Transcriptome-Wide Analysis of RNA \( \text{m}^6\text{A} \) Methylation and Gene Expression Changes Among Two Arabidopsis Ecotypes and Their Reciprocal Hybrids

Zhihui Xu\(^1,2\), Xiaobo Shi\(^2\), Mengmei Bao\(^2\), Xiaoqian Song\(^2\), Yuxia Zhang\(^2\), Haiyan Wang\(^2\), Hairong Xie\(^2\), Fei Mao\(^2\), Shuai Wang\(^2\), Hongmei Jin\(^3\), Suomeng Dong\(^4\), Feng Zhang\(^4\), Zhe Wu\(^5\) and Yufeng Wu\(^2\)*

\(^1\) College of Life Science, Nanjing Agricultural University, Nanjing, China, \(^2\) State Key Laboratory for Crop Genetics and Germplasm Enhancement, Jiangsu Key Laboratory for Information Agriculture, Bioinformatics Center, Academy for Advanced Interdisciplinary Studies, Nanjing Agricultural University, Nanjing, China, \(^3\) Institute of Agricultural Resources and Environment, Jiangsu Academy of Agricultural Sciences, Nanjing, China, \(^4\) College of Plant Protection, Nanjing Agricultural University, Nanjing, China, \(^5\) Department of Biology, SUSTech-PKU Institute of Plant and Food Science, Southern University of Science and Technology, Shenzhen, China

The remodeling of transcriptome, epigenome, proteome, and metabolome in hybrids plays an important role in heterosis. N(6)-methyladenosine (\( \text{m}^6\text{A} \)) methylation is the most abundant type of post-transcriptional modification for mRNAs, but the pattern of inheritance from parents to hybrids and potential impact on heterosis are largely unknown. We constructed transcriptome-wide mRNA \( \text{m}^6\text{A} \) methylation maps of Arabidopsis thaliana Col-0 and Landsberg erecta (Ler) and their reciprocal F\(_1\) hybrids. Generally, the transcriptome-wide pattern of \( \text{m}^6\text{A} \) methylation tends to be conserved between accessions. Approximately 74\% of \( \text{m}^6\text{A} \) methylation peaks are consistent between the parents and hybrids, indicating that a majority of the \( \text{m}^6\text{A} \) methylation is maintained after hybridization. We found a significant association between differential expression and differential \( \text{m}^6\text{A} \) modification, and between non-additive expression and non-additive methylation on the same gene. The overall RNA \( \text{m}^6\text{A} \) level between Col-0 and Ler is clearly different but tended to disappear at the allelic sites in the hybrids. Interestingly, many enriched biological functions of genes with differential \( \text{m}^6\text{A} \) modification between parents and hybrids are also conserved, including many heterosis-related genes involved in biosynthetic processes of starch. Collectively, our study revealed the overall pattern of inheritance of mRNA \( \text{m}^6\text{A} \) modifications from parents to hybrids and a potential new layer of regulatory mechanisms related to heterosis formation.

**Keywords:** RNA \( \text{m}^6\text{A} \) methylation, hybrid, heterosis, Arabidopsis, RNA modification dynamics
SIGNIFICANCE STATEMENT

The reprogramming and corresponding effect of mRNA m<sup>6</sup>A methylation on hybrids remain highly unknown. We demonstrated the pattern of conserved inheritance of m<sup>6</sup>A methylation from parents to hybrids and the potential impact on heterosis formation, uncovering mRNA m<sup>6</sup>A methylation as a new layer of regulatory mechanisms in the formation of hybrid vigor.

INTRODUCTION

Heterosis refers to the increased performance of hybrid offspring relative to their parents in many traits, such as growth rate and biomass (Birchler et al., 2003, 2010; Hochholdinger and Hoecker, 2007; Chen, 2010; Birchler, 2015). Both genetic and epigenetic mechanisms are thought to be involved in heterosis (Chen, 2013). Epigenetic changes have been found to impact hybrid vigor (Cubas et al., 1999; Manning et al., 2006; Shindo et al., 2006; Ni et al., 2009; He et al., 2010). DNA methylation level is altered by trans-chromosomal methylation (TCM) and trans-chromosomal demethylation (TCDM) (Greaves et al., 2014), which changes the overall DNA methylation level in the F<sub>1</sub> hybrids, especially in regions that are differentially methylated in two parents (Shen et al., 2012). Histone modification patterns in hybrids of rice or maize have shown correlations between altered gene expression and changes in histone marks compared with the parents (He et al., 2010, 2013; Lv et al., 2019). In Arabidopsis hybrids, global histone modifications of the parents are largely transmitted to the F<sub>1</sub> generation (Moghaddam et al., 2011; Dong et al., 2012; Yang et al., 2016). DNA methylation and histone modifications are altered at many loci, such as circadian clock associated1 (CCA1) and late elongated hypocotyl (LHY), which are associated with growth vigor in Arabidopsis F<sub>1</sub> hybrids (Ni et al., 2009; Shen et al., 2012).

Recently, chemical modifications of mRNAs, such as N(6)-methyladenosine (m<sup>6</sup>A), N(1)-methyladenosine (m<sup>2</sup>A), and 5-methylcytosine (m<sup>5</sup>C), have emerged as an additional level of transcript regulation (Dominissini et al., 2012, 2016; Meyer et al., 2012; Li et al., 2016, 2017). m<sup>6</sup>A methylation is the most abundant type of modification for mRNAs, occurring in more than one-third of mammalian transcripts and half of the plant transcripts (Dominissini et al., 2012; Meyer et al., 2012; Li et al., 2014; Luo et al., 2014; Luo et al., 2020; Wan et al., 2015; Zhou et al., 2019; Miao et al., 2020). The m<sup>6</sup>A modification is reversible and dynamic, with m<sup>6</sup>A demethylase acting as an eraser and methyltransferase acting as a writer (Jia et al., 2011; Meyer and Jaffrey, 2017). Recognition of these dynamic m<sup>6</sup>A modifications by YTH domain-containing proteins leads to a broad range of functions associated with the change in mRNA stability, cap-independent translation, splicing, translation efficiency, and mRNA structure (Dominissini et al., 2012; Meyer et al., 2015; Meyer and Jaffrey, 2017), but the location of m<sup>6</sup>A in mRNA determines different functions (Gilbert et al., 2016). In the 5′ UTR, m<sup>6</sup>A participates in mRNA cap-independent translation by directly binding to eukaryotic initiation factor 3 (eIF3) and then recruiting the 40S ribosomal subunit to initiate translation (Meyer et al., 2015). On the other hand, m<sup>6</sup>A in the 3′ UTR has been reported to have several functions, such as promoting translation by binding with METTL3 and eIF3h to facilitate formation of the translation loop (Choe et al., 2018), regulating mRNA lifetime by binding with YTHDF2, which relocates transcripts to the P-body (Wang et al., 2014), and changing mRNA structure to affect RNA-protein interactions (Liu et al., 2015).

In Arabidopsis thaliana, m<sup>6</sup>A is essential in embryo development (Zhong et al., 2008). Further research revealed that m<sup>6</sup>A is also essential in post-embryonic development (Bodi et al., 2012), for example, for normal trichome morphology and correct timing of leaf formation (Arribas-Hernandez et al., 2018; Scutenaire et al., 2018; Wei et al., 2018), partly because it regulates the expression of key shoot meristem genes to control shoot apical meristem (SAM) proliferation (Shen et al., 2016). Transcriptome-wide mapping of m<sup>6</sup>A in Arabidopsis wild-type (WT) and related mutants indicated a complex relationship between m<sup>6</sup>A modifications and gene expression. Lack of FKBP12 interacting protein 37 (FIP37), a component of the methyltransferase complex in Arabidopsis, results in a dramatically reduced abundance of m<sup>6</sup>A, as most transcripts bearing m<sup>6</sup>A in WT are decreased in the mutant (Shen et al., 2016). In addition, further study showed that m<sup>6</sup>A inhibits mRNA degradation through inhibition of site-specific cleavage (Anderson et al., 2018). Nevertheless, it was reported that the highly expressed transcripts had fewer m<sup>6</sup>A modifications, as revealed by transcriptome-wide m<sup>6</sup>A patterns in Arabidopsis (Wan et al., 2015). Although m<sup>6</sup>A abundance varies among Arabidopsis accessions and affects transcript abundance, how m<sup>6</sup>A changes in F<sub>1</sub> hybrids relative to their parents and its potential role in determining F<sub>1</sub> hybrid vigor have not been clarified.

In this study, we selected two Arabidopsis ecotypes, namely, Col-0 and Landsberg erecta (Ler), and their F<sub>1</sub> reciprocal hybrids, to investigate the potential effect of m<sup>6</sup>A on heterosis. We identified the distribution pattern and the intensity change in m<sup>6</sup>A in Col-0, Ler, and their F<sub>1</sub> reciprocal hybrids. We showed that the peaks and distribution features of m<sup>6</sup>A methylation are highly conserved between accessions. Although changes in m<sup>6</sup>A intensity and transcript abundance within accessions are weakly positively correlated, upregulation of m<sup>6</sup>A between accessions tends to be associated with a downregulated abundance of mRNA and vice versa. We found that the overall m<sup>6</sup>A difference between the parents is attenuated at allelic sites in the hybrids, and that there is a negative correlation between the expression and corresponding m<sup>6</sup>A intensity of allelic genes. Interestingly, even though hundreds of m<sup>6</sup>A peaks are changed between the parents and hybrids, many biological functions of the corresponding genes are consistently affected, including the biosynthetic processes of starch, which have been reported to be associated with growth vigor. The data, therefore, suggest the overall pattern of mRNA m<sup>6</sup>A remodeling in hybrids, which may contribute to heterosis formation.
RESULTS
Transcriptome-Wide Profile of m^6^A Methylation Among Col-0, Ler, and Their F\textsubscript{1} Reciprocal Hybrids

To explore RNA m^6^A abundance variation between the two ecotypes and its alteration in hybrids, we first analyzed transcriptome-wide m^6^A profiles among Col-0, Ler, and their F\textsubscript{1} reciprocal hybrids (Supplementary Figure 1A) by applying N^6^-methyladenosine sequencing (m^6^A-seq) with two biological replicates. Sequencing data of RNA input and immunoprecipitation (RIP) are highly correlated between replicates, indicating the high quality of m^6^A-seq in this study (Supplementary Figures 1B,C). We found that the normalized reads from m^6^A-RIP of all samples are enriched in the 3' UTR of the transcripts (Figure 1A), which is similar to the results of previous research (Luo et al., 2014; Wan et al., 2015). The normalized read depth in Ler is significantly lower than that of the other three samples, suggesting that the overall m^6^A abundance of Ler was lower (Figure 1A). To exclude the possible bias introduced by the reference genome, we performed exact analysis using the Ler reference genome rather than Col-0 and still obtained identical results (Supplementary Figure 2). Interestingly, we did not find low m^6^A abundance in the 3' UTR of the two hybrids, similar to Ler.

To further study global patterns of m^6^A in Col-0, Ler and their hybrids, we identified m^6^A peaks using a transcriptomic peak caller, METPeak (Cui et al., 2016). A total of 13,145, 13,562, 12,956, and 12,542 peaks are detected in Col-0, Ler, F\textsubscript{1}CL, and F\textsubscript{1}LC, respectively (Supplementary Table 1); and these peaks were located in ~9,778, 9,920, 10,066, and 10,017 protein-coding
Variations of m^6A Modification Among the Parental Lines and Hybrids

Previous research has shown that m^6A is highly conserved between two accessions of Arabidopsis, namely, Col-0 and Hen-16 (Luo et al., 2014). We found that 10,584 m^6A peaks (80.5% of Col-0, 78% of Ler; Figure 2A) are common between Col-0 and Ler, and that these peaks are located in 8,302 expressed transcripts (49.4% of the total). In addition, we found that the majority of the m^6A peaks are common among the parental lines and F_1 hybrids. There are 9,641 (74.4% of F_{1CL}) and 9,331 (74.4% of F_{1LC}) m^6A peaks that are common between the parents and the F_1 hybrids, respectively (Figure 2B). These peaks are also located in 7,844 and 7,723 of the expressed transcripts in F_{1CL} and F_{1LC}, respectively (Supplementary Figure 3C). The common peaks (11,000) between F_{1CL} and F_{1LC} account for 85.6–88.5% of the total peaks in F_1 hybrids (Figure 2A), which is slightly higher than that in the two parents. Collectively, these data indicate a more general conservation pattern of RNA m^6A modification among accessions and hybrids of Arabidopsis.

Considering the obvious difference in m^6A levels between Col-0 and Ler, it is necessary to determine whether common m^6A peaks between any two samples are significant differentially methylated areas (DMPs). We established two criteria for DMPs: (1) passed Fisher’s exact test after multiple comparison corrections (FDR < 0.05); (2) the difference in peak enrichment between any two samples was larger than a 1.5-fold change. Eventually, we identified 1,776 DMPs (16.8% of the common peaks) between the parents, among which the intensity of 1,721 (16.3%) peaks, as expected, is higher in Col-0 (Figures 2C,D; Supplementary Table 2). For the comparison between F_1 reciprocal hybrids, we found only 2 DMPs (0.02%), suggesting that paternal or maternal effects on the level of m^6A modifications are weak in Arabidopsis (Figure 2C). For the m^6A peaks shared between the parents and F_{1CL} or between the parents and F_{1LC}, we identified 315, 479, 477, and 1,273 DMPs, respectively (Figure 2C). Taken together, the intensity of common m^6A peaks tends to be conserved between accessions or during inheritance from parents to hybrids.

Relationship Between Transcript Abundance and m^6A Modification Level

Multiple recent studies have indicated complex functions of m^6A in transcription regulation with the ability to stabilize (Luo et al., 2014; Anderson et al., 2018) or destabilize mRNAs in Arabidopsis (Wan et al., 2015). We analyzed the relationships of transcript abundance and the corresponding m^6A levels. We found a weak positive correlation between the expression abundance and intensity of m^6A modification on one gene within each accession (Figure 3A). Overall, the genes with m^6A modification show significantly higher expression than non-m^6A-containing genes (Figure 3B). In addition, more than 60% of the expressed genes are associated with at least one m^6A peak (Supplementary Figure 3A, Supplementary Table 1).

Next, we investigated the relationship between changes in m^6A methylation and transcript abundance in the parent lines and their F_1 reciprocal hybrids. We first identified differentially expressed genes (DEGs) between the lines (Supplementary Figures 4A,B) and checked the overlap between DEGs and DMPs. We found that the proportion of DEGs associated with DMPs is significantly higher than that of non-DEGs (Figure 3C). Even so, only 3.29–13.26% of the DEGs are associated with DMPs between the parent lines and hybrids (Figure 3C). Taken together, these results indicated that changes in m^6A intensity on transcripts tend to be associated with changes in abundance, and that most DEGs are not directly associated with m^6A changes in Arabidopsis.

We then focused on genes with significant changes in both expression and m^6A modification between accessions. Most DMPs showed upregulated m^6A intensity in Col-0 between Col-0/Ler (comparison between Col-0 and Ler), as well as between Col-0/F_{1CL} and between Col-0/F_{1LC} (Figure 3D). A total of 862 DEGs between Col-0/Ler are associated with DMPs upregulated in Col-0, among which there are significantly more downregulated expressed genes than upregulated genes in Col-0 (Figure 3D, \( p = 2.41e^-5 \), chi-square test). A similar pattern is also found in Col-0/F_{1CL} (\( p = 1.7e^-10 \)) and Col-0/F_{1LC} (\( p = 0.037 \), Figure 3D). There are more DMPs showing downregulated m^6A intensity in Ler between Ler/F_{1CL} and between Ler/F_{1LC}, and these DMPs are also associated with more genes with upregulated expression in Ler (Supplementary Figure 4C, \( p = 1.19e^-12 \) for Ler/F_{1CL}; \( p = 2.38e^-20 \) for Ler/F_{1LC}). This result indicates that downregulated DMPs tend to be associated with more upregulated DEGs and vice
versa between accessions of *Arabidopsis*, implying that the complexity of m\(^6\)A function affects the transcript abundance of genes.

### Relationship Between Non-additive Expression and Non-additive m\(^6\)A Modification

We identified 2,758 and 4,123 genes showing non-additive expression in F\(_{1}\)CL and F\(_{1}\)LC, respectively. Similar to gene expression, the inheritance of m\(^6\)A modifications in hybrids can be additive or non-additive. We defined m\(^6\)A peaks with a significant change between enrichment value in hybrid and the average enrichment value of parents (MPV) (FDR < 0.05, see methods for detail) as non-additive m\(^6\)A modified peaks. The majority (95.6 and 95.2\%) of the m\(^6\)A peaks show additive patterns in both hybrids, while only 538 and 563 peaks in F\(_{1}\)CL and F\(_{1}\)LC, respectively, are non-additive (Supplementary Table 3). Moreover, non-additive m\(^6\)A peaks are significantly associated with non-additively expressed genes in both hybrids (Supplementary Table 3, \(p < 2.2\times10^{-16}\) for both F\(_{1}\)CL and F\(_{1}\)LC, chi-square test). We still observed that only 6.53–6.82\% of non-additively expressed genes show a non-additive pattern of m\(^6\)A modification, indicating that m\(^6\)A may play a role in the regulation of non-additive gene expression in *Arabidopsis* hybrids.

### Relationship Between Allelic Gene Expression and Allelic m\(^6\)A Methylation in F\(_{1}\) Hybrids

To analyze the allelic bias in gene expression and m\(^6\)A modifications in hybrids, we identified single-nucleotide polymorphisms (SNPs) between Col-0 and Ler with stringent criteria (see methods) and used these SNPs to determine the reads of RNA-seq or m\(^6\)A-seq generated from the allele of Col-0 or Ler. A total of 76,983 SNPs with high confidence are identified. These SNPs associate with 8,972 and 8,991 genes and with 2,509 and 2,325 m\(^6\)A peaks in F\(_{1}\)CL and F\(_{1}\)LC, respectively, which are used in the following analysis. As expected, we still observed significantly higher m\(^6\)A modification in Col-0 than in Ler on these SNPs (Figure 4A). Nevertheless, this bias tends to disappear between the two parental alleles in the hybrids. The log-transformed mean value of the m\(^6\)A ratio between the two allelic SNPs is close to zero, and the majority of the ratio (94.1\% for F\(_{1}\)CL and 93.6\% for F\(_{1}\)LC) falls within the interval (−1, 1) in both hybrids (Figure 4A), indicating that the overall m\(^6\)A difference between the parents is attenuated at allelic sites in the hybrids. The pattern of attenuation is not observed for the expression of allelic genes (Figure 4B). We have identified only
four and seven peaks showing significant allele-specific RNA m^6^A methylation (FDR < 0.05, see Methods) in the hybrids, implying extremely rare allele bias of RNA m^6^A methylation after the combination of the two parental genomes. Despite the smaller difference in m^6^A abundance between the alleles, the correlation between the allelic abundance of mRNA and the
corresponding allelic intensity of m$^6$A methylation is negative (Figure 4C). This result is consistent with the relationship between DEGs and DMPs.

**Biological Function of Genes Associated With Significant Changes in m$^6$A**

F$_1$ hybrids crossed by ecotypes of *Arabidopsis*, as well as Col-0 × Ler (Groszmann et al., 2014), showed clear growth vigor (Supplementary Figure 1), but the relationship between heterosis and changes in m$^6$A abundance between the parent lines and hybrids was unknown. We first focused on the function of genes showing significantly differential m$^6$A methylation (Supplementary Table 4), which were referred to as differentially m$^6$A-modified genes (DMGs). We identified 462 enriched GO terms of DMGs generated from the comparisons between the parents and hybrids (Supplementary Table 5). Interestingly, the enriched GO terms of DMGs between the parents and hybrids tend to be consistent. For instance, we found 319 enriched GO terms of DMGs between Col-0/F$_{1}$CL, among which 231–267 (72–84%) are also identified in the Col-0/F$_{1}$LC, Ler/F$_{1}$CL, and Ler/F$_{1}$LC comparisons. These data implied that there is clear heterogeneity of biological functions affected by differential m$^6$A modification between Col/Ler (between-parent difference) and between parent/hybrid (parent-hybrid difference). We kept the enriched GO terms of DMGs from the parent-hybrid comparison but not from the between-parent comparison and found a clear trend of enriched biological functions, such as biosynthetic and metabolic processes of multiple carbohydrates, secondary metabolic processes, and development of shoot, root hair, and so on (Figure 5A, Supplementary Table 4), among which starch biosynthetic process was reported to be involved in heterosis (Chen, 2013).

Hybrid vigor has been found to be related to changes in transcription, epigenetic modifications, and protein abundance (Chen, 2013). Considering that m$^6$A is involved in multiple biological processes related to RNA fate at the post-transcriptional level, it is worthwhile to focus on the function of genes showing differential m$^6$A modification without changes in gene expression. We found that the enriched GO terms of DMGs and not DEGs between parents/hybrids are associated with membrane- or chloroplast-located proteins, transport, or the proteasome complex (Supplementary Figure 5A). This pattern is clearly different from the enriched GO terms of the DEGs but not DMGs between parent/hybrid, which are associated with stress response genes, mitochondria-located genes, etc (Supplementary Figure 5B). These data implied that m$^6$A modification could be involved in the formation of F$_1$ hybrid vigor through post-transcriptional regulation of mRNA without changing the abundance.

Several genes involved in starch and carbohydrate metabolism promote growth and biomass vigor in *Arabidopsis* (Chen, 2013), so we focused on the DMGs involved in the starch biosynthetic process (Figure 5B). There are 51 DMGs from all four parent/hybrid comparisons annotated as genes of the starch biosynthetic process (GO: 0019252). We checked the published biological functions against the TAIR database one by one and found 20 genes associated with biomass and growth rate of *Arabidopsis* (Supplementary Table 6). We visualized seven genes of 51 DMGs that are annotated as starch metabolism genes through Mapman (Thimm et al., 2004). Interestingly, six genes are located in chloroplasts and involved in the same pathway, and four of them control biomass and growth rate in *Arabidopsis* based on published results (Figure 5B). Collectively, these data indicated the strong association between changes in m$^6$A methylation and the growth vigor of F$_1$ hybrids.

**DISCUSSION**

Multiple transcriptome-wide maps revealed highly conserved patterns of m$^6$A methylation among *Arabidopsis* accessions...
FIGURE 5 | Enriched biological functions of differentially m^6^A-methylated genes. (A) GO terms of genes associated with DMPs that are enriched only in comparisons of parents/hybrids, e.g., m^6^A-methylated peaks showing significant differences between Col-0 and F_1CL_. Only some enriched GO terms are shown in the figure, and all GO terms are listed in Supplementary Table 4. (B) A diagram for genes with DMPs involved in starch biosynthetic process related to biomass in Arabidopsis. The pathway is modified based on the Mapman database. Only genes in the pathway with DMPs are highlighted, of which m^6^A methylation differences between parents/hybrids are shown in the heatmap. *indicates the genes showing non-additive m^6^A methylation in the hybrids simultaneously.
(Can-0 and Hen-16) or organs (leaf, root, and flower) (Luo et al., 2014; Wan et al., 2015). More than 70% of m6A peaks are shared between Arabidopsis Can-0 and Hen-16 (Luo et al., 2014), similar to the percentage (78%) of peaks shared between Col-0 and Ler in this study. We also found that ∼74% of m6A peaks are shared between the parents and hybrids. In addition, our results indicated that m6A modifications in hybrids are enriched around the 3′ UTR, stop and start codons of transcripts, showing consistent features across accessions and organs (Luo et al., 2014; Wan et al., 2015). Moreover, RNA m6A methylation peaks are also conserved between two inbred lines (B73 and Mo17) of maize (Luo et al., 2020) and two tissues of rice (Li et al., 2014). In summary, these results implied a more general conservation pattern of m6A methylation in plants, which could be related to the fundamental role of m6A methylation in plant development (Zhong et al., 2008; Shen et al., 2016; Anderson et al., 2018). Nevertheless, a recently published study showed that there are much more genes with differentially m6A level or non-additively m6A variation in maize hybrid (B73 × Mo17) compared with the parents (Luo et al., 2021), implying that the pattern of m6A reprogramming in hybrid is related to species or parent lines with different degree of variation.

The effects of m6A modification on gene expression vary among genes. In an Arabidopsis demethylase ALKBH10B loss-of-function mutant, mRNAs of flower development genes, such as FT, SPL3, and SPL9 show increased m6A modification but reduced stability (Duan et al., 2017). Nevertheless, lack of m6A modifications on the mRNA of the WUS and STM genes enhances their stability in the FIP37 mutant line of Arabidopsis (Shen et al., 2016). Additional studies have indicated the biological functions of stabilizing or destabilizing mRNAs in Arabidopsis (Luo et al., 2014; Wan et al., 2015; Anderson et al., 2018). The data also indicated the conflicting functions of m6A in regulating gene expression. Overall, we observed a very weak positive correlation between the abundance of mRNA and the intensity of m6A modification within each of the accessions. However, we also found that mRNAs with significantly decreased methylation of m6A tend to show upregulated expression between accessions or between parents and hybrids. The complex regulatory roles of m6A in transcript abundance might be correlated with its location (Luo et al., 2014), differences between readers (Wei et al., 2018), or the RNA structure dependent on m6A (Liu et al., 2015).

The molecular mechanism of heterosis is quite complex; and omics methods, ranging from transcriptomics to metabolomics, have provided novel insights into the mechanism (Chen, 2013). Changes in epigenetic modifications, such as histone methylation in hybrids, could promote growth by altering gene expression (Ni et al., 2009). As a newly identified reversible modification of RNA, the reprogramming of m6A in hybrids and the corresponding functions related to heterosis remain largely elusive. The data indicated that most of the differentially expressed genes are not associated with differential m6A methylation, and that only a few hundred m6A peaks are significantly changed between parents and hybrids. However, these peaks are associated with many biological functions, of which 20 of 51 starch- and carbohydrate-related genes are confirmed as being associated with biomass vigor in Arabidopsis (Supplementary Table 6). We did not identify the genes showing differential m6A methylation involved in the circadian rhythm regulatory network, for instance, LHY, GI, CCA1, and TOC1, which are also related to biomass vigor in hybrids crossed by two accessions, namely, Col-0 and C24 (Chen, 2013). We propose two possible reasons. One could be that the different molecular bases of heterosis between F1 hybrids are crossed by different ecotypes. The hybrids of C24 × Col and Col × Ler showed differences in growth vigor at various time points of vegetative development (Groszmann et al., 2014). Another reason could be that circadian rhythm-related genes tend to promote growth through the regulation of transcription. We found that some circadian genes, such as GI and TOC1, are differentially expressed between the parents and hybrids, while a considerable number of the 20 genes involved in the starch biosynthetic process showed only differential m6A methylation rather than differential expression (Supplementary Table 7). Since m6A controls RNA fate-related processes, such as mRNA stability, transport, or translation (Dominissini et al., 2012; Meyer et al., 2015; Meyer and Jaffrey, 2017), this study indicates a new layer of regulatory mechanisms contributing to heterosis at the post-transcriptional level in Arabidopsis.

**Experimental Procedures**

**Plant Materials and Growth Conditions**

Plant materials included two Arabidopsis accessions (Col-0, Ler) and their F1 reciprocal hybrids. F1 seeds were produced by hand pollination between Col-0 and Ler. Seeds were sown on soil, stratified at 4°C for 3 days to synchronize germination. Plants were then shifted into greenhouse and grown under a long-day condition (16 h in light and 8 h in dark) at 22°C for 21 days. Above-ground tissues were harvested and stored at −80°C for the following experiments.

**MeRIP Libraries Construction and Sequencing**

MeRIP libraries preparation mainly followed a published procedure (Dominissini et al., 2013). Briefly, total RNA was extracted from leaves in 50 mL conicals using TRIzol (15596018, Ambion, Austin, TX, United States). Poly(A) RNA was enriched (MRN10, Sigma-Aldrich, Saint Louis, MO, United States) and fragmented into ∼100 nt by fragmentation reagent (AM8740, Invitrogen, Carlsbad, CA, United States) for 15 min at 70°C. Few microliters of fragmented RNA was saved as input control, and the left was incubated with m6A antibody (202003, Synaptic Systems, Goettingen, Germany), in 1x IP buffer supplemented with RNasin Plus (N2611, Promega, Madison, WI, United States) for 4 h at 4°C. The antibody-bound RNA was then incubated with pre-blocked protein A beads (10001D, Invitrogen, Carlsbad, CA, United States) at 4°C for 2 h. The immunoprecipitated RNA was released using an elution buffer (1x IP buffer supplemented with 6.7 mM N6-methyladenosine, M2780, Sigma-Aldrich, Saint Louis, MO, United States). Input and IP libraries were constructed using NEBNext Ultra RNA Library Prep Kit for Illumina (E7645S, NEB, Ipswich, MA, United States) and subjected to sequencing on the Illumina Hiseq X-10 platform.
Reads Pre-processing and Alignment
Raw reads of input and IP samples were processed by trim-galore (version 0.4.1) to remove adaptors and low quality reads and then mapped to the Arabidopsis Col-0 reference genome (TAIR 10) using Tophat2 (version 2.1.1) (Kim et al., 2013) with Araport11 annotation in the analyses for parental lines and hybrid lines. We also used Ler reference genome and corresponding annotation (downloaded from NCBI, accession number GCA_001651475.1) to check for possible bias introduced by the reference genome. The parameters were modified (–read-edit-dist 5, -N 5) to obtain more SNP information of Ler and F1 hybrids. Multiple mapped reads were filtered using the SAMtools package (version 1.9) (Li et al., 2009). Only paired unique reads were used for downstream analysis.

N(6)-Methyladenosine Peak Identification and Annotation
McTPeak (Cui et al., 2016), a transcriptomic peak caller, was used to identify m6A peaks. In order to get confidence peaks, we maintained peaks on genes with FPKM ≥ 1. Moreover, to avoid huge differences in the calculation of peak enrichment due to insufficient coverage, we performed a random sampling of genomic regions and calculated reads of all input samples, and high confidence peaks were selected if the peak region satisfied Input FPKM ≥ 5.

To define m6A peak summits, two repeats of input and IP sample were merged, and the coverage of each base of peaks was counted by in-house script (Supplementary Scripts 1–3). The residual was calculated by IP reads subtracted by input reads, and the point with the largest residual was referred to as peak summit. The peak summits were intersected with protein-coding gene sequences, which were integrated into a tiered order—3’UTR, 5’ UTR, and CDS, to determine their locations (Supplementary Script 4). Additionally, m6A peaks were assigned to start codon and stop codon segments, which was 200 nt centered to start codon and stop codon, respectively, to identify the preference of m6A peaks.

Identification of Differentially Methylated Peaks and Additive/Non-Additive Methylated Peaks
The common m6A peaks between any two samples were defined according to whether they intersected with each other. We calculated read counts of IP and input replicates for each m6A peak of every comparison group (Supplementary Script 5). A 2 × 2 contingency table was filled by IP and input normalized reads of samples, respectively. A Fisher’s exact test was performed to identify m6A differentially methylated peaks, and p-value was adjusted by Bonferroni–Holm correction using R scripts. The differentially methylated peaks should satisfy two requirements: (1) p(adj) < 0.05; (2) the difference between any two samples > 1.5.

To classify non-additive and additive methylated peaks, Fisher’s exact test was performed by comparing the input and IP normalized reads of hybrid and the average of parents’ input and IP normalized reads. Only common peaks with p(adj) < 0.05 were considered as non-additive methylated peaks. Otherwise, they were referred to as additively methylated peaks.

Identification of Differentially Expressed Genes and Additive/Non-Additive Expressed Genes
The number of reads for each gene was counted using HTSeq (Anders et al., 2015) with a default setting. R package DESeq2 (version 1.22.2) was used for analyzing differentially expressed genes, and only genes with p(adj) < 0.05 were considered as DEGs. If the expression of genes in hybrids was significantly different from mid-parent value (p(adj) < 0.05), these genes were classified as non-additive expressed genes, and the others were referred to as additive expressed genes.

Gene Ontology Analysis
The gene sets were submitted to agriGO database (Tian et al., 2017) to perform GO enrichment analysis. Functional enrichment was performed using the singular enrichment analysis (SEA) tool and TAIR genome locus (TAIR 10) as background. The GO terms with FDR ≤ 0.01 were considered to be enriched.

Analysis of Allelic Expression and Allelic N(6)-Methyladenosine Enrichment
To obtain confidence SNPs between Col-0 and Ler, the Ler (downloaded from NCBI) and Col-0 reference genomes (TAIR 10) were cut into 100 bp fragments with 1 bp shift, and then mutually mapped to the reference genome. The read counts of each position were called using the SAMTools “mpileup” command with the parameter “-i.” SNPs were first identified if site coverage ≥90X and mutant ratio (mutants/covered reads) ≥ 90%. The input and IP reads of F1CL and F1LCL were separately mapped to the Col-0 reference and the Ler reference, and the reads covered SNPs were calculated. Theoretically, the reads mapped to the corresponding coordinate of the Col-0 and Ler references should be identical, or at least with small bias. Thus, SNPs with severe biased reads (the difference of reads mapped to the corresponding SNPs of two references was more than 10%) were excluded. Additionally, the SNPs that were not homozygous in parent lines were filtered. For allele-specific methylation analysis (Supplementary Script 6), we first calculated reads at SNPs within m6A peaks of IP and input replicates of F1 hybrids, and then filled a 2 × 2 contingency table with normalized reads. A Fisher’s exact test was performed to identify allele-specific methylated peaks, and p-value was adjusted by Bonferroni–Holm correction using R scripts. Peaks with significant allelic methylation difference (FDR < 0.05) were identified as allele-specific peaks.

DATA AVAILABILITY STATEMENT
The original contributions presented in the study are publicly available. All the sequencing data have been deposited in The National Genomics Data Center (NGDC) under accession number CRA003884.

AUTHOR CONTRIBUTIONS
YW conceived the project. ZX, SD, FZ, and ZW designed the experiments. ZX, XBS, MB, YZ, HW, HX, SW, and HJ

Frontiers in Plant Science | www.frontiersin.org 10 June 2021 | Volume 12 | Article 685189
performed the experiments. ZX, XQS, FM, and YW conducted bioinformatics analyses. ZX, SW, and YW wrote the article. All authors contributed to the article and approved the submitted version.

FUNDING
This study was supported by the National Key Research and Development Program of China (2016YFD0101001), Innovative Project of State Key Laboratory for Crop Genetics and Germplasm Enhancement, Fundamental Research Funds for the Central Universities (JCQY201901), Jiangsu Collaborative Innovation Center for Modern Crop Production, and an Innovation and Enterprise Scholar of Jiangsu Province to YW.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.685189/full#supplementary-material

REFERENCES
Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169. doi: 10.1093/bioinformatics/btu638
Anderson, S. J., Kramer, M. C., Gosai, S. J., Yu, X., Vandivier, L. E., Nelson, A. D. L., et al. (2018). N6-methyladenosine inhibits local ribonucleolytic cleavage to stabilize mRNAs in Arabidopsis. Cell Rep. 25, 1146–1157. doi: 10.1016/j.celrep.2018.10.020
Arribas-Hernandez, L., Bressendorff, S., Hansen, M. H., Poulsen, C., Erdmann, S., and Brodersen, P. (2018). An m6A-YTH module controls developmental timing and morphogenesis in Arabidopsis. Plant Cell 30, 952–967. doi: 10.1105/tpc.17.00833
Birchler, J. A. (2015). Heterosis: the genetic basis of hybrid vigor. Nat. Plants 1, 15020. doi: 10.1038/nplants.2015.20
Birchler, J. A., Auger, D. L., and Riddle, N. C. (2003). In search of the molecular basis of heterosis. Plant Cell 15, 2236–2239. doi: 10.1105/tpc.1151030
Birchler, J. A., Yao, H., Chudalayandi, S., Vaiman, D., and Veitia, R. A. (2010). Heterosis. Plant Cell 22, 2105–2112. doi: 10.1105/tpc.110.076133
Bodi, Z., Zhong, S., Mehra, S., Song, J., Graham, N., Li, H., et al. (2012). Adenosine methylation in Arabidopsis mRNAs is associated with the 3’ end and reduced levels cause developmental defects. Front. Plant Sci. 3:48. doi: 10.3389/fpls.2012.00048
Chen, Z. J. (2010). Molecular mechanisms of polyploidy and hybrid vigor. Trends Plant Sci. 15, 57–71. doi: 10.1016/j.tplants.2009.12.003
Chen, Z. J. (2013). Genomic and epigenetic insights into the molecular bases of heterosis. Nat. Rev. Genet. 14, 471–482. doi: 10.1038/nrg3503
Choe, J., Lin, S., Zhang, W., Liu, Q., Wang, L., Ramirez-Moya, J., et al. (2018). mRNA circularization by METTL3-4F3b enhances translation and promotes oncogenesis. Nature 561, 556–560. doi: 10.1038/s41586-018-0338-8
Cubas, P., Vincent, C., and Coen, E. (1999). An epigenetic mutation responsible for natural variation in floral symmetry. Nature 401, 157–161. doi: 10.1038/43657
Cui, X., Meng, J., Zhang, S., Chen, Y., and Huang, Y. (2016). A novel algorithm for calling mRNA m6A peaks by modeling biological variances in MeRIP-seq data. Bioinformatics 32, i378–i385. doi: 10.1093/bioinformatics/btw281
Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amargilo, N., and Rechavi, G. (2013). Transcriptome-wide mapping of N6-methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing. Nat. Protoc. 8, 176–189. doi: 10.1038/nprot.2012.148
Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Olseneg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485, 201–206. doi: 10.1038/nature11112
Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., et al. (2016). The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. Nature 530, 441–446. doi: 10.1038/nature16998
Dong, X., Reimer, J., Göbel, U., Engellhorn, J., He, F., Schoof, H., et al. (2012). Natural variation of H3K27me3 distribution between two Arabidopsis accessions and its association with flanking transposable elements. Genome Biol. 13:R117. doi: 10.1186/gb-2012-13-12-r117..
