

In this study, we comprehensively analyzed ASE by using transcriptome data from reciprocal hybrids and their parent strains, B73 and Mo17, under different light treatments. A total of 5,273 ASEGs were identified, which were categorized into three expression types. Gene Ontology (GO) analysis was performed to identify ASEG functions under different light conditions. In addition, we explored promoter variations in heterosis-related ASEGs and assessed how nucleotide diversity contributes to ASEG domestication. These findings not only advance

our understanding of maize heterosis but also provide a crucial empirical foundation for crop breeding.

Methods

Plant materials

The samples for RNA-seq were maize inbred lines B73 and Mo17 and their reciprocal hybrids B73×Mo17 and Mo17×B73. These experiments were conducted under darkness, far-red, red, and blue light conditions [22]. Briefly, maize seedlings were initially grown in darkness (26 °C) for 6 days, followed by exposure to far-red (737 nm, 2.5 $\mu\text{mol m}^{-2} \text{s}^{-1}$), red (658 nm, 30.0 $\mu\text{mol m}^{-2} \text{s}^{-1}$), and blue (447 nm, 6.0 $\mu\text{mol m}^{-2} \text{s}^{-1}$) light. After 24 h, samples of the seedlings were collected for RNA-seq analysis. Three biological replicates were performed for each sample.

RNA-seq data processing and ASEG identification

RNA-seq data were obtained from the National Center for Biotechnology Information Sequence Read Archive (SRA) database (<https://www.ncbi.nlm.nih.gov/sra>; accession no. PRJNA780806). Firstly, raw reads were pre-processed with Trimmomatic software v0.39 [23], where reads with mass less than 30 were filtered out (Fig. S1A). All clean reads from B73 (four treatments, three replications) were mapped to the reference genome of B73 V4 [24] by using Hisat2 v2.1.0 [25] with default parameters. SAMtools software v1.9 [26] was used to remove aligned reads with a mapping quality score of <30. Single nucleotide polymorphisms (SNPs) were called using the short variant identification pipeline of GATK v4.1.3.0 [27]. In addition, VCFtools software v0.1.13 [28] was employed to filter out variants with an allele frequency of <0.9 and a depth below 20. The base calls at SNP locations were then substituted with the corresponding bases from the B73 V4 reference genome to reconstruct a pseudo-genome reference through GATK (Fig. S1B).

Clean reads from all samples (four genotypes, four treatments, and three replications) were aligned to the pseudo-genome by using STAR software v2.7.2a [29]. Gene expression levels were quantified as transcripts per million (TPM) (Fig. S1C). GATK and SnpEff software v5.0e [30] were used to identify and annotate SNP variants, respectively. The phASER software v0.9.9.4 [31] was used for ASE identification and annotation at SNP loci, excluding alleles with a depth of <10. The chi-square test was performed to assess the differential expression of allele abundance, adopting a significance threshold of $P < 0.05$. Gene with imbalanced expression between the parental alleles in F_1 hybrids (≥ 2 replicates) were designed as ASEG (Fig. S1D). On the basis of their expression patterns, genes were classified into four groups: bias-consistent (BC) ASEGs, bias-specific (BS) ASEGs, bias-reversal (BR) ASEGs, and non-bias genes.

The BC ASEGs consistently favored one parental allele under all light conditions. BS ASEGs exhibited a preference for one parental allele under certain light conditions. However, BR ASEGs displayed a directional shift, favoring one parental allele under some light conditions and the other parental allele under other light conditions (Fig. S2).

Gene expression analysis

Differentially expressed gene (DEG) analysis was performed using the DESeq2 v1.40.1 R package [32], with a false discovery rate (FDR) threshold of <0.01 and a fold change (FC) of >1.5. DEGs between hybrids and their parental lines were categorized into three primary groups (additive, complete-incomplete dominant, and overdominant) or further detailed into 12 subcategories (types I to XII) following previously described criteria [22, 33]. Gene expression levels of F₁ hybrids in additive categories (types I and II) were between those of the two parental inbred lines. Furthermore, gene expression levels of F₁ hybrids in the complete-incomplete dominant categories (types III, IV, V, and VI) were similar to those in male or female parents. Gene expression levels of F₁ hybrids in the overdominant categories (types VII, VIII, IX, X, XI, and XII) were lower or higher than those of either parent. In addition, complete-incomplete dominant and overdominant categories were regarded as non-additive genes.

GO enrichment analysis

GO enrichment analysis of ASEGs was performed using the AgriGO v2.0 database [34], with B73 V4 selected as the background. The enrichment results ($P < 0.05$) were visualized using the clusterProfiler v4.8.2 R package [35].

Promoter variation analysis

NUCmer v4.0.0beta5 [36] was used to compare between B73 and Mo17 genomes [37]. Filtering results were analyzed using the delta-filter program (parameter '-1 -qr') and parsed using show-coords (parameter '-qclT'). Only collinear regions on identical chromosomes were retained. Finally, single-nucleotide polymorphism (SNP) and insertion/deletion (InDel) were identified using show-snps (parameter '-ClrT -x 1'). VCFtools was used to extract variants (SNPs, deletions, and insertions) in the 3000 bp promoter region of ASEGs, and frequencies were calculated. Variant information was annotated using SnpEff.

Nucleotide diversity analysis

Nucleotide diversity (π) was evaluated among 223 accessions, comprising 23 teosinte accessions and 200 maize inbred lines. On the basis of the third-generation Hapmap3 data of *Z. mays* [38], nucleotide diversity was

analyzed using the PopGenome v2.7.7 R package [39] with parameters 'sliding width=1000', and 'sliding jump=300'. Average nucleotide diversity in the 100-kb upstream region, middle region including ASEGs, and 100-kb downstream region was assessed for each group.

Results

Global identification of ASEG expression in maize

To investigate the effect of light quality on ASEGs in maize hybrids, we examined the transcriptome data of maize inbred lines B73 and Mo17 and their reciprocal hybrids B73×Mo17 (BM) and Mo17×B73 (MB) subjected to darkness or exposure to far-red, red, or blue light [22]. A total of 5,273 ASEGs were identified and categorized into three groups: bias-consistent (BC, 395), bias-specific (BS, 4,754), and bias-reversal (BR, 124) ASEGs (Table S1). No significant difference was observed in the number of ASEGs identified in BM and MB hybrids under various light conditions (Fig. 1A, Table S2). Notably, the number of Mo17-biased ASEGs (1,321–1,441) was significantly higher than those of biased B73 ASEGs (411–539) under all four light conditions (Fig. 1A, Table S2). BM and MB have 571 and 557 common ASEGs under four treatment conditions, respectively (Fig. 1B and C). Among these, 395 BC genes (70 biased toward B73 and 325 biased toward Mo17) were shared between BM and MB (Fig. 1D, Table S3). Despite a strong correlation in bias frequency between the reciprocal hybrids ($R^2 = 0.89$) (Fig. S3A), the bias frequency of hybrids exhibited a weak correlation with the gene expression levels of parents ($R^2 = 0.38$) (Fig. S3B). These results indicate that hybridization alters gene expression levels from theoretical predictions; however, reciprocal hybrids do not affect the bias of ASEGs.

Biased B73 and Mo17 BC ASEGs participate in the basal metabolism of hybrids in a functionally complementary manner

Hierarchical clustering analysis of the 395 BC ASEGs in parents and hybrids revealed their segregation into two main clusters (cluster1 and cluster2). In cluster1, the expression level of the B73 allele was higher than that of the Mo17 allele (90/395). By contrast, in cluster2, the expression level of the Mo17 allele was higher than that of the B73 allele. In addition, the expression levels of alleles in hybrids (305/395) was intermediate between those of the two parents (Fig. 2A). Furthermore, the expression patterns of these 395 BC ASEGs were categorized into 12 subgroups, primarily including additive, dominant (expression level dominance-female, ELD-M; expression level dominance-male, ELD-F) and overdominant (transgressive down/up regulation) effects. Additive effects constituted the majority (30.45–50.32%), followed by ELD-M (13.25–34.08%), ELD-F (15.43–29.55%), and overdominance (4.59–19.70%).

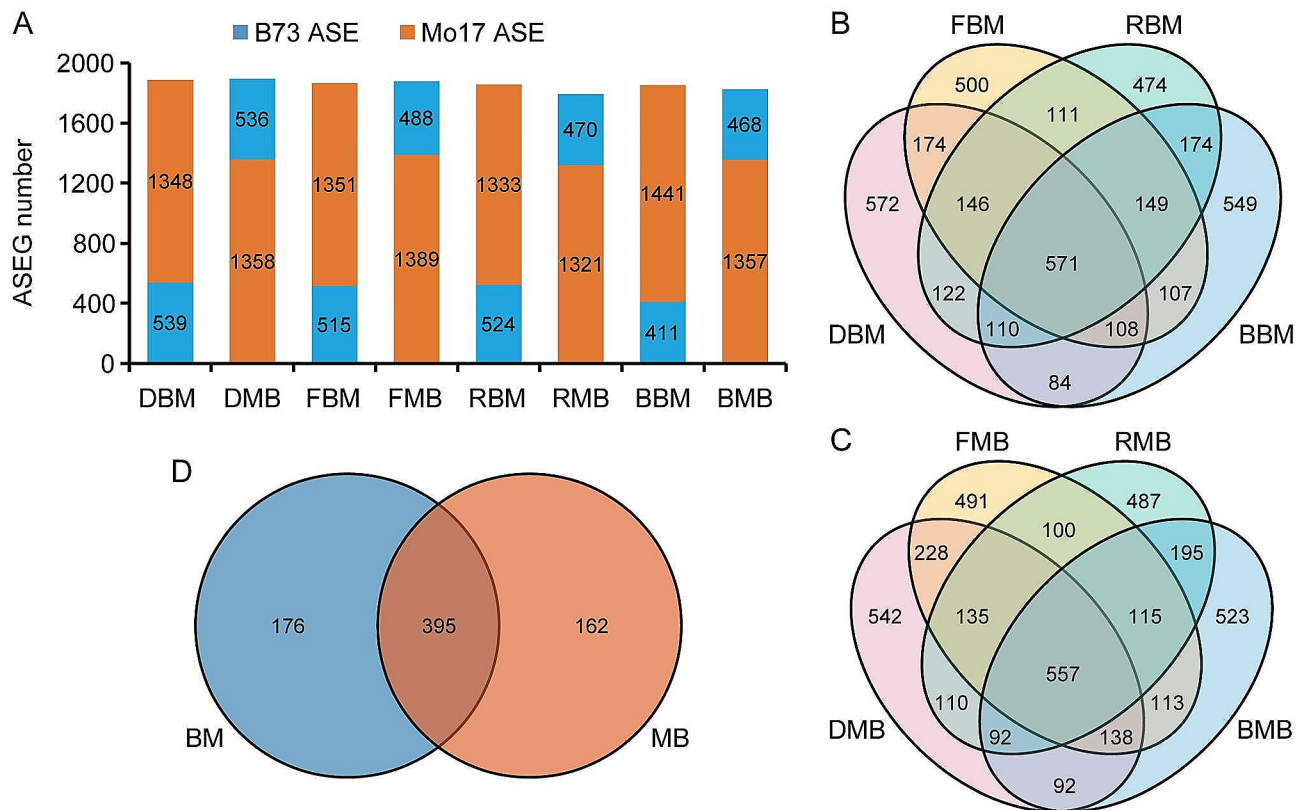


Fig. 1 Summary of ASEGs. **(A)** Number of ASEGs in the backcross under various light conditions. **(B-C)** Venn diagrams demonstrate the number of ASEGs in BM and MB under various light conditions. **(D)** Overlapping ASEGs in BM and MB under various light conditions. BM and MB represent B73×Mo17 and Mo17×B73, respectively. DBM, FBM, RBM, and BBM represent BM grown in darkness and under far-red, red, and blue light conditions, respectively. DMB, FMB, RMB, and BMB represent MB grown in darkness and under far-red, red, and blue light conditions, respectively

To investigate the potential effects of variations within the BC ASEGs, we compared the coding sequences of the 395 ASEGs between two parents, using B73 as the reference and employing the SnpEff software for annotation. The results showed that 51 (12.91%) of the 395 ASEGs, comprising nine B73 ASEGs and 42 Mo17 ASEGs, contained InDels and SNPs with high impact, potentially leading to protein truncation, dysfunction, or the initiation of nonsense-mediated decay. The most frequent variation was frameshift mutations caused by non-triplet nucleotide insertions/deletions ($n=33$, 64.71%). Furthermore, 233 (58.99%) of the 395 ASEGs, including 34 B73 ASEGs and 199 Mo17 ASEGs, possessed non-disruptive variants with moderate impacts, which might alter protein function. The predominant category within this group was missense variants ($n=210$), likely causing changes in protein function but not major structural alterations. The remaining 111 genes comprised low-impact variations (unlikely to alter protein sequences) and modifier-impact mutations (impact undetermined), accounting for 97 (24.56%) and 14 (3.54%) genes, respectively. These mutations do not affect gene expression and structure (Fig. 2C). The diversity of variations suggests

that changes in coding sequence are not the primary drivers of ASEGs.

GO enrichment analysis of the 395 BC ASEGs using InterPro classification demonstrated that biased B73 ASEGs were mainly involved in ribosome-related functions, including ribosome assembly, ribosome biogenesis, ribonucleoprotein complex assembly, and ribonucleoprotein complex subunit organization. Biased Mo17 ASEGs were primarily associated with abiotic stress responses and organic compound biosynthesis and metabolism, including photosynthesis, light reaction, organic hydroxy compound biosynthetic process, single-organism metabolic process, lipid catabolic process, and oxidation-reduction process (Fig. 2D). These findings indicate that BC ASEGs originating from different parental sources exhibit functional complementarity in the F_1 hybrids and participate in basic metabolic processes through additive expression modes. Furthermore, BC ASEGs were consistent in both BM and MB hybrids, suggesting that the complementary manner is the same in reciprocal hybrids.

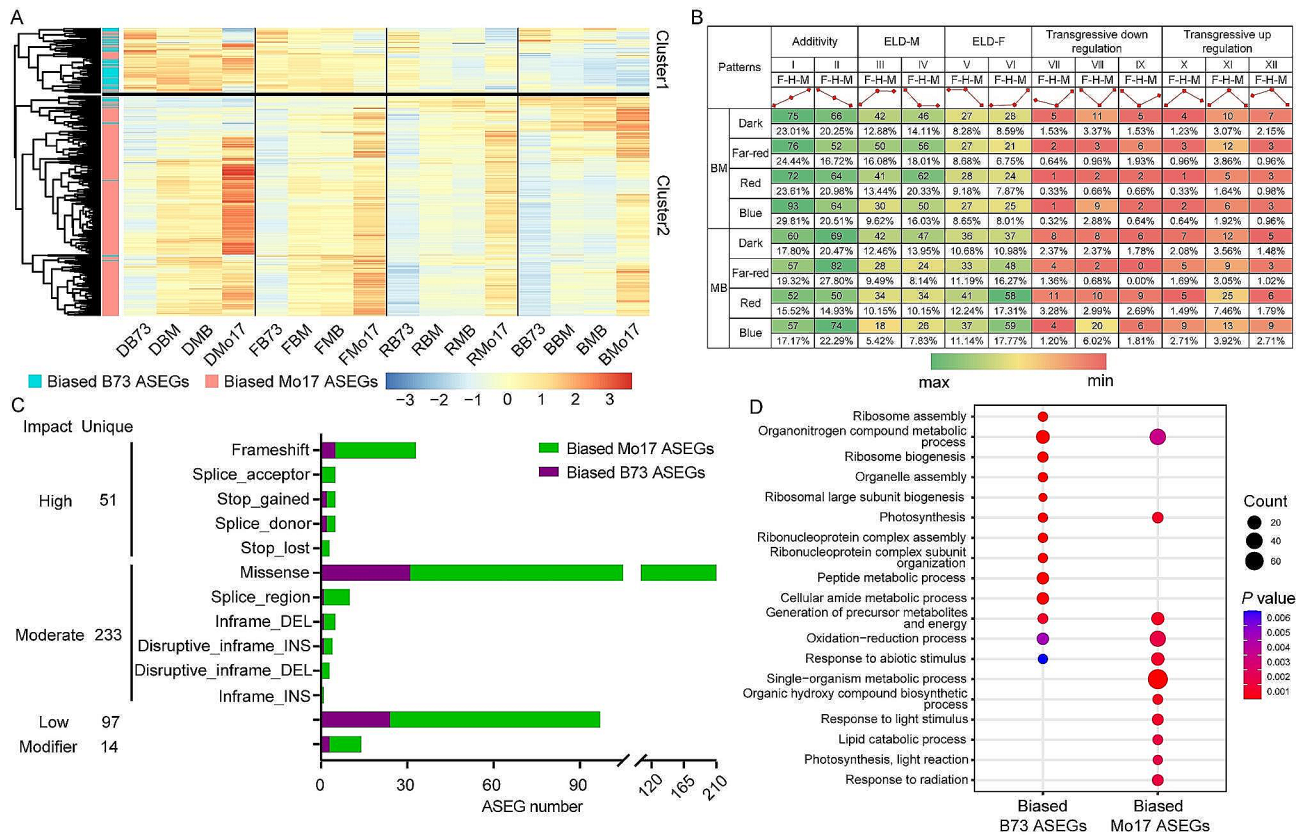


Fig. 2 Characteristic analysis of 395 BC genes. **(A)** Expression levels of 395 genes in parents and hybrids. DB73, DMo17, DBM, and DMB represent B73, Mo17, BM, and MB grown in darkness. FB73, FMo17, FBM, and FMB represent B73, Mo17, BM, and MB grown under far-red light condition, respectively; RB73, RMo17, RBM, and RMB represent B73, Mo17, BM, and MB grown under red light condition, respectively; and BB73, BMo17, BBM, and BMB represent B73, Mo17, BM, and MB grown under blue light condition, respectively. **(B)** Expression patterns of 395 genes in hybrids. ELD, expression level dominance; F, female; H, hybrid; M, male. **(C)** Features of BC ASEGs with low, moderate, high, and modifier impact variations. The unique numbers of each impact category are indicated. DEL, deletion; INS, insertion. **(D)** GO enrichment analysis of 395 BC genes

BR ASEGs are beneficial for hybrids to adapt to different light environments

Under the four light conditions, 124 BR ASEGs were identified. A total of 91 genes exhibited biased transition under different conditions, including 56 ASEGs in BM, 58 ASEGs in MB, and 23 ASEGs common to both reciprocal hybrids (Fig. 3, Table S4). Moreover, 33 genes displayed biased conversion between different genotypes, with genes such as *Zm00001d022421*, *Zm00001d032956*, *Zm00001d034034*, *Zm00001d047349*, *Zm00001d051804*, *Zm00001d042906*, *Zm00001d001966*, and *Zm00001d009717* showing consistent bias in the same genotypes (Fig. S4).

GO functional annotation analysis revealed that the 124 BR ASEGs were predominantly concentrated in processes related to the chloroplast stroma, photosynthetic membrane, and cellular homeostasis (Table S5). Among the ASEGs shared by BM and MB, *Zm00001d044686*, a lipid transport protein, is involved in the mid-stage response to sustained endoplasmic reticulum stress and plays a role in protecting plants from adverse

environmental conditions [40]. In BM, the expression level of *Zm00001d044686* showed a Mo17 bias under darkness, whereas it exhibited a B73 bias under the other three light conditions. By contrast, in MB, its expression level was B73-biased under darkness and far-red light but Mo17-biased under red and blue light conditions. *Zm00001d049650* encodes the photosystem II core complex protein PsbY [41]. Its expression in BM and MB was B73-biased under red and far-red light conditions but Mo17-biased under darkness and far-red light conditions. Notably, it showed a Mo17 bias specifically under far-red light conditions. *Zm00001d016926*, homologous to the *Arabidopsis ABC1KB*, is an atypical protein kinase induced by heavy metals [42]. In BM, its expression was Mo17-biased under darkness and far-red light conditions but B73-biased under red and blue light conditions. In MB, its expression was Mo17-biased under far-red light and B73-biased under red and blue light conditions. *Zm00001d023707*, a member of the thio-redoxin superfamily, regulates various plant functions, such as germination, development, photosynthesis, and

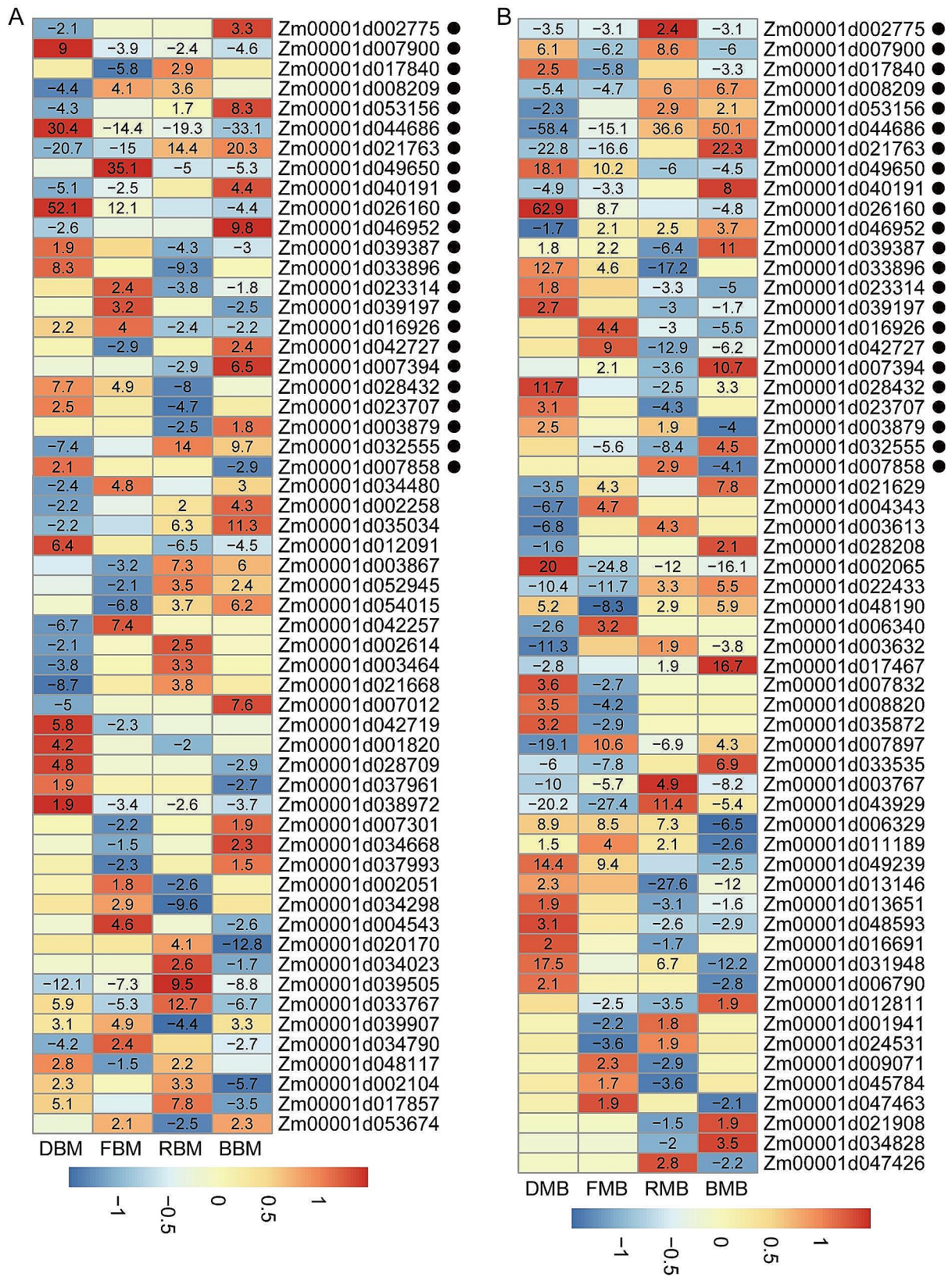


Fig. 3 Directions of expression bias for BM (A) and MB (B) ASEGs under four light conditions. Black dots represent that ASEGs are shared in both BM and MB. The heatmap is labeled using the $-\log_{10}$ (P value) of significantly biased SNPs. The marker is a negative number for biased B73 and a positive number for biased Mo17

flowering. Its expression in both BM and MB was Mo17-biased in darkness and B73-biased under red light conditions (Fig. 3). These findings suggest that the conversion of allele bias plays a significant role in plant adaptation to diverse light environments.

BS ASEGs maintain cell stability through both dominant and additive expression

In the BS ASEGs, B73-biased and Mo17-biased genes constituted 28.40% (1,350/4,754) and 71.60% (3,404/4,754), respectively. Notably, the majority of biased genes were exclusive to a single light condition (2,816, 59.23%), with decreasing frequencies under two (998, 20.99%), three (610, 12.83%), and all four light conditions (330, 6.94%) (Fig. S5). Hierarchical clustering analysis indicated that although the expression levels of BS ASEGs significantly varied under different light conditions, the majority of genes in the F₁ generation mirrored the expression trends of their respective parental lines (Fig. 4A). The 4,754 BS ASEGs were further categorized into 12 types, with gene expression in the F₁ predominantly characterized by dominant (18.09–42.77%) and additive effects (15.27–38.88%) (Fig. 4B). The proportion

of genes exhibiting low parental expression in F₁ dominant effects (17.05–19.94% in BM and 16.12–22.27% in MB) surpassed those with high parental expression (11.65–17.81% in BM and 12.23–20.50% in MB) (Fig. 4B).

GO enrichment analysis revealed that B73-biased genes were mainly associated with cellular components, such as plastid, cytoplasm, intracellular space, and chloroplast, whereas Mo17-biased genes were predominantly related to membrane structures, including membrane parts, and intrinsic and integral components of the membrane (Fig. 4C). Furthermore, BS ASEG functions showed complete overlap under dark (2,238), far-red (2,220), red (2,166), and blue (2,184) light conditions (Fig. S6). These findings suggest that BS genes contribute to the adaptability of hybrid plants to diverse light environments through dominant and additive expression modes.

Promoter variations in heterosis-related ASEGs cause ASE formation

Xiao et al. conducted a genome-wide association analysis on days to tasseling, plant height, and ear weight in 42,840 F₁ hybrids and identified 4,307 heterosis-related genes [43]. The 758 ASEGs identified in this study were

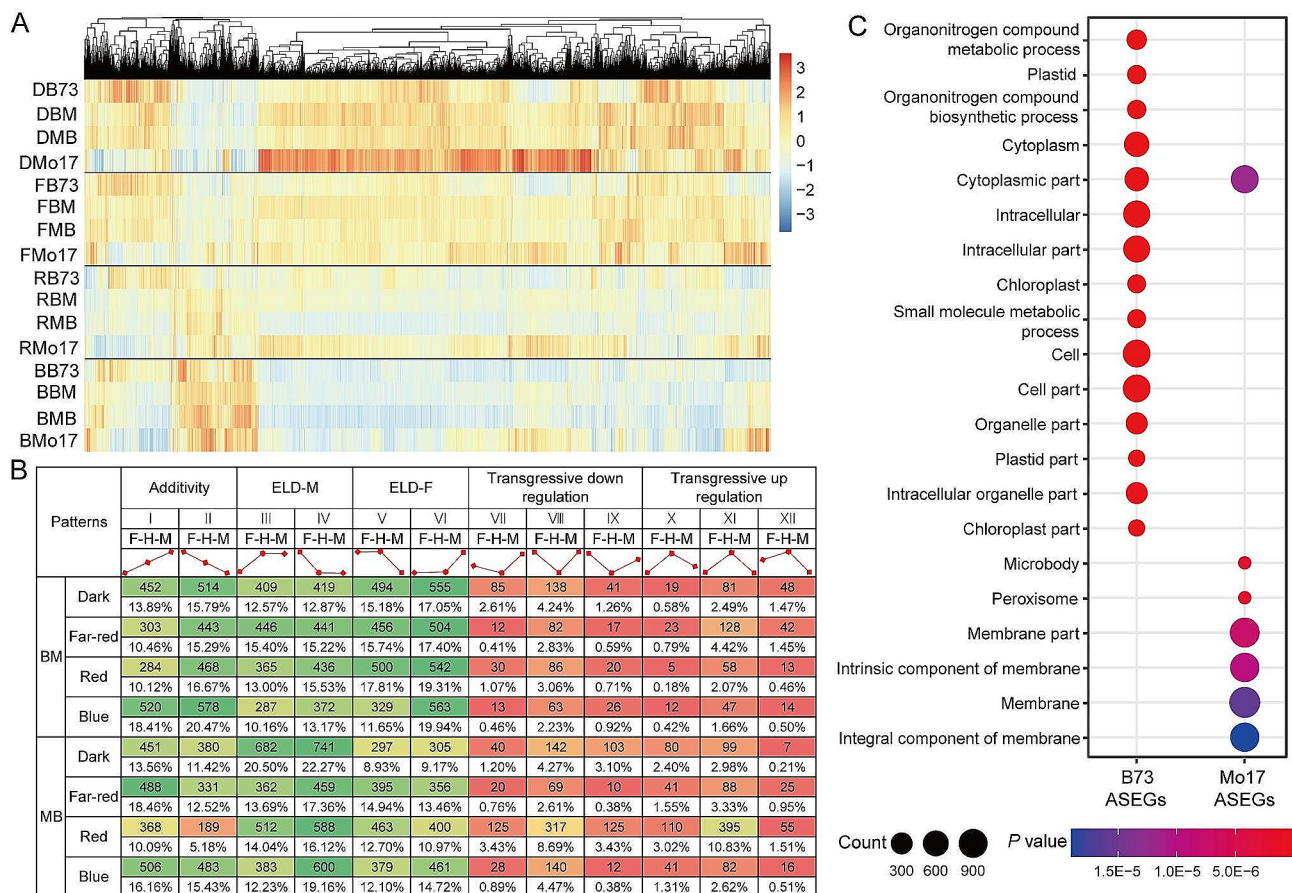


Fig. 4 Expression pattern and function analysis of 4,754 BS ASEGs. **(A)** Expression levels of BS ASEGs in hybrids and parents. **(B)** Twelve expression patterns of BS ASEGs in hybrids. ELD, expression level dominance; F, female; H, hybrid; M, male. **(C)** GO enrichment analysis of biased parent-related BS ASEGs

significantly overexpressed in these regions ($P=5.04E-3$, hypergeometric test) (Fig. 5A), including 684 BS, 60 BC, and 14 BR ASEGs (Fig. 5B, Table S6). When analyzing the coding regions of the 758 heterosis-associated ASEGs, the variation types between B73 and Mo17 sequences were found to be considerably complex. Some genes, such as *Zm00001d011210*, had multiple differential sites (22 loci), whereas others, such as *Zm00001d028180*, had only two SNPs (Fig. 5C). Even in shorter coding regions, significant variations were noted; for instance, *Zm00001d039432* had 15 variant sites. These findings along with previous results (Fig. 2C) suggest that changes in the coding region sequence of heterosis-related ASEGs may not be the primary factor in ASE formation. However, the impact of ASEG coding regions on heterosis cannot be ruled out. To explore whether differences in promoter regions contribute to ASEG formation, 3,000 bp sequences of the promoter regions of the 758 ASEGs between B73 and Mo17 were examined. The results revealed a high frequency of InDels and SNPs in these promoter regions (Fig. 5D), indicating that

variations in the promoter region were associated with heterosis-related ASE.

Ten heterosis-related ASEGs underlie artificial selection during maize domestication

Out of 758 heterosis-related ASEGs, 10 have been identified as playing a role in maize heterosis, specifically in stress resistance (*ZmRap2.7/Zm00001d010987*, *Zm00001d042314*, and *ZmPHD17/Zm00001d010974*) and yield (*ZmACO2/Zm00001d024952*, *ZmASN4/Zm00001d047736*, *ZmLYCE1/Zm00001d011210*, *Zm00001d028180*, *Zm00001d053090*, *ZmGSK/Zm00001d016188*, and *ZmGAE1/Zm00001d039432*) heterosis [44]. To determine whether artificial selection has impacted these 10 heterosis-related ASEGs, we examined the nucleotide diversity of 23 teosinte and 200 maize accessions by using hapmap3.2.1 data [38]. For most ASEGs, the average nucleotide diversity in flanking regions was significantly higher in teosinte than in maize accessions (Fig. 6A–J, Table S7). These results suggest that some ASEGs

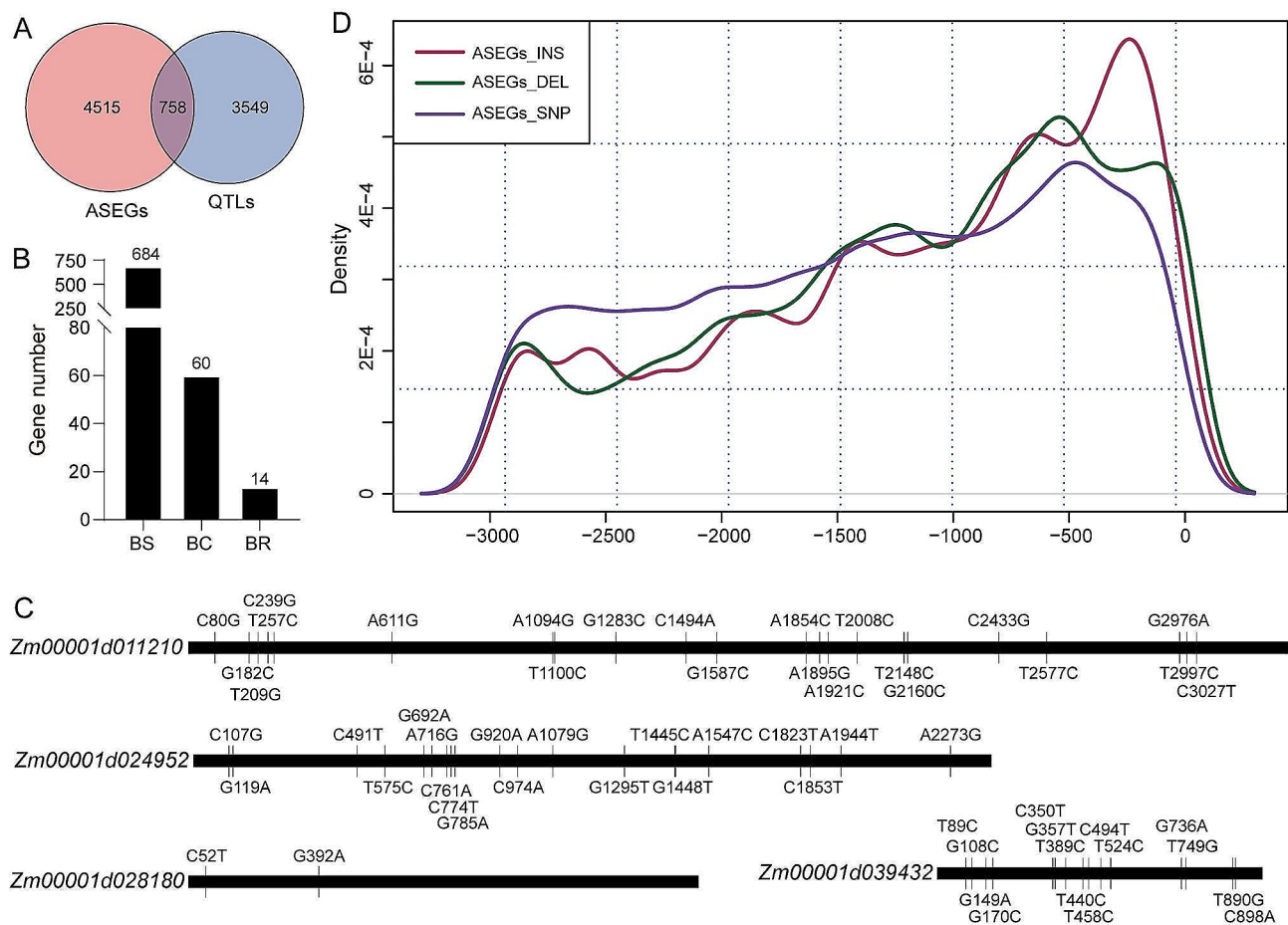


Fig. 5 Variations in alleles between parents. **(A)** Venn diagram presenting ASEGs and heterosis-related QTLs. **(B)** Classification of heterosis-related ASEGs. **(C)** Sequence variations in the ASEG coding sequence between B73 and Mo17. B73 served as the reference. **(D)** Differences in the promoter regions of 758 ASEGs between B73 and Mo17. DEL, deletion; INS, insertion

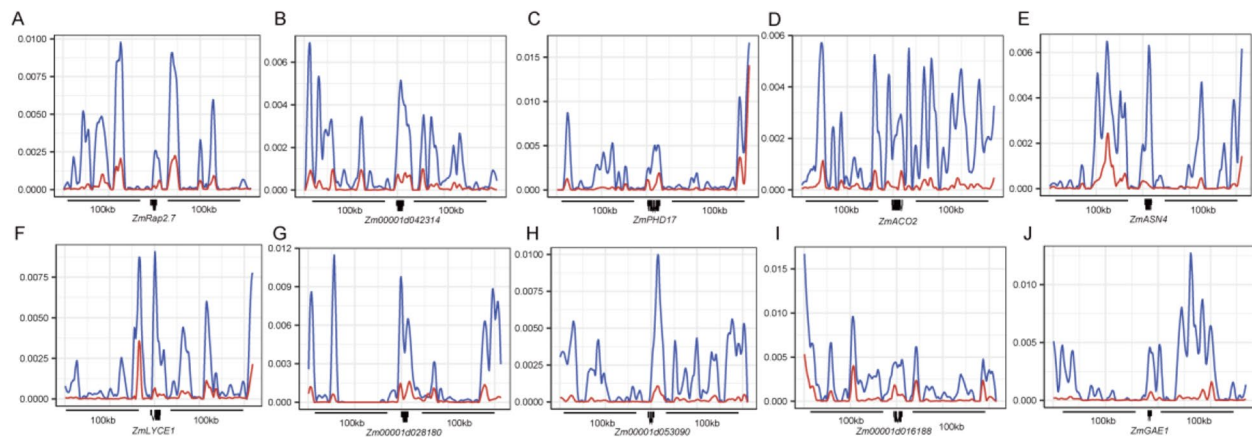


Fig. 6 Selective sweep signals of 10 heterosis-related ASEGs. Blue and red lines represent the average nucleotide diversity of heterosis-related ASEGs based on 23 teosinte and 200 maize sequences

participating in heterosis formation were subjected to artificial selection during maize domestication.

Discussion

ASEGs adapt to different light environments through various expression patterns

ASEGs, varying in number and function among different organisms and conditions, play a crucial role in adapting to growth stages [17, 45, 46]. Maize, a typical C₄ crop, is sensitive to light. Low red/far-red or low blue light causes the inactivation of phytochromes or cryptochromes, leading to shade avoidance syndrome [47–49]. However, the effects of different light conditions on maize ASEGs is less explored. In this study, the maize transcriptomes of reciprocal crosses (BM and MB) were analyzed under dark, far-red, red, and blue light conditions. The results revealed both the light-general and light-specific expression patterns of ASEGs.

Previous studies have suggested that single additive or dominant mode is the main expression pattern in hybrids [50–52]. In this study, we classified ASEGs on the basis of their expression patterns between parents and hybrids, and determined that BC ASEGs primarily exhibited additive expression (30.45–50.32%). Biased B73 BC ASEGs were mainly involved in ribosome-related functions, whereas biased Mo17 BC ASEGs were associated with abiotic stress responses and organic compound biosynthesis and metabolism. In BS ASEGs (4754/5273, 90.16%), dominant (18.09–42.77%) and additive (15.27–38.88%) expression patterns accounted for a significant proportion. Biased B73 and Mo17 BS ASEGs were primarily associated with cell components and membrane composition, respectively (Fig. 4B and C). Therefore, ASEGs adapt to different light environments through multiple complementary expression patterns, which enhances the robustness of hybrid.

Promoter variations induce genotype-dependent ASE formation

ASE differences exist between reciprocal hybrids, a phenomenon observed in rice, maize, and *Arabidopsis* [53–55]. This phenomenon is believed to be caused by genes in the maternal mitochondria and chloroplasts [56]. Other studies have indicated a linear correlation between ASE and genotype during maize seed germination and grain development in reciprocal crosses [57–59]. In our study, reciprocal maize seedlings exhibited a highly genotype-dependent ASE ($R^2=0.89$) (Fig. 1A, Fig. S3A), indicating its prevalence during the growth and development process of maize hybrids.

Transcription factors regulate gene expression by binding to functional elements in promoter regions. The identification of *ZmBZR1* binding sites through ChIP-seq revealed that motif variants contribute to ASE in F₁ [60]. Transposon insertions in promoter regions account for 35.4% of *cis*-regulation in maize hybrids [61]. Furthermore, ASE caused by *cis*-regulation significantly differed between high- and low-altitude maize populations [62]. Our comprehensive analysis revealed a high frequency of variations in the promoters of heterosis-related ASEGs, particularly within the upstream 1,000 bp of promoter regions (Fig. 5D). The use of CRISPR-Cas9 editing technology to modify the *CLV3/ESR-RELATED (CLE)* promoter changes the expression of network genes, resulting in an increase in maize yield [63]. Thus, understanding the genetic basis of genotype-dependent promoter variations in ASEGs can inform breeding strategies aimed at enhancing maize yield.

Heterosis-related ASEGs undergo domestication selection

Teosinte possesses abundant genetic diversity and numerous beneficial genes with substantial breeding value. Favorable alleles in teosinte are domesticated into

specific alleles in different subgroups through gene infiltration and artificial selection, laying the foundation for heterosis formation [8, 55, 64, 65]. Our analysis demonstrated the significant enrichment of 758 ASEGs in regions associated with heterosis-related genes, with 10 heterosis-related BS ASEGs being significantly domesticated (Fig. 6). Surprisingly, 80 and 229 ASEGs in this study were consistent with heterosis-related genes in buds and seeds [58, 59], respectively. However, no significant enrichment region associated with heterosis-related genes was observed, which could be attributed to the single environmental condition. Future efforts could enrich maize resources and enhance hybrid vigor by analyzing ASEG characteristics and selecting specific alleles.

Conclusion

This study comprehensively identified ASEGs under different light conditions and analyzed the expression and functional characteristics of three types of ASEGs (BC, BS, and BR). Among them, 395 BC ASEGs maintained basic metabolism in hybrids in a functionally complementary manner. In addition, 124 BR ASEGs contributed to environmental adaptability, whereas 4,757 BS ASEGs ensured cellular stability through both dominant and additive expression. Moreover, 10 ASEGs associated with the heterosis formation have undergone domestication, with variations in promoter regions being one of the reasons for ASEG formation. These findings offer valuable insights for future maize breeding and selection strategies.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10395-y>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

Not applicable.

Author contributions

Yanpei Zhang and Jianping Yang designed the study. Weimin Zhan, Lianhua Cui, Shuling Yang and Kangni Zhang analyzed the data. Weimin Zhan, Lianhua Cui, Yanpei Zhang and Jianping Yang wrote and revised the manuscript. All authors read and approved the final manuscript.

Funding

This research was supported by the Major Science and Technology Project of the Science and Technology of the Xizang Autonomous Region (XZKJT202405), Lhasa Regional Science and Technology Collaborative Innovation Project in 2022 (QYXTZX-LS2022-01), the National Natural Science Foundation of China (31871709), Henan Province Joint Fund for Science and Technology Research (232103810006), and the Startup Grant of Henan Agricultural University (30500823).

Data availability

Transcriptome raw data in this study was downloaded from the National Center for Biotechnology Information Sequence Read Archive (SRA) database (<https://www.ncbi.nlm.nih.gov/sra>; accession no. PRJNA780806).

Declarations

Ethics approval and consent to participate

We declare that all experimental research comply with relevant institutional, national, and international guidelines and legislation. All experimental protocols were approved by the Academic Committee of the College of Agriculture, Henan Agricultural University.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 29 March 2024 / Accepted: 8 May 2024

Published online: 14 May 2024

References

- Dyer WTT. The effects of cross and self-fertilisation in the vegetable kingdom. *Nature*. 1877;15:329–32.
- Jones DF. Dominance of linked factors as a means of accounting for heterosis. *Proc Natl Acad Sci U S A*. 1917;3:310–2.
- Crow JF. Alternative hypotheses of hybrid vigor. *Genetics*. 1948;33:477–87.
- Powers L. Relative yields of inbred lines and F1 hybrids of tomato. *Bot Gaz*. 1945;106:247–68.
- Liu W, Zhang Y, He H, He G, Deng XW. From hybrid genomes to heterotic trait output: challenges and opportunities. *Curr Opin Plant Biol*. 2022;66:102193.
- Wang B, Lin Z, Li X, Zhao Y, Zhao B, Wu G, et al. Genome-wide selection and genetic improvement during modern maize breeding. *Nat Genet*. 2020;52:565–71.
- Wagner MR, Tang C, Salvato F, Clouse KM, Bartlett A, Vintila S, et al. Microbe-dependent heterosis in maize. *Proc Natl Acad Sci*. 2021;118:e2021965118.
- Li C, Guan H, Jing X, Li Y, Wang B, Li Y, et al. Genomic insights into historical improvement of heterotic groups during modern hybrid maize breeding. *Nat Plants*. 2022;8:750–63.
- Springer NM, Stupar RM. Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. *Plant Cell*. 2007;19:2391–402.
- Aguilar-Rangel MR, Chávez Montes RA, González-Segovia E, Ross-Ibarra J, Simpson JK, Sawers RJH. Allele specific expression analysis identifies regulatory variation associated with stress-related genes in the Mexican highland maize landrace palomero toluqueño. *PeerJ*. 2017;5:e3737.
- Wang Y, Gao S, Zhao Y, Chen W, Shao J, Wang N, et al. Allele-specific expression and alternative splicing in horse x donkey and cattle x yak hybrids. *Zool Res*. 2019;40:293–304.
- Li D, Lu X, Zhu Y, Pan J, Zhou S, Zhang X, et al. The multi-omics basis of potato heterosis. *J Integr Plant Biol*. 2022;64:671–87.
- Tian Y, Thrimawithana A, Ding T, Guo J, Gleave A, Chagné D, et al. Transposon insertions regulate genome-wide allele-specific expression and underpin flower colour variations in apple (*malus* spp). *Plant Biotechnol J*. 2022;20:1285–97.
- Cleary S, Seoighe C. Perspectives on allele-specific expression. *Annu Rev Biomed Data Sci*. 2021;4:101–22.
- Pierre CLS, Macias-Velasco JF, Wayhart JP, Yin L, Semenkovich CF, Lawson HA. Genetic, epigenetic, and environmental mechanisms govern allele-specific gene expression. *Genome Res*. 2022;gr.276193.121.
- Batista RA, Köhler C. Genomic imprinting in plants—revisiting existing models. *Genes Dev*. 2020;34:24–36.
- Shao L, Xing F, Xu C, Zhang Q, Che J, Wang X, et al. Patterns of genome-wide allele-specific expression in hybrid rice and the implications on the genetic basis of heterosis. *Proc Natl Acad Sci*. 2019;116:5653–8.
- Ma X, Xing F, Jia Q, Zhang Q, Hu T, Wu B, et al. Parental variation in CHG methylation is associated with allelic-specific expression in elite hybrid rice. *Plant Physiol*. 2021;186:1025–41.

19. Chory J, Chatterjee M, Cook RK, Elich T, Fankhauser C, Li J, et al. From seed germination to flowering, light controls plant development via the pigment phytochrome. *Proc Natl Acad Sci*. 1996;93:12066–71.
20. Sullivan JA, Deng XW. From seed to seed: the role of photoreceptors in *Arabidopsis* development. *Dev Biol*. 2003;260:289–97.
21. Bae G, Choi G. Decoding of light signals by plant phytochromes and their interacting proteins. *Annu Rev Plant Biol*. 2008;59:281–311.
22. Zhan W, Guo G, Cui L, Rashid MAR, Jiang L, Sun G, et al. Combined transcriptome and metabolome analysis reveals the effects of light quality on maize hybrids. *BMC Plant Biol*. 2023;23:41.
23. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
24. Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, et al. Improved maize reference genome with single-molecule technologies. *Nature*. 2017;546:524–7.
25. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37:907–15.
26. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078–9.
27. Brouard J-S, Schenkel F, Marete A, Bissonnette N. The GATK joint genotyping workflow is appropriate for calling variants in RNA-seq experiments. *J Anim Sci Biotechnol*. 2019;10:44.
28. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27:2156–8.
29. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21.
30. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms. SnpEff Fly (Austin). 2012;6:80–92.
31. Castel SE, Mohammadi P, Chung WK, Shen Y, Lappalainen T. Rare variant phasing and haplotypic expression from RNA sequencing with phASER. *Nat Commun*. 2016;7:12817.
32. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010;11:R106.
33. Shen Y, Sun S, Hua S, Shen E, Ye C-Y, Cai D, et al. Analysis of transcriptional and epigenetic changes in hybrid vigor of allopolyploid *Brassica napus* uncovers key roles for small RNAs. *Plant J*. 2017;91:874–93.
34. Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, et al. AgriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res*. 2017;45:W122–9.
35. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. ClusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innov Camb Mass*. 2021;2:100141.
36. Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol*. 2018;14:e1005944.
37. Sun S, Zhou Y, Chen J, Shi J, Zhao H, Zhao H, et al. Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes. *Nat Genet*. 2018;50:1289–95.
38. Bukowski R, Guo X, Lu Y, Zou C, He B, Rong Z, et al. Construction of the third-generation *Zea mays* haplotype map. *GigaScience*. 2018;7:gix134.
39. Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol*. 2014;31:1929–36.
40. Li C, Xie W, Bai W, Li Z, Zhao Y, Liu H. Calmodulin binds to maize lipid transfer protein and modulates its lipids binding ability. *FEBS J*. 2008;275:5298–308.
41. Du Q, Yang J, Syed muhammad sadiq S, Yang R, Yu J, Li W. Comparative transcriptome analysis of different nitrogen responses in low-nitrogen sensitive and tolerant maize genotypes. *J Integr Agric*. 2021;20:2043–55.
42. Watkins KP, Williams-Carrier R, Chotewutmontri P, Friso G, Teubner M, Belcher S, et al. Exploring the proteome associated with the mRNA encoding the D1 reaction center protein of photosystem II in plant chloroplasts. *Plant J*. 2020;102:369–82.
43. Xiao Y, Jiang S, Cheng Q, Wang X, Yan J, Zhang R, et al. The genetic mechanism of heterosis utilization in maize improvement. *Genome Biol*. 2021;22:148.
44. Wang B, Hou M, Shi J, Ku L, Song W, Li C et al. De novo genome assembly and analyses of 12 founder inbred lines provide insights into maize heterosis. *Nat Genet*. 2023;1–12.
45. Fujimoto R, Taylor JM, Shirasawa S, Peacock WJ, Dennis ES. Heterosis of *Arabidopsis* hybrids between C24 and col is associated with increased photosynthesis capacity. *Proc Natl Acad Sci*. 2012;109:7109–14.
46. Wu J, Sun D, Zhao Q, Yong H, Zhang D, Hao Z, et al. Transcriptome reveals allele contribution to heterosis in maize. *Front Plant Sci*. 2021;12:739072.
47. Liu Y, Jafari F, Wang H. Integration of light and hormone signaling pathways in the regulation of plant shade avoidance syndrome. *aBIOTECH*. 2021;2:131–45.
48. Zhao Y, Zhao B, Wu G, Ma X, Wang B, Kong D, et al. Creation of two hyperactive variants of phytochrome B1 for attenuating shade avoidance syndrome in maize. *J Integr Agric*. 2022;21:1253–65.
49. Fan X, Chen S, Wu W, Song M, Sun G, Yao S, et al. Maize cryptochromes 1a1 and 1a2 promote seedling photomorphogenesis and shade resistance in *Zea mays* and *Arabidopsis*. *Crop J*. 2023;11:1192–203.
50. Stupar RM, Springer NM. cis-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. *Genetics*. 2006;173:2199–210.
51. Wang Z, Xue Z, Wang T. Differential analysis of proteomes and metabolomes reveals additively balanced networking for metabolism in maize heterosis. *J Proteome Res*. 2014;13:3987–4001.
52. Zhao Y, Hu F, Zhang X, Wei Q, Dong J, Bo C, et al. Comparative transcriptome analysis reveals important roles of nonadditive genes in maize hybrid an'nong 591 under heat stress. *BMC Plant Biol*. 2019;19:273.
53. Dapp M, Reinders J, Bédiee A, Balsera C, Bucher E, Theiler G, et al. Heterosis and inbreeding depression of epigenetic *Arabidopsis* hybrids. *Nat Plants*. 2015;1:1–8.
54. Santos JF, Dirk LMA, Downie AB, Sanches MFG, Vieira RD. Reciprocal effect of parental lines on the physiological potential and seed composition of corn hybrid seeds. *Seed Sci Res*. 2017;27:206–16.
55. Lin Z, Qin P, Zhang X, Fu C, Deng H, Fu X, et al. Divergent selection and genetic introgression shape the genome landscape of heterosis in hybrid rice. *Proc Natl Acad Sci*. 2020;117:4623–31.
56. Kamphorst SH, Amaral Júnior AT, do, Vergara-Diaz O, Gracia-Romero A, Fernandez-Gallego JA, Chang-Espino MC, et al. Heterosis and reciprocal effects for physiological and morphological traits of popcorn plants under different water conditions. *Agric Water Manag*. 2022;261:107371.
57. Song G, Guo Z, Liu Z, Cheng Q, Qu X, Chen R, et al. Global RNA sequencing reveals that genotype-dependent allele-specific expression contributes to differential expression in rice F1 hybrids. *BMC Plant Biol*. 2013;13:221.
58. Wan J, Wang Q, Zhao J, Zhang X, Guo Z, Hu D, et al. Gene expression variation explains maize seed germination heterosis. *BMC Plant Biol*. 2022;22:1–12.
59. Dong X, Luo H, Yao J, Guo Q, Yu S, Zhang X, et al. Characterization of genes that exhibit genotype-dependent allele-specific expression and its implications for the development of maize kernel. *Int J Mol Sci*. 2023;24:4766.
60. Hartwig T, Banf M, Prietsch GP, Zhu J, Mora-Ramírez I, Schippers JHM, et al. Hybrid allele-specific ChIP-seq analysis identifies variation in brassinosteroid-responsive transcription factor binding linked to traits in maize. *Genome Biol*. 2023;24:108.
61. Li Z, Zhou P, Della Coletta R, Zhang T, Brohammer AB, O'Connor H. Single-parent expression drives dynamic gene expression complementation in maize hybrids. *Plant J*. 2021;105:93–107.
62. Hu H, Crow T, Nojoomi S, Schulz AJ, Estévez-Palmas JM, Hufford MB, et al. Allele-specific expression reveals multiple paths to highland adaptation in maize. *Mol Biol Evol*. 2022;39:msac239.
63. Liu L, Gallagher J, Arevalo ED, Chen R, Skopelitis T, Wu Q, et al. Enhancing grain-yield-related traits by CRISPR-Cas9 promoter editing of maize *CLE* genes. *Nat Plants*. 2021;7:287–94.
64. Wang Q, Liao Z, Zhu C, Gou X, Liu Y, Xie W, et al. Teosinte confers specific alleles and yield potential to maize improvement. *Theor Appl Genet*. 2022;135:3545–62.
65. Yang N, Wang Y, Liu X, Jin M, Vallebuena-Estrada M, Calfee E, et al. Two teosintes made modern maize. *Science*. 2023;382:eadg8940.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.