Comparative Genomics and Functional Studies of Putative m^6A Methyltransferase (METTL) Genes in Cotton

Junfeng Cao, Chaochen Huang, Jun'e Liu, Chenyi Li, Xia Liu, Zishou Zheng, Lipan Hou, Jinquan Huang, Lingjian Wang, Yugao Zhang, Xiaoxia Shangguan, and Zhiwen Chen

Abstract: N6-methyladenosine (m^6A) RNA modification plays important regulatory roles in plant development and adapting to the environment, which requires methyltransferases to achieve the methylation process. However, there has been no research regarding m^6A RNA methyltransferases in cotton. Here, a systematic analysis of the m^6A methyltransferase (METTL) gene family was performed on twelve cotton species, resulting in six METTLs identified in five allotetraploid cottons, respectively, and three to four METTLs in the seven diploid species. Phylogenetic analysis of protein-coding sequences revealed that METTL genes from cottons, Arabidopsis thaliana, and Homo sapiens could be classified into three clades (METTL3, METTL14, and METTL-like clades). cis-element analysis predicated the possible functions of METTL genes in G. hirsutum. RNA-seq data revealed that GhMETTL14 (GH_A07G0817/GH_D07G0819) and GhMETTL3 (GH_A12G2586/GH_D12G2605) had high expressions in root, stem, leaf, torus, petal, stamen, pistil, and calycle tissues. GhMETTL14 also had the highest expression in 20 and 25 dpa fiber cells, implying a potential role at the cell wall thickening stage. Suppressing GhMETTL3 and GhMETTL14 by VIGS caused growth arrest and even death in G. hirsutum, along with decreased m^6A abundance from the leaf tissues of VIGS plants. Overexpression of GhMETTL3 and GhMETTL14 produced distinct differentially expressed genes (DEGs) in A. thaliana, indicating their possible divergent functions after gene duplication. Overall, GhMETTLs play indispensable but divergent roles during the growth of cotton plants, which provides the basis for the systematic investigation of m^6A in subsequent studies to improve the agronomic traits in cotton.

Keywords: m^6A methyltransferase; cotton; phylogenetics; VIGS; divergent function

1. Introduction

N6-methyladenosine (m^6A) methylation is the most abundant post-transcriptional modification of nucleotides in mRNA [1,2]. m^6A RNA methylation is a dynamic process and highly conserved in mammals, flies, plants, etc., that is mediated by m^6A writer, eraser, and reader proteins [3]. The m^6A methylation writer complex is composed of the m^6A methyltransferases and other factors, e.g., WTAP. METTL3 was the first m^6A RNA methyltransferase identified in Homo sapiens [4]. Subsequently, METTL14 was proposed as the second m^6A methyltransferase in humans with high homology to METTL3 [5–7]. Actually, subsequent studies showed that METTL14 had no methyltransferase activity, but could bind to RNA substrates by interacting with METTL3 to form a heterodimer [8–10].
WTAP was also discovered to be a major binding partner of the complex [11]. WTAP interacted with METTL3 and METTL14, and was required for METTL3–METTL14 heterodimer localization into the nucleus for the catalytic activity of the m^6^A methyltransferase complex [11].

Recent studies on m^6^A writers have also been conducted in plants extensively, including the m^6^A methyltransferase components being characterized in *Arabidopsis*. For example, AtMTA (METTL3 human homolog protein) is one of the earliest discovered m^6^A methyltransferases in *Arabidopsis* [12]. AtMTA proteins were actively distributed in the reproduction organs and apical meristems of *Arabidopsis* plants. AtMTA protein was found to be required for m^6^A mRNA methylation, and its function disruption prevented plant embryo development, which further resulted in embryo lethality [12]. In addition, PtrMTA, a *Populus trichocarpa* methyltransferase MTA homologous protein, was colocalized in the nucleus. PtrMTA-overexpressing plants significantly enhanced their tolerance to drought stress by increasing the density of trichomes and roots in poplar [13]. AtMTB is the human METTL14 homologous protein identified in *Arabidopsis*, but its function is unknown [14]. However, CIMTB, an AtMTB homolog protein in watermelon, and its expression was induced under drought stress. Overexpression of CIMTB enhanced drought tolerance by elevating the reactive oxygen species (ROS) scavenging ability and alleviating photosynthesis inhibition in tobacco [15]. Both these MTAs and MTBs belong to the MT-A70-like (MT) proteins [3]. These results imply that the m^6^A methyltransferases might serve as the positive regulatory factors in response to the different adverse stresses.

As m^6^A methyltransferases are highly conserved, the protein methyltransferases in *H. sapiens* and *Arabidopsis* would provide the necessary sequence references to identify the orthologous proteins or genes in other plant species. Cotton (*Gossypium*) is the most important fiber crop worldwide, and this genus contains more than 50 species [16–19]. There are many publicly available genomes for *Gossypium*, such as diploid D sub-genome species *Gossypium thurberi* (D1), *G. raimondii* (D5), *G. turneri* (D10) [20–23]; diploid A sub-genome species *G. herbaceum* (A1), *G. arboreum* (A2) [24–26]; diploid F sub-genome species *G. longicalyx* (F1) [27]; diploid G sub-genome species *G. hickii* (G1) [28], *G. australis* (G) [29]; and allotetraploid species *G. hirsutum* (AD1), *G. barbadense* (AD2), *G. tomentosum* (AD3), *G. mustelinum* (AD4), and *G. darwinii* (AD5) [26,30–37], as well as *Gossypium* sister genera *Gossypioides kirkii* [38]. These genome sequences provide a good platform for performing comparative genomics analysis and dissecting their gene functions. As the m^6^A methyltransferase (METTL) genes in *Gossypium* have not been investigated, several representative *Gossypium* species were selected to characterize the m^6^A methyltransferase genes and their roles in cotton.

In this study, we performed genome-wide screening for METTL genes in cotton, based on data gathered from recent whole-genome sequencing efforts. We used an in silico approach to identify and characterize METTL genes, and then focused on the characterization and phylogenetic relationships of *Gossypium METTLs*. The METTL gene structures, conserved domains, as well as cis-elements were explored in this systematical analysis. Moreover, the tissue-specific expression patterns and the transcriptional responses of *GhMETTLs* to abiotic stresses were examined. The possible biological functions of *GhMETTL3* and *GhMETTL14* were investigated in *G. hirsutum* and *A. thaliana* using the VIGS method and overexpression of transformation, respectively. Our data indicate the conservation of m^6^A methyltransferase (METTL) genes in cotton, and provide evidence for the divergent roles of *GhMETTL3* and *GhMETTL14* during *Gossypium* evolution.

2. Results

2.1. Identification and Chromosomal Location of RNA Methyltransferase Genes in *G. hirsutum*

Six methyltransferase-like genes (METTL) were identified in the *G. hirsutum* genome using Hmmersearch with conserved domains (PF05063 for mRNA: m^6^A methyltransferase) (Table 1). These 6 *GhMETTL* genes were dispersed over 6 of the 26 *G. hirsutum* chromosomes, with all homoeologs conserved (Figure 1). Among them, *GH_A12G2586/GH_D12G2605*...
(GhMETTL3_A/D) were the homologs of METTL3, and GH_A07G0817/GH_D07G0819 (GhMETTL14_A/D) were the homologs of METTL14 in H. sapiens. GH_A06G0112/GH_D06G0096 were another two copies of methyltransferase-like genes in G. hirsutum (Figure 1).

Table 1. Sequence characteristics of GhMETTL (G. hirsutum methyltransferase) genes and proteins.

| Gene Name      | Locus Name       | Chr | Genomics Position   | CDS  | No. of Introns | Size (aa) | MW   | pI   |
|----------------|------------------|-----|---------------------|------|----------------|-----------|------|------|
| GhMETTL3_A     | GH_A12G2586      | A12 | 104,206,704-104,211,250 | 2106 | 6              | 701       | 78.50 | 5.98 |
| GhMETTL3_D     | GH_D12G2605      | D12 | 58,367,589-58,370,785    | 2106 | 6              | 701       | 78.42 | 5.92 |
| GhMETTL14_A    | GH_A07G0817      | A07 | 10,130,789-10,135,646   | 3558 | 5              | 1185      | 132.59| 7.10 |
| GhMETTL14_D    | GH_D07G0819      | D07 | 8,927,821-8,932,500     | 3555 | 5              | 1184      | 132.49| 6.92 |
| GhMETTL-Like_A | GH_A06G0112      | A06 | 1,007,934-1,011,291     | 1269 | 8              | 422       | 48.45 | 7.18 |
| GhMETTL-Like_D | GH_D06G0096      | D06 | 890,591-893,931         | 1269 | 8              | 422       | 48.47 | 6.88 |

Note: bp, base pair; Chr, chromosome; aa, amino acid; MW (estimated), molecular weight; kDa, kilodalton; pI, isoelectric point. The gene accessions were downloaded from the CottonGen database [39].

Figure 1. Dispersed distribution of METTL genes in G. hirsutum (AD1) chromosomes: 6 GhMETTL genes are scattered over 6 of the 26 G. hirsutum chromosomes.

2.2. Structural Organization of GhMETTL Genes

Less than a threefold variation in length was detected in the predicted coding sequences (CDS) for these GhMETTLs, from 1269 bp for GH_A06G0112/GH_D06G0096 to 3558 bp for GH_A07G0817 (Table 1), which translates to proteins ranging from 422 amino acids (aa) to 1185 aa. Predicted isoelectric points (pI) for members of this family range from 5.92 to 7.18 (Table 1). All putative GhMETTL genes contain introns (Figure S1), which exhibit considerable variations in length and number. In general, homoeologous GhMETTL genes show highly similar intron patterns; however, intron structure among homoeologous pairs can exhibit variation in intron number (5 to 8) and length (Figure S1). One of the homoeologous gene pairs did exhibit divergence in intron length, namely, the fourth intron of GH_D12G2605 (GhMETTL3_D) was significantly less than that of GH_A12G2586 (GhMETTL3_A). Characterizing the gene structure in the diploid parental species for the homoeologs suggests that the length variation of the fourth intron occurred before the divergence of the allotetraploid (Figure S2). In addition, the phylogenetic relationship of
the GhMETTL gene family was consistent with the intron/exon structure characterized (Figure S1).

2.3. Phylogenetic Analysis of METTL Genes in Gossypium

We further evaluated the general preservation of METTL genes in 12 Gossypium species, Gossypioides kirkii, Arabidopsis thaliana, and Homo sapiens (Table 2). Two METTL genes (METTL3 and METTL14) were identified in H. sapiens. In A. thaliana, three METTL genes (AtMTA, AtMTB, and AtMTC) were identified; the relative of Gossypium, Gossypioides kirkii, had three METTL genes identified. The 12 Gossypium species surveyed recovered a minimum of three putative METTL genes in diploid species and six in five allotetraploid species (Table 2). Among D genome species, three METTL gene were identified in G. thurberi (D1), G. raimondii (D3), and G. turneri (D10), respectively. The two cultivated diploid species of G. herbaceum (A1) and G. arboreum (A2) had four copies. The sister species of A-genome G. longicalyx (F1) contained three METTL genes. METTL copy numbers in the allotetraploid species surveyed were conserved in six METTL genes in five species: G. hirsutum (AD1), G. barbadense (AD2), G. tomentosum (AD3), G. mustelinum (AD4), G. darwinii (AD5) (Table 2).

Table 2. Characteristics of the putative methyltransferase (METTL) genes in 12 Gossypium species, Gossypium kirkii, Arabidopsis thaliana, and Homo sapiens.

| Species                | METTL Number | METTL3 Homologs            | METTL14 Homologs          | METTL-Like Homologs |
|------------------------|--------------|-----------------------------|---------------------------|---------------------|
| Gossypium hirsutum     | 6            | GH_A12G2586/                | GH_A07G0817/              | GH_A06G0112/       |
|                        |              | GH_D12G2605                 | GH_D07G0819               | GH_D06G0096/       |
| Gossypium barbadense   | 6            | GB_A12G2683/                | GB_A07G0808/              | GB_A06G0089/       |
|                        |              | GB_D12G2687                 | GB_D07G0816               | GB_D06G0116/       |
| Gossypium tomentosum   | 6            | Ga12G272500/                | Ga07G086300/              | Ga06G01300/        |
|                        |              | Gd12G281700                 | Gd07G087100               | Gd06G01400/        |
| Gossypium mustelinum   | 6            | Ga12G260700/                | Ga07G087300/              | Ga06G01900/        |
|                        |              | Gm12G271800                 | Gm07G087500               | Gm06G012500/       |
| Gossypium darwini      | 6            | Ga12G271500/                | Ga07G086700/              | Gd06G013300/       |
|                        |              | Gd12G280500                 | Gd07G087200               | Gd06G012500/       |
| Gossypium herbaceum    | 4            | Gh12G29960                  | Gh07G08800                | Gh06G01080/        |
|                        |              | Ghe07G08800                 | Ghe06G01080/              | Gh06G01370/        |
| Gossypium arborum      | 4            | Gar12G30100                 | Gar07G08930               | Gar06G00970/       |
|                        |              | Gar07G08930                 | Gar06G00970/              | Gar06G01270/       |
| Gossypium longicalyx   | 3            | Golon.012G274900-1          | Golon.007G08200-1         | Golon.006G006700-1 |
| Gossypium thurberi     | 3            | Gothu.00035416-RA           | Gothu.00035901-RA         | Gothu.00028549-RA  |
| Gossypium raimondii    | 3            | Gorai.008G257500.1          | Gorai.001G08900.1         | Gorai.010G1900.1   |
| Gossypium turneri      | 3            | turneri_TURN.02g361810       | turneri_TURN.07g196460    | turneri_TURN.06g175650 |
| Gossypium australae    | 3            | G. australie_KAA3460050.1   | G. australie_KAA3458407.1 | G. australie_KAA3476391.1 |
| Gossypioides kirkii    | 3            | Kirkii_Version3_Juced.      | Kirkii_Version3_Juced.    | Kirkii_Version3_Juced. |
| Arabidopsis thaliana   | 3            | AT4G10760.1_MTA             | AT4G09980.1_MTB           | AT1G19340.1        |
| Homo sapiens           | 2            | NM_019852.4_METTL3         | NM_020961.3_METTL14       |                     |

Note: The gene accessions from 12 Gossypium species and Gossypium kirkii were downloaded from the CottonGen database [39]. The gene accessions of Arabidopsis thaliana and Homo sapiens were downloaded from the TAIR database [41] and NCBI database [https://www.ncbi.nlm.nih.gov (accessed on 1 July 2008)], respectively.

To further investigate the evolutionary relationships between the METTL genes provided here, we specifically assessed these relationships using the protein sequences of 61 METTL genes, including 53 Gossypium, 3 Gossypioides kirkii, 3 A. thaliana, and 2 H. sapiens METTL genes for phylogenetic analysis (Figure 2). Three clades (METTL3, METTL14, and METTL-like) were robustly supported in Gossypium, with one A. thaliana or Gossypioides kirkii
gene associated with each clade, respectively, but shared METTL3 and METTL14 clades with *H. sapiens*.

**Figure 2.** Phylogenetic analysis of METTL genes from 12 *Gossypium* species, *Gossypium kirkii*, *Arabidopsis thaliana*, and *Homo sapiens*. The phylogenetic tree was established with entire amino acid sequences using NJ methods. The numbers on the branches indicate bootstrap support values from 1000 replications.

Overall, the expected diploid–polyploid topology is reflected in the tree for each set of orthologous/homoeologous genes, indicating general preservation during diploid divergence and through polyploid evolution. That is, the number of METTL genes in allotetraploids was generally additive with respect to the model diploid progenitors, with each homoeolog (A<sub>t</sub> or D<sub>t</sub>) sister to their respective parental copies. METTL3 and METTL14 clades had equal numbers of METTL genes (Figure 2). In clades METTL3 and METTL14, genes related to *AtMTA* and *AtMTB* were exhibited parallelly, and there was no duplication in *Gossypium* species, compared to *A. thaliana* and *H. sapiens*. Therefore, it is speculated that the function of METTL genes in these clades of cotton is similar to that of the corresponding METTL genes in *Arabidopsis* and *H. sapiens*. 
2.4. Cis-Element Analysis of METTL Genes in G. hirsutum

Cis-elements are responsive to corresponding stimulations to regulate the expression of genes [41]. In this study, a 1.5 kb upstream region from the start codon of each METTL gene in *G. hirsutum* was extracted to investigate putative cis-elements involved in the mediation of gene expression using the PlantCARE server [42]. We completely identified 213 cis-elements among 6 *GhMETTL* genes, ranging from 24 in *GH_A12G2586* to 44 in *GH_D06G0096* (Figure S3). All *GhMETTL* gene promoters had several TATA-box and CAAT-box elements (Figure S3). Four *GhMETTL* gene promoters processed at least one abscisic acid responsiveness element (ABRE), and three *GhMETTL* genes had at least one AE-box (part of a module for light response in *A. thaliana*). *GhMETTL14_A* (*GH_A07G0817*) processed the element involved in HD-Zip 1 (involved in differentiation of the palisade mesophyll cells in *A. thaliana*). Six *GhMETTLs* had light responsiveness elements (GA-motif, GT1-motif, and G-box) and at least two MYB or one MYC binding site element (MYB/MYC). In addition, MYB binding sites involved in drought inducibility (MBS) also existed in these *GhMETTLs* genes.

2.5. Expression Patterns of GhMETTL Genes in Different G. hirsutum Tissues

The expression patterns of a gene family can provide valuable information to predict the possible biological functions of each gene. Analysis of six *GhMETTL* genes showed that most genes have different spatial expression patterns. For instance, the expression levels of *GhMETTL14* (*GH_A07G0817*/*GH_D07G0819*) and *GhMETTL3* (*GH_A12G2586*/*GH_D12G2605*) in the root, stem, leaf, torus, petal, stamen, pistil, and calyce were significantly higher than the other two *GhMETTL*-like (*GH_A06G0112*/*GH_D06G0096*) genes (Figure 3A). *GhMETTL14* (*GH_A07G0817*/*GH_D07G0819*) and *GhMETTL3* (*GH_A12G2586*/*GH_D12G2605*) presented the highest expression level in the pistil than other tissues (Figure 3A). In addition, *GhMETTL14* (*GH_A07G0817*/*GH_D07G0819*) still showed significantly higher expression in seed and root samples at different time points after seed germination (Figure 3B). In ovule samples at different development stages, *GhMETTL14* (*GH_A07G0817*/*GH_D07G0819*) had the highest expression levels, followed by *GhMETTL3* (*GH_A12G2586*/*GH_D12G2605*) (Figure 3C). In different development periods of fiber samples, *GhMETTL14* (*GH_A07G0817*/*GH_D07G0819*) had the highest expression in 20 and 25 dpa fiber cells (Figure 3D), suggesting that this gene might play an important role at the cell wall thickening stage. *GhMETTL3* (*GH_A12G2586*/*GH_D12G2605*) also presented higher expression in 20 and 25 dpa fiber cells than in 5 and 10 dpa cells (Figure 3D). The above two genes were the homologs of the *H. sapiens* METTL3 and METTL14, respectively (Figure 2). These results suggest that *GhMETTL* genes may not be involved in the regulation of fiber cell elongation (Figure 3D). *GhMETTL3* and *GhMETTL14* might serve as the functional genes in charge of cotton fiber cell wall development. Based on the gene expression data, two *GhMETTL*-like (*GH_A06G0112*/*GH_D06G0096*) genes were barely expressed across different tissues, which might diverge to become pseudogenes.

2.6. Expression Changes of GhMETTL Genes in G. hirsutum under Different Stresses

A variety of abiotic stresses pose threats to the growth and development of cotton plants. Therefore, we comprehensively analyzed the expression changes of *GhMETTL* genes under simulated drought (PEG 6000), salt (NaCl), heat, and cold abiotic stresses using RNA-seq data (Figure S4). At different time points of PEG6000-simulated drought conditions, the expression levels of six *GhMETTL* genes were slightly decreased, but the difference was not significant (l_{lg2} (fold change) \geq 1) as the threshold of differentially expressed genes. Among them, the *GhMETTL14_A* (*GH_A07G0817*) gene was downregulated after 1 h PEG treatment (Figure S4A), but not changed at the time points of 3, 6, and 12 h. The expression of other *GhMETTL* genes did not change after PEG treatment for 1, 3, 6, and 12 h (Figure S4A). These results indicate that few *GhMETTL* genes were involved in response to drought stress in *G. hirsutum*. Meanwhile, at the four time points of salt or high-temperature...
stress, the expressions of six GhMETTL genes both did not change after 1, 3, 6, and 12 h of NaCl or heat treatment (Figure S4B,C).

As for the four time points of low-temperature stress in G. hirsutum, GhMETTL14 (GH_A07G0817/GH_D07G0819) were upregulated after 6 h of low-temperature treatment (Figure S4D). However, the expressions of GhMETTL3 (GH_A12G2586/GH_D12G2605) were repressed after 12 h low-temperature treatment. These results indicate their divergent regulating roles under low-temperature stress. The expression levels of these genes did not change at other time points. The overall results show that GhMETTL3 and GhMETTL14 played an important role in response to the adverse stress in cotton, and it might be worth further exploring their biological functions.

2.7. Suppressing GhMETTL3 and GhMETTL14 Caused Growth Arrest in G. hirsutum

To explore the role of the identified GhMETTLS during the growth of cotton plants, the virus-induced gene silencing (VIGS) approach was used to suppress the expression of GhMETTL3 and GhMETTL14 (Figure 4). About 400 bp fragments from the conserved regions of both genes were inserted into the tobacco rattle virus (TRV) vector, respectively, and TRV-vector-carrying Agrobacterium strains were then co-inoculated into 1-week-old cotton plant cotyledons by the needleless syringe method. From those silenced cotton plants, we selected each independent VIGS line for phenotypic analyses (Figure 4A–C). The qPCR results show that the expression levels of GhMETTL3 and GhMETTL14 genes in silenced plants were significantly downregulated (Figure 4D,E). Along with the decreased expression of the GhMETTL genes, suppressing GhMETTL3 and GhMETTL14 both caused growth arrest.

Figure 3. Expression patterns of GhMETTL genes in different cotton tissues and fiber cells of different stages based on the RPKM values of RNA-seq data. (A) Expression profiles of GhMETTL genes in eight cotton tissues. (B) Expression patterns of GhMETTL genes in seed germination, cotyledons, and roots after germination. (C) Expression patterns of GhMETTL genes in ovules of different stages. (D) Expression patterns of GhMETTL genes in fibers of different stages.
arrest in *G. hirsutum*, with a more severe withered phenotype produced by *GhMETTL14* (Figure 4A). The resulting phenotypes also indicate that the inhibition of *GhMETTL14* resulted in slight yellowing of the leaves (Figure 4B). One month later, downregulation of *GhMETTL3* and *GhMETTL14* by VIGS led to the death of cotton plants (Figure 4C). Then, the m^6^A abundances in the leaf tissues of *GhMETTL3* and *GhMETTL14* VIGS plants were determined by quantitative mass spectrometry (MS) analysis (Figure 4F). For the TVR2 control leaf tissue, the m^6^A/A ratio of total RNA was approximately 0.075% (Figure 4F), which was significantly higher than that in *GhMETTL3*-silenced and *GhMETTL14*-silenced plants, ranging from 0.050% to 0.057% of m^6^A/A ratio, respectively. To provide more evidence of the molecular functions of *GhMETTLs*, we checked the expression levels of growth-related genes whose mRNA contains m^6^A according to the public cotton m^6^A-seq data [31]. The results show that the expression level of MYB (*Gohir.D06G213100*), bHLH (*Gohir.A11G121700*), and epidermal growth factor receptor (*Gohir.A01G103800*) decreased both in the *GhMETTL3* as well as the *GhMETTL14* VIGS plants (Figure S5A). Furthermore, 10 µL actinomycin D was used for the VIGS plants to inhibit the transcription, and the mRNA stability of the three genes was reduced dramatically (Figure S5B). Overall, the results show that the decreased m^6^A abundance might cause the growth arrest of *G. hirsutum* juvenile plants.

Figure 4. Suppressing *GhMETTL3* and *GhMETTL14* by virus-induced gene silencing (VIGS) approach. The 400 bp CDS fragments of two genes were inserted into the tobacco rattle virus (TRV) vector, and
TRV-vectors-carrying Agrobacterium strains were then co-inoculated into 1-week-old cotton plant cotyledons by the needleless syringe method. (A) Suppressing GhMETTL3 and GhMETTL14 both caused development arrest in *G. hirsutum*. (B) The leaf phenotypes of GhMETTL3 and GhMETTL14 VIGS cotton plants compared with the TRV2 blank vector plants. (C) Downregulation of GhMETTL3 and GhMETTL14 by VIGS led to the death of cotton plants. (D, E) The qPCR data identified the repressed expression levels of GhMETTL3 (D) and GhMETTL14 (E) genes in silenced plants. (F) The decreased m^6^A abundance from the leaf tissues of GhMETTL3 and GhMETTL14 VIGS plants determined by quantitative mass spectrometry (MS) analysis compared with the TVR2 control. Means of triplicates ± SD, **p < 0.01, *p < 0.05, Student’s t-test.

2.8. Subcellular Localization of GhMETTL3 and GhMETTL14 Proteins

To predict the subcellular localization of GhMETTL3 and GhMETTL14 proteins, we constructed two fusion vectors. A C-terminal fusion to the GFP gene was generated, driven by the CaMV 35S promoter, to determine the subcellular localization. Then, GhMETTL3:GFP and GhMETTL14:GFP fusion proteins were transiently expressed in tobacco epidermal cells and viewed using a confocal fluorescence microscope (Figure 5). The transient expression of GhMETTL3:GFP and GhMETTL14:GFP fusion proteins were specifically localized in the nucleus, not other subcellular compartments. Combined with the m^6^A ratio in the VIGS cotton, these results indicate that the GhMETTL3 and GhMETTL14 proteins act as the writers of m^6^A in cotton.

Figure 5. Subcellular localization assays of GhMETTL3 and GhMETTL14, showing nuclear localization of GhMETTL3-GFP (upper) and GhMETTL14-GFP (bottom). Agrobacterium cells harboring the fusion protein genes were infiltrated into leaves of *Nicotiana benthamiana* through the abaxial surface, and 72 h later the samples were observed under a confocal microscope (LSM510, Zeiss). Scale bar, 25 μm.
2.9. Overexpression of GhMETTL3 and GhMETTL14 Produced Distinct Differentially Expressed Genes (DEGs) in A. Thaliana

To further analyze the biological function of GhMETTL3 and GhMETTL14, we transformed these two genes into Arabidopsis using the floral dip method, and then we obtained the T₁ generation transgenic plants. However, we did not observe any significant phenotypic difference between two transgenic and wild type Arabidopsis plants during the whole growth period. To gain insight into the molecular mechanisms by overexpressing GhMETTL3 and GhMETTL14 into Arabidopsis plants, we performed RNA-seq of leaf tissues from two transgenic plants compared with the wild type.

Transcriptome data showed that there indeed existed differentially expressed transcripts in the leaf samples after transformation. In total, we identified 71 differentially expressed genes (DEGs) in GhMETTL3 overexpressed Arabidopsis plants compared with the wild type, among which 42 were induced and 29 were repressed (Figure S6A, Table S2). Meanwhile, 1274 DEGs were identified in GhMETTL14 overexpressed Arabidopsis plants with 315 upregulating and 958 downregulating (Figure S6B, Table S3).

Subsequently, these DEGs were subjected to KEGG pathway analysis to identify the functional categorization. Figure S7 lists the results of the KEGG analysis for DEGs. DEGs between OEGhMETTL3 and WT were assigned to 17 KEGG pathways, including RNA degradation, phenylpropanoid biosynthesis, ascorbate and aldarate metabolism, inositol phosphate metabolism, and the MAPK signaling pathway in plants, etc. (Figure S7A). In addition, DEGs between OEGhMETTL14 and WT were significantly enriched in 93 KEGG pathways, mainly including phenylpropanoid biosynthesis, glutathione metabolism, protein processing in endoplasmic reticulum, ascorbate and aldarate metabolism, starch and sucrose metabolism, pentose and glucuronate interconversions, pyruvate metabolism, fatty acid elongation, fatty acid biosynthesis, fatty acid degradation, nitrogen metabolism, RNA degradation, as well as the MAPK signaling pathway in plants (Figure S7B). The above results reveal that these two methyltransferase genes (GhMETTL3 and GhMETTL14) exerted extensive and distinct effects on the life processes in A. thaliana plants.

2.10. GhMETTL3 and GhMETTL14 Exhibited Divergent Functions in Transgenic Arabidopsis

Although no significant phenotypic changes were observed after these two genes were transformed into A. thaliana, transcriptome data still proved that GhMETTL3 and GhMETTL14 exhibited functional differentiation in transgenic Arabidopsis. We selected a total of 88 genes that were differentially expressed between transgenic and wild type Arabidopsis, and these DEGs could be classified into 11 families, including ethylene responsive transcription factor gene, bHLH transcription factor genes, WRKY transcription factor genes, dirigent protein genes, cytochrome P450 genes, calcium-binding protein genes, ABC transporter genes, glutathione S-transferase genes, detoxification genes, xyloglucan hydrolase genes, and CCR4-associated factor 1 homolog genes. Among them, six gene families comprising nine ABC transporter genes (Figure 6A), thirteen cytochrome P450 genes (CYP) (Figure 6B), six detoxification genes (DTX) (Figure 6C), four bHLH transcription factor genes (Figure 6D), three dirigent protein genes (DIR) (Figure 6E), and nine glutathione S-transferase genes (Figure 6F) were all downregulated in OEGhMETTL14 transgenic Arabidopsis plants, but unchanged in OEGhMETTL3 compared with the wild type plants. These results indicate that these two genes were functionally inconsistent.
Figure 6. Expressions of six family genes were reduced in OEGhMETTL14 but not changed in OEGhMETTL3 transgenic Arabidopsis plants compared with the wild type. (A) ABC transporter genes. (B) cytochrome P450 genes (CYP). (C) detoxification genes (DTX). (D) bHLH transcription factor genes. (E) dirigent protein genes (DIR). (F) glutathione S-transferase genes.

In addition, among thirteen calcium-binding protein (CML) genes (Figure S7A), the expression levels of eight CML genes (CML12, CML27, CML39, CML41, CML42, CML44, CML46, and CML47) were downregulated in OEGhMETTL14, but not changed in OEGhMETTL3 transgenic Arabidopsis plants compared with the wild type. Four CML genes (CML23, CML37, CML38, and CML40) were repressed in OEGhMETTL14, but induced in OEGhMETTL3.
transgenic Arabidopsis plants. Expression of one CML gene (CML45) was increased in OEGhMETTL3, but unchanged in OEGhMETTL14 transgenic Arabidopsis plants (Figure S8A). For another fifteen ethylene responsive transcription factor genes (ERF) (Figure S8B), there were six ERF genes (ERF019, ERF043, ERF060, ERF061, ERF114, and ERF115) being downregulated in OEGhMETTL14 plants compared with the wild type, but not changed in OEGhMETTL3. Meanwhile, three ERF genes (ERF1, ERF13, and ERF1B) were reduced in OEGhMETTL14, but elevated in OEGhMETTL3 plants. There still existed six ERF genes (ERF5, ERF016, ERF017, ERF018, ERF104, and ERF105) that were upregulated in OEGhMETTL3 plants, but unchanged in OEGhMETTL14 (Figure S7B). In addition, the presence of divergence between these two GhMETTL genes was also revealed from the expression patterns of twelve WRKY transcription factor genes (Figure S7C). Thereinto, the expressions of eight WRKY genes (WRKY8, WRKY28, WRKY29, WRKY48, WRKY51, WRKY54, WRKY62, and WRKY70) were downregulated in OEGhMETTL14 compared with the wild type, but not changed in OEGhMETTL3. Four WRKY genes (WRKY33, WRKY40, WRKY46, and WRKY53) presented reduced expressions in OEGhMETTL14, but induced in OEGhMETTL3 (Figure S8C). Expressions of three CCR4-associated factor 1 homolog genes (CAF1-11, CAF1-5, and CAF1-9) were upregulated in OEGhMETTL3 plants, but not changed in OEGhMETTL14 (Figure S8D). Finally, discordant expression patterns of four xyloglucan hydrolase genes (XTH) were also observed between OEGhMETTL14 and OEGhMETTL3 transgenic Arabidopsis plants (Figure S8E). Among them, the expressions of two XTH genes were upregulated in OEGhMETTL3 plants compared with that of wild type, but not changed in OEGhMETTL14. Two XTH genes exhibited repressed expressions in OEGhMETTL14, but upregulated in OEGhMETTL3 compared with the wild type (Figure S8E). The above results further confirm that the incongruent expression levels of these DEGs were caused by the transformations of GhMETTL3 and GhMETTL14 into Arabidopsis thaliana. Their biological functions might also be diverged and is worth further exploring in subsequent studies.

3. Discussion

N6-methyladenosine (m6A) in RNA represents an indispensable post-transcriptional modification in gene regulation. It was involved in RNA stability and degradation [43] and mRNA alternative splicing [44], thus affecting translation efficiency [45] and miRNA processing [46]. Meanwhile, the m6A modification plays a critical role in plant growth, development, and response to the adverse environmental stresses. The m6A writer components include MTA, MTB, FIP37, VIRILIZER, and HAKAI in A. thaliana [3,47–49]. When the expression level of m6A writer components, such as m6A methyltransferase, were reduced, the abundance of m6A also showed a dramatic reduction [12]. Knockout of m6A methyltransferase MTA prevented development and eventually led to embryonic lethality, while inactivation of other writer components resulted in increasing trichome branch phenotypes, as well as the over-proliferation of shoot apical meristem [3,47,48].

Previous reports have indicated that m6A modification is involved in abiotic and biotic stress responses [50–53]. The increase in 5’ UTR m6A level would promote protein translation in response to the heat shock stress [54]. Another study showed that m6A methylation was involved in salt stress tolerance in Arabidopsis. Both mutants of m6A writer components, such as MTA and MTB, displayed salt-sensitive phenotypes along with a reduced m6A level [55]. On the other hand, m6A methylation level showed a reduction after exposure to virus infection in Nicotiana tabacum and drought stress in Zea mays, respectively [56,57]. In addition, the expression of m6A methyltransferase CMTB in watermelon was also induced by drought stress. Overexpression of CMTB in tobacco plants increased their drought tolerance by enhancing ROS scavenging activity [15]. Overexpressing Populus trichocarpa m6A methyltransferase also enhanced the tolerance to drought stress by increasing the trichome and root density in poplar [13]. In our study, we also found the expression of GhMETTL14_A (GH_A07G0817) gene was repressed after simulated drought PEG treatment. These results indicate that METTL genes that mediated m6A modification were involved in response to environmental stresses in plants.
For the emergence of homologous genes in a single gene family after duplication, homologous genes generally have different fates: both genes retained for dosage effect or diverged to produce new function \[58,59\]. In METTL3 mutants, the translation of mRNAs containing m$^6$A modifications in the 5'UTR was reduced, indicating translation efficiency was influenced by 5'UTR m$^6$A in cells \[60,61\]. Early reports also found that m$^6$A modifications in 5'UTR regions were positively correlated with gene expression in *Arabidopsis* \[62\]. These investigations suggested that the METTL3 protein was the most active m$^6$A methyltransferase enzyme to catalyze m$^6$A RNA methylation, and a core m$^6$A methyltransferase component. However, studies found that METTL14 had no methyltransferase activity, but could form a stable antiparallel heterodimer with METTL3 to further bind to the RNA substrates in cells \[8–10\]. In *Arabidopsis*, the AtMTA protein was necessary for m$^6$A mRNA methylation, and loss of function led to plant embryo development arrest, and further to lethality \[12\]. However, the AtMTB identified might have no m$^6$A methyltransferase activity in *Arabidopsis* \[14\]. Furthermore, only one METTL3 or MTA homologous protein was identified in barley and *Micromonas pusilla*. In these two species, the METTL14 or MTB protein may have undergone functional gene loss during evolution or evolved into a distinct molecular role of a new gene family, but not the MT-A70-like protein \[3\]. In our study, the similar phenomenon of GhMETTL gene’s function divergence was observed in *Gossypium*. For instance, the expressions of *GhMETTL14* (*GH_A07G0817/GH_D07G0819*) were induced after low-temperature treatment, but *GhMETTL3* (*GH_A12G2586/GH_D12G2605*) were repressed under low-temperature stress. Meanwhile, *GhMETTL3* and *GhMETTL14* also exhibited divergent functions in transgenic *Arabidopsis* based on the transcriptome data. Compared to the wild type, the expressions of DEGs between OEGhMETTL3 and OEGhMETTL14 transgenic *Arabidopsis* were always inconsistent. These results imply that the homologous genes in a single gene family, such as m$^6$A methyltransferases METTL3 and METTL14, have divergent regulatory roles during the process of growth and development of plants and animals.

Although there have been many studies on the response of METTL in regulating plant growth and development, as well as responses to adverse stresses in plants, there are few studies on the exploration the potential roles of METTLs in cotton. The overall results of this study show that *GhMETTL3* and *GhMETTL14* play different roles in cotton; it would be worth further exploring their biological functions with modern biotechnologies, such as gene editing.

4. Materials and Methods

4.1. Identification of METTL Family Genes and METTL Proteins in Diploid and Tetraploid *Gossypium* Species

The genome sequences of cotton species were downloaded from the CottonGen database (https://www.cottongen.org/ (accessed on 1 April 2020)) \[63\], including *G. raimondii* \[20\], *G. herbaceum* \[26\], *G. arboreum* \[24\], *G. hirsutum* \[30\], *G. barbadense*, *G. tomentosum*, *G. mustelinum*, and *G. darwinii* \[26,30–37\]. To identify all putative METTL proteins in each genome assembly, the METTL protein conserved domains (PF05063 for MTA70 superfamily) were used to develop a hidden Markov model \[64\] profile matrix via the hmmbuild program from the HMMER 3.3.1 package \[65\] with default parameters. This HMM profile matrix was used in conjunction with hmmersearch with default parameters against these *Gossypium* genome databases to identify putative METTL genes (*GhMETTLs*) \[66,67\]. Previously identified METTL gene sequences from *A. thaliana* (AtMTA: AT4G10760.1 and AtMTB: AT4G09980.1) and *H. sapiens* (METTL, NM_019852.4, and NM_020961.3) were retrieved from the TAIR database \[40\] and NCBI database (https://www.ncbi.nlm.nih.gov) for phylogenetic comparison, respectively. The presence of conserved domains in each *Arabidopsis* and *Gossypium* gene was verified by the SMART version 9 conserved domain search tool \[68\] and Pfam 35.0 databases \[69\].
4.2. Chromosomal Location and Gene Structure Analyses

Chromosomal locations for the identified GhMETTLs were extracted from the genome annotation gff3 file [30,36]. Chromosomal locations of the predicted GhMETTLs were visualized using TBtools windows-x64_1_098769 [70], and the exon–intron structure of each gene was displayed using the online tool GSDS 2.0 [71]. The number of amino acids, molecular weight (MW), and theoretical isoelectric point (pI) of putative GhMETTL proteins were determined using the ProtParam tool [72]. The GFF3 file was included as Supplementary File named as “METTL-GFF3”.

4.3. Sequence Alignment and Phylogenetic Tree Construction

Complete amino acid sequences for METTL genes from 12 Gossypium species, Gossypioideae kirkii, A. thaliana, and H. sapiens were aligned using MAFFT version 6 with the G-INS-i algorithm [73]. As METTLs are single-domain proteins with a conserved MTA70 domain, the NJ phylogenetic tree was constructed using MEGA version 11.0 [74] with p-distance model and uniform rates by sampling 1000 bootstrap replicates. The alignments, and tree files were included as Supplementary Files named as “METTL-NJ3 and METTL-aa”, respectively.

4.4. Analysis of Cis-Acting Element in Promoter Regions of GhMETTLs

The upstream sequences (1.5 kb) [75] of the GhMETTL genes were retrieved from the G. hirsutum genome sequence based on the gene locations [30,36]. Then, the retrieved promoter sequences were submitted to PlantCARE [42] to identify the potential Cis-acting element.

4.5. Expression Patterns of GhMETTLs in Different Tissues and Stress Conditions

Raw RNA-seq data for G. hirsutum seed, root, stem, leaf, torus, petal, stamen, ovary, calyx, ovule (-3 dpa, -1 dpa, 0 dpa, 1 dpa, 3 dpa, 5 dpa, 10 dpa, 20 dpa, 25 dpa, 35 dpa) and fiber (5 dpa, 10 dpa, 20 dpa, 25 dpa) were downloaded from the NCBI Sequence Read Archive (PRJNA 248163) [30,36], represented by one library each. Reads were mapped to the G. hirsutum genome [36] via HISAT2 2.2.1 software with default parameters, and read abundance was calculated via StringTie v2.2.0 [76,77]. Read counts were normalized in R 3.2 using the RUVSeq package version 3.0 [78] and the internal control reference gene GhUBQ7, which is detected at relatively constant levels across different cotton samples [79]. Potential batch effects were corrected by the ComBat-seq sva v3.36.0 [80]. Gene expression was estimated by Ballgown [77], using fragments per kilobase million (FPKM) values to calculate the gene expression levels across libraries. Expression levels of G. hirsutum leaf RNA-seq data (in FPKM) for each GhMETTL gene under drought, salt, heat, and cold stress (time points: 0, 1, 3, 6, 12 h) were retrieved from the ccNET database [81]. Genes were considered differentially expressed if expression varied more than twofold with a p-value of less than 0.05. TBtools windows-x64_1_098769 [70] was used to display the gene expression patterns from the calculated FPKM values [82].

4.6. Virus-Induced Gene Silencing (VIGS) Assays

The virus-induced gene silencing (VIGS) assays were performed based on previous reports [83–85]. Briefly, to knockdown the expression of the GhMETTL genes, ~400 bp fragments of the GhMETTL3 and GhMETTL14 cDNA from upland cotton R15 were PCR-amplified using TransTaq® DNA Polymerase High Fidelity (AP131-11, TransGen Biotech, Beijing, China) with specific primers (Table S1). The resulting PCR products were infused to pTRV2 to produce the VIGS vectors, named pTRV2-GhMETTL3 and pTRV2-GhMETTL14, respectively. The pTRV1 and pTRV2-GhMETTL3 or pTRV2-GhMETTL14 vectors were introduced into the Agrobacterium strain GV3101 through electroporation (Bio-Rad, Hercules, CA, USA). For the VIGS assay, the transformed Agrobacterium colonies containing pTRV1 and pTRV2-GhMETTL3 or pTRV2-GhMETTL14 were grown overnight at 28 °C. The Agrobacterium cells were collected and resuspended in the infiltration medium (10 mM MgCl₂, 10 mM MES, and 200 mM acetosyringone) and subsequently adjusted to an OD600
of 0.5. The Agrobacterium strains containing the pTRV1 and pTRV2-GhMETTL3 or pTRV2-GhMETTL14 vectors were mixed at a ratio of 1:1. Cotton seedlings with mature cotyledons but without a visible rosette leaf (7 days after germination) were infiltrated by inserting the Agrobacterium suspension into the cotyledons via a syringe. The plants were then grown at 22 °C in pots arranged in a growth chamber under a 16 h light–8 h dark cycle and at a humidity of 60%. The mutant of GhSU has an obvious leaf-yellowing phenotype during the seeding stage, and we used GhSU as the positive control for VIGS. For mRNA stability assay, 10 µL actinomycin D was used to treat the VIGS plants as well as the negative control. After 1, 3, and 6 h, the treated leaves were collected to detect the expression levels.

4.7. RNA Extraction, cDNA Synthesis, and RT-qPCR Expression Analyses

These experiments were conducted with the methods reported previously [86–88]. In brief, total RNAs from cotton leaf were extracted using a RNAprep pure plant kit (TIANGEN, Shanghai, China). The resulting RNAs were treated with DNase I prior to synthesizing cDNA with TransScript® First-Strand cDNA Synthesis SuperMix (AT301-02, TransGen Biotech, Beijing, China); these products were diluted fivefold before use. For quantitative real-time PCR, Primer v5.0 software was used to design gene-specific forward and reverse primers (Table S1). Analyses were performed with SYBR-Green PCR Mastermix (TaKaRa) on a cycler (Mastercycler RealPlex; Eppendorf Ltd., Shanghai, China). The G. hirsutum histone-3 (GhHIS3) genes were used as internal references, and the relative amount of amplified product was calculated following the 2−ΔΔCt method [89].

4.8. Quantitative Analysis of m6A Level in Cotton Leaf Tissue

The quantification of m6A level in cotton leaf was conducted following the methods of a previous report [90]. In brief, mRNA was extracted from 10 µg total RNA for each sample with Dynabeads™ mRNA Purification Kit (Invitrogen, 61006). Then, 150 ng purified RNA was digested by 1 U nuclease P1 (Sigma, N8630) in 10 mM NH4Ac reaction buffer, made up to 20 µL, at 42 °C for 3 h. Subsequently, 1 U rSAP (NEB, M0371S) and 5 µL rCutSmart buffer were added to the mixture, and incubated at 37 °C for an additional 6 h. In addition, 1 µL digested RNA was injected into a LC-MS/MS (AB SCIEX QTRAP 5500) to detect the mononucleotide. The m6A level was detected and quantified in the positive ion multiple reaction monitoring (MRM) mode with nucleoside to base ion mass transitions (282.0 to 150.1 for m6A and 268.0 to 136.0 for A). The concentrations of m6A level in cotton leaf were calculated from the standard curve generated from pure nucleoside standards.

4.9. Subcellular Localization of GhMETTL3 and GhMETTL14 Proteins and Microscopic Observation

The protein subcellular localization experiments were carried out as described in a previous study [90]. The coding region of GhMETTL3 and GhMETTL14 was fused to GFP by pEASY®-Basic Seamless Cloning and Assembly Kit (CU201-02, TransGen Biotech, Beijing, China). The vectors carrying p35S::GhMETTL3-GFP and p35S::GhMETTL14-GFP were introduced into Agrobacterium cells, which were infiltrated with a syringe into the abaxial surface of N. benthamiana leaves for transient expression. Three days later, the infiltrated areas were observed under a confocal microscope (LSM510, Zeiss, Oberkochen, Germany).

4.10. Arabidopsis Transformation

The floral dip method was used for Arabidopsis transformation [91]. In brief, to overexpress GhMETTL3 and GhMETTL14 in A. thaliana, the full-length cDNA of GH_A12G2586 (GhMETTL3_A) and GH_A07G0817 (GhMETTL14_A) was cloned into the pCAMBIA2301 vector driven by the CaMV35S promoter, respectively. The recombinant plasmid was transferred into Agrobacterium cells and then transformed into A. thaliana plants via the floral dip method [91,92]. Kanamycin painting and PCR analysis confirmed the transgenic lines.
4.11. RNA Sequencing of OEGhMETTL3, OEGhMETTL14, and Wild Type A. thaliana Plants

Total RNA was isolated from the OEGhMETTL3, OEGhMETTL14, and wild type A. thaliana plants. Library construction and sequencing were performed following previous reports [93,94]. In brief, 1 μg of purified mRNA was selected for cDNA library construction. mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. cDNA fragments 240 bp in length were preferentially selected, and the library fragments were purified using an AMPure XP system (Beckman Coulter, Beverly, USA). Finally, the PCR products were purified (AMPure XP system), and library quality was assessed on an Agilent Bioanalyzer 2100 system. After cluster generation, the library preparations were sequenced on an Illumina HiSeqTM 2500 platform, and 150 bp paired-end reads were generated. Three biological replicates were performed for the transgenic and wild type A. thaliana plants. RNA-sequencing data that support the findings of this study have been deposited in the Genome Sequence Archive in the BIG Data Center of Sciences (https://bigd.big.ac.cn/ (accessed on 28 September 2021)) under accession CRA008376.

4.12. Read Data Quality Control, Mapping, and Calculations of the Differentially Expressed Genes

Raw data (raw reads) in fastq format were first processed through FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/ (accessed on 2 February 2010)). Clean data (clean reads) were obtained by removing reads containing adapters, reads containing poly-N sequences, and low-quality reads from the raw data. The downstream analyses were based on clean high-quality data. RNA-seq data were mapped to the reference genomes using HISAT2 software [76,77]. Only reads with a perfect match or one mismatch were further analyzed to calculate the expression values. Differential expression analysis of the two groups was performed using DESeq2 (v1.0) [95]. The resulting $p$-values were adjusted using Benjamini and Hochberg’s approach for controlling the false discovery rate (FDR). Genes with an adjusted $p$-value < 0.01 and twofold change (up and down) were defined as differentially expressed. TBtools windows-x64_1_098769 [70] was used to display the gene expression patterns from the TPM values.

4.13. Gene Annotation and Enrichment Analyses

Differentially expressed gene functions were annotated based on the following databases: Nr (NCBI nonredundant protein sequences, ftp://ftp.ncbi.nih.gov/blast/db/ (accessed on 1 July 2008)), Nt (NCBI nonredundant nucleotide sequences, ftp://ftp.ncbi.nih.gov/blast/db/ (accessed on 1 July 2008)), and Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/ (accessed on 1 January 2020)). KOBAS was used to test the statistical enrichment of the differentially expressed genes in the KEGG pathways [96].

5. Conclusions

We reported the structures, domains, phylogenetics, divergence, and cis-elements analyses results, systematically, of Gossypium METTL genes. We also investigated the expression patterns of METTL family genes in different cotton tissues and under different stress conditions to predict their possible biological functions. Both GhMETTL3 and GhMETTL14 genes were variably expressed in different cotton tissues, with particularly high expression in pistil and ovules. Suppressing GhMETTL3 and GhMETTL14 by VIGS led to development arrest in G. hirsutum. Overexpression of GhMETTL3 and GhMETTL14 into Arabidopsis did not produce significant phenotype differences, but effected the expressions of variant family genes, implying their divergent functions in the regulation of growth and development of cotton plants. Together, our results provide candidate genes to facilitate the functional identification of METTL genes in subsequent studies.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms232214111/s1.
Author Contributions: Conceptualization, X.S. and Z.C.; methodology, J.C.; software, C.H.; validation, J.C., C.H. and Z.Z.; formal analysis, L.H. and C.L.; investigation, X.L. and J.H.; resources, J.L.; data curation, J.C.; writing—original draft preparation, J.C. and Z.C.; writing—review and editing, Z.C.; visualization, L.W. and Y.Z.; supervision, Z.C.; project administration, Z.C.; funding acquisition, Z.C. All authors have read and agreed to the published version of the manuscript.

Funding: The work reported in this publication was supported by the Hainan Yazhou Bay Seed Laboratory (B22E49716, B22C39101), the Basic Research Program of Shanxi Province (20210302123409). The funding bodies did not participate in the design of the study; the collection, analysis, or interpretation of the data; or in the writing of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: RNA-sequencing data that support the findings of this study have been deposited in the Genome Sequence Archive in the BIG Data Center of Sciences (https://bigd.big.ac.cn) under accession CRA008376.

Acknowledgments: We thank Tian-Zhen Zhang for providing the RNA-seq expression data of the RPKM values and Xiao-Ya Chen for discussions.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

METTL, m^6^A methyltransferase; DEGs, differentially expressed genes; FPKM, fragments per kilobase of transcript per million mapped fragments; G. herbaeum, Gossypium herbaeum; G. arboreum, Gossypium arboreum; G. longicalyx, Gossypium longicalyx; G. thurberi, Gossypium thurberi; G. raimondii, Gossypium raimondii; G. turneri, Gossypium turneri; G. austral, Gossypium austral; G. hirsutum, Gossypium hirsutum; G. barbadense, Gossypium barbadense; G. tomentosum, Gossypium tomentosum; G. mustelinum, Gossypium mustelinum; G. darvinii, Gossypium darvinii; A. thaliana, Arabidopsis thaliana; H. sapiens, Homo sapiens; MW, molecular weight; pI, isoelectric point; qRT-PCR, quantitative real-time polymerase chain reaction.

References

1. Liu, N.; Pan, T. N6-methyladenosine–encoded epitranscriptomics. Nat. Struct. Mol. Biol. 2016, 23, 98–102. [CrossRef] [PubMed]
2. Zhou, L.; Gao, G.; Tang, R.; Wang, W.; Wang, Y.; Tian, S.; Qin, G. m^6^A-methyltransferase: Structural insights into the molecular mechanism of the m(6)A writer complex. Mol. Cell 2019, 77, 426–440. [CrossRef]
3. Yue, H.; Nie, Y.; Yan, Z.; Weinig, S. N6-methyladenosine regulatory machinery in plants: Composition, function and evolution. Plant Biotechnol. J. 2019, 17, 1194–1208. [CrossRef] [PubMed]
4. Bokar, J.A.; Shambaugh, M.E.; Polayes, D.; Matera, A.G.; Rottman, F.M. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA 1997, 3, 1233–1247. [PubMed]
5. Bujnicki, J.M.; Feder, M.; Radlinska, M.; Blumenthal, R.M. Structure prediction and phylogenetic analysis of a functionally diverse family of proteins homologous to the MT-A70 subunit of the human mRNA:m(6)A methyltransferase. J. Mol. Biol. 2002, 55, 431–444. [CrossRef]
6. Liu, J.; Yue, Y.; Han, D.; Wang, X.; Fu, Y.; Zhang, L.; Jia, G.; Yu, M.; Lu, Z.; Deng, X.; et al. A METTL3–METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat. Chem. Biol. 2013, 10, 93–95. [CrossRef]
7. Liu, J.; Li, K.; Cai, J.; Zhang, M.; Zhang, X.; Xiong, X.; Meng, H.; Xu, X.; Huang, Z.; Peng, J.; et al. Landscape and Regulation of m^6^A and m6Am Methylose across Human and Mouse Tissues. Mol. Cell 2019, 77, 426–440. [CrossRef]
8. Śledź, P.; Jinek, M. Structural insights into the molecular mechanism of the m(6)A writer complex. eLife 2016, 5, e18434.
9. Wang, P.; Doxtader, K.A.; Nam, Y. Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. Mol. Cell 2016, 63, 306–317. [CrossRef]
10. Wang, X.; Feng, J.; Xue, Y.; Guan, Z.; Zhang, D.; Liu, Z.; Gong, Z.; Huang, J.; Tang, C.; Zou, T.; et al. Structural basis of N6-adenosine methyltransferase by the METTL3–METTL14 complex. Nature 2016, 534, 575–578. [CrossRef]
11. Ping, X.-L.; Sun, B.-F.; Wang, L.; Xiao, W.; Yang, X.; Wang, W.-J.; Adhikari, S.; Shi, Y.; Lv, Y.; Chen, Y.-S.; et al. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. 2014, 24, 177–189. [CrossRef]
12. Zhong, S.; Li, H.; Bodi, Z.; Button, J.; Vespa, L.; Herzog, M.; Fray, R.G. MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. Plant Cell 2008, 20, 1278–1288. [CrossRef]
13. Lu, L.; Zhang, Y.; He, Q.; Qi, Z.; Zhang, G.; Xu, W.; Yi, T.; Wu, G.; Li, R. MTA, an RNA m^6A methyltransferase, enhances drought tolerance by regulating the development of trichomes and roots in poplar. *Int. J. Mol. Sci.* **2020**, *21*, 2462. [CrossRef]

14. Arribas-Hernández, L.; Bressendorff, S.; Hansen, M.H.; Poulsen, C.; Erdmann, S.; Brodersen, P. An m^6A-YTH module controls developmental timing and morphogenesis in arabidopsis. *Plant Cell* **2018**, *30*, 952–967. [CrossRef]

15. He, Y.; Li, Y.; Yao, Y.; Zhang, H.; Wang, Y.; Gao, J.; Fan, M. Overexpression of watermelon m^6A methyltransferase CMTB enhances drought tolerance in tobacco by mitigating oxidative stress and photosynthesis inhibition and modulating stress-responsive gene expression. *Plant Physiol. Biochem.* **2021**, *168*, 349–352. [CrossRef]

16. Wendel, F.J.; Grover, C. *Taxonomy and Evolution of the Cotton Genus, Gossypium*, 2nd ed.; American Society of Agronomy Inc.: Madison, WI, USA; Crop Science Society of America Inc.: Madison, WI, USA; Soil Science Society of America Inc.: Madison, WI, USA, 2015.

17. Wang, K.; Wendel, J.F.; Hua, J. Designations for individual genomes and chromosomes in Gossypium. *J. Cotton Res.* **2018**, *1*, 3. [CrossRef]

18. Chen, Z.; Grover, C.E.; Li, P.; Wang, Y.; Nie, H.; Zhao, Y.; Wang, M.; Liu, F.; Zhou, Z.; Wang, X.; et al. Molecular evolution of the plastid genome during diversification of the cotton genus. *Mol. Phylogenetics Evol.* **2017**, *112*, 268–276. [CrossRef]

19. Huang, G.; Huang, J.-Q.; Chen, X.-Y.; Zhu, Y.-X. Recent advances and future perspectives in cotton research. *Annu. Rev. Plant Biol.* **2021**, *72*, 437–462. [CrossRef]

20. Paterson, A.H.; Wendel, J.F.; Gundlach, H.; Guo, H.; Jenkins, J.; Jin, D.; Llewellyn, D.; Showmaker, K.C.; Shu, S.; Udall, J.; et al. Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. *Nature* **2012**, *492*, 423–427. [CrossRef]

21. Wang, K.; Wang, Z.; Li, F.; Ye, W.; Wang, J.; Song, G.; Yue, Z.; Cong, L.; Shang, H.; Zhu, S.; et al. The draft genome of a diploid cotton Gossypium raimondii. *Nature Genet.* **2012**, *44*, 1098–1103. [CrossRef]

22. Udall, J.A.; Long, E.; Hanson, C.; Yuan, D.; Ramaraj, T.; Conover, J.L.; Gong, L.; Arick, M.A.; Grover, C.E.; Peterson, D.G.; et al. De novo genome sequence assemblies of Gossypium raimondii and Gossypium turneri. *G3 Genomes Genet.* **2019**, *9*, 3079–3085. [CrossRef] [PubMed]

23. Grover, C.E.; Arick, M.A., II; Thrash, A.; Conover, J.L.; Sanders, W.S.; Peterson, D.G.; Frelichowski, J.E.; Scheffler, J.A.; Scheffler, B.E.; Wendel, J.F. Insights into the evolution of the New World diploid cottons (Gossypium, Subgenus Houzingenia) based on genome sequencing. *Genome Biol. Evol.* **2018**, *11*, 53–71. [CrossRef] [PubMed]

24. Li, F.; Fan, G.; Wang, K.; Sun, F.; Yuan, Y.; Song, G.; Li, Q.; Ma, Z.; Lu, C.; Zou, C.; et al. Genome sequence of the cultivated cotton *Gossypium arboresum*. *Nat. Genet.* **2014**, *46*, 567–572. [CrossRef] [PubMed]

25. Du, X.; Huang, G.; He, S.; Yang, Z.; Sun, G.; Ma, X.; Li, N.; Zhang, X.; Sun, J.; Liu, M.; et al. Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. *Nat. Genet.* **2018**, *50*, 796–802. [CrossRef] [PubMed]

26. Huang, G.; Wu, Z.; Percy, R.G.; Bai, M.; Li, Y.; Frelichowski, J.E.; Hu, J.; Wang, K.; Yu, J.Z.; Zhu, Y. Genome sequence of Gossypium herbaceum and genome updates of *Gossypium arboresum* and *Gossypium hirsutum* provide insights into cotton A-genome evolution. *Nat. Genet.* **2020**, *52*, 516–524. [CrossRef] [PubMed]

27. Grover, C.E.; Pan, M.; Yuan, D.; Arick, M.A.; Hu, G.; Brase, L.; Stelly, D.M.; Lu, Z.; Schmitz, R.J.; Peterson, D.G.; et al. The *Gossypium longicalyx* Genome as a Resource for Cotton Breeding and Evolution. *G3 Genes Genomes Genet.* **2020**, *10*, 1457–1467. [CrossRef] [PubMed]

28. Sheng, K.; Sun, Y.; Liu, M.; Cao, Y.; Han, Y.; Li, C.; Muhammad, U.; Daud, M.K.; Wang, W.; Li, H.; et al. A reference-grade genome assembly for *Gossypium bickii* and insights into its genome evolution and formation of pigment glands and gossypol. *Plant Commun.* **2022**, 100421. [CrossRef]

29. Cai, Y.; Cai, X.; Wang, Q.; Wang, P.; Zhang, Y.; Cai, C.; Xu, Y.; Wang, K.; Zhou, Z.; Wang, C.; et al. Genome sequencing of the Australian wild diploid species *Gossypium australe* highlights disease resistance and delayed gland morphogenesis. *Plant Biotechnol. J.* **2019**, *18*, 814–828. [CrossRef]

30. Zhang, T.; Hu, Y.; Jiang, W.; Fang, L.; Guan, X.; Chen, J.; Zhang, J.; Sasaki, C.A.; Scheffler, B.E.; Stelly, D.M.; et al. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* **2015**, *33*, 531–537. [CrossRef]

31. Chen, Z.J.; Sreedasyam, A.; Ando, A.; Song, Q.; De Santiago, L.M.; Hulse-Kemp, A.M.; Ding, M.; Ye, W.; Kirkbride, R.C.; Jenkins, J.; et al. Genomic diversifications of five Gossypium allotetraploid species and their impact on cotton improvement. *Nat. Genet.* **2020**, *52*, 525–533. [CrossRef]

32. Li, F.; Fan, G.; Lu, C.; Xiao, G.; Zou, C.; Kohel, R.J.; Ma, Z.; Shang, H.; Ma, X.; Wu, J.; et al. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat. Biotechnol.* **2015**, *33*, 524–530. [CrossRef]

33. Liu, X.; Zhao, B.; Zheng, H.-J.; Hu, Y.; Lu, G.; Yang, C.; Chen, J.-D.; Chen, J.-J.; Chen, D.-Y.; Zhang, L.; et al. *Gossypium barbadense* genome sequence provides insights into the evolution of extra-long staple fiber and specialized metabolites. *Sci. Rep.* **2015**, *5*, 14139. [CrossRef]

34. Yuan, D.; Tang, Z.; Wang, M.; Gao, W.; Tu, L.; Jin, X.; Chen, L.; He, Y.; Zhang, L.; Zhu, L.; et al. The genome sequence of Sea-Island cotton (*Gossypium barbadense*) provides insights into the allopolyploidization and development of superior spinnable fibres. *Sci. Rep.* **2015**, *5*, 17662. [CrossRef]
35. Wang, M.; Tu, L.; Yuan, D.; Zhu, D.; Shen, C.; Li, J.; Liu, F.; Pei, L.; Wang, P.; Zhao, G.; et al. Reference genome sequences of two cultivated tetraploid cottons, Gossypium hirsutum and Gossypium barbadense. Nat. Genet. 2019, 51, 224–229. [CrossRef]

36. Hu, Y.; Chen, J.; Fang, L.; Zhang, Z.; Ma, W.; Niu, Y.; Hu, L.; Deng, J.; Zhao, T.; Lian, J.; et al. Gossypium barbadense and Gossypium hirsutum genomes provide insights into the origin and evolution of tetraploid cotton. Nat. Genet. 2019, 51, 739–748. [CrossRef]

37. Yang, Z.; Ge, X.; Yang, Z.; Qin, W.; Sun, G.; Wang, Z.; Li, Z.; Liu, J.; Wu, J.; Wang, Y.; et al. Extensive intraspecific gene order and gene structural variations in upland cotton cultivars. Nat. Commun. 2019, 10, 1–15. [CrossRef]

38. Udall, J.A.; Long, E.; Ramaraj, T.; Conover, J.L.; Yuan, D.; Grover, C.E.; Gong, J.; Arick, M.A., 2nd; Masonbrink, R.E.; Peterson, D.G.; et al. The genome sequence of Gossypiodae kirkii illustrates a descending dysploidy in plants. Front. Plant. Sci. 2019, 10, 1541. [CrossRef]

39. Yu, J.; Jung, S.; Cheng, C.-H.; Lee, T.; Zheng, P.; Buble, K.; Crabb, J.; Humann, J.; Hough, H.; Jones, D.; et al. CottonGen: The community database for cotton genomics, genetics, and breeding research. Plants 2021, 10, 2805. [CrossRef]

40. Poole, R.L. The TAIR database. Methods Mol. Biol. 2007, 406, 179–212.

41. Hernandez-Garcia, C.M.; Finer, J.I. Identification and validation of promoters and cis-acting regulatory elements. Plant Sci. 2014, 217, 109–119. [CrossRef]

42. Lescot, M.; D Nicholson, J.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002, 30, 325–327. [CrossRef] [PubMed]

43. Zhang, F.; Zhang, Y.-C.; Liao, J.-Y.; Yu, Y.; Zhou, Y.-F.; Feng, Y.-W.; Yang, Y.-W.; Lei, M.-Q.; Bai, M.; Wu, H.; et al. The subunit of RNA N6-methyladenosine methyltransferase OsFIP regulates early degeneration of microspores in rice. PLoS Genet. 2019, 15, e1008120. [CrossRef] [PubMed]

44. Sharma, P.; Chatterjee, M.; Burman, N.; Khurana, J.P. Cryptochrome 1 regulates growth and development in Brassica through alteration in the expression of genes involved in light, phytohormone and stress signalling. Plant Cell Environ. 2011, 35, 994–1012. [CrossRef] [PubMed]

45. Lopez, L.; Carbone, F.; Bianco, L.; Giuliano, G.; Facella, P.; Perrotta, G. Tomato plants overexpressing cryptochrome 2 reveal altered expression of energy and stress-related gene products in response to diurnal cues. Plant Cell Environ. 2021, 106, 1759–1775. [CrossRef]

46. Visvanathan, A.; Somasundaram, K. mRNA traffic control reviewed: N6-methyladenosine (m6A) takes the driver’s seat. Bioessays 2018, 40, 1700093. [CrossRef]

47. Bodi, Z.; Zhong, S.; Mehra, S.; Song, J.; Graham, N.; Li, H.; May, S.; Fray, R.G. Adenosine Methylation in Arabidopsis mRNA is Associated with the 3’ End and Reduced Levels Cause Developmental Defects. Front. Plant Sci. 2012, 3, 48. [CrossRef]

48. Růžička, K.; Zhang, M.; Campilho, A.; Bodi, Z.; Kashif, M.; Saleh, M.; Eckhout, D.; El-Showk, S.; Li, H.; Zhong, S.; et al. Identification of factors required for m6A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. New Phytol. 2017, 215, 157–172. [CrossRef]

49. Zhang, F.; Zhang, Y.-C.; Liao, J.-Y.; Yu, Y.; Zhou, Y.-F.; Feng, Y.-Z.; Yang, Y.-W.; Lei, M.-Q.; Bai, M.; Wu, H.; et al. The subunit of RNA N6-methyladenosine methyltransferase OsFIP regulates early degeneration of microspores in rice. PLoS Genet. 2019, 15, e1008120. [CrossRef] [PubMed]

50. Sharma, P.; Chatterjee, M.; Burman, N.; Khurana, J.P. Cryptochrome 1 regulates growth and development in Brassica through alteration in the expression of genes involved in light, phytohormone and stress signalling. Plant Cell Environ. 2011, 35, 994–1012. [CrossRef] [PubMed]

51. Lopez, L.; Carbone, F.; Bianco, L.; Giuliano, G.; Facella, P.; Perrotta, G. Tomato plants overexpressing cryptochrome 2 reveal altered expression of energy and stress-related gene products in response to diurnal cues. Plant Cell Environ. 2011, 35, 994–1012. [CrossRef] [PubMed]

52. D’Amico-Damião, V.; Carvalho, R.F. Cryptochrome-related abiotic stress responses in plants. Front. Plant Sci. 2018, 9, 1897. [CrossRef]

53. D’Amico-Damião, V.; Lúcio, J.C.B.; Oliveira, R.; Gaion, L.A.; Barreto, R.F.; Carvalho, R.F. Cryptochrome 1a depends on blue light fluence rate to mediate osmotic stress responses in tomato. J. Plant Physiol. 2021, 258–259, 153374. [CrossRef]

54. Meyer, K.D.; Patil, D.P.; Zhou, J.; Zinoviev, A.; Skabkin, M.A.; Elemento, O.; Pestova, T.V.; Qian, S.-B.; Jaffrey, S.R. 5′ UTR m6A promotes cap-independent translation. Cell 2015, 163, 999–1010. [CrossRef]

55. Hu, J.; Cai, J.; Park, S.J.; Lee, K.; Li, Y.; Chen, Y.; Yun, J.Y.; Xu, T.; Kang, H. N6(-)-Methyladenosine mRNA methylation is important for salt stress tolerance in Arabidopsis. Plant J. 2021, 106, 1759–1775. [CrossRef]

56. Li, Z.; Shi, J.; Yu, L.; Zhao, X.; Ran, L.; Hu, D.; Song, B. N 6-methyl-adenosine level in Nicotiana tabacum is associated with tobacco mosaic virus. Virol. J. 2018, 15, 87. [CrossRef]

57. Miao, Z.; Zhang, T.; Qi, Y.; Song, J.; Han, Z.; Ma, C. Evolution of the RNA N (6)-Methyladenosine Methylyme Mediated by Genomic Duplication. Plant Physiol. 2020, 182, 345–360. [CrossRef]

58. Wendel, J.F. The wondrous cycles of polyploidy in plants. Annu. J. Bot. 2015, 102, 1753–1756. [CrossRef]

59. Chen, Z.; Cao, J.-F.; Zhang, X.-F.; Shangguan, X.-X.; Mao, Y.-B.; Wang, L.-J.; Chen, X.-Y. Cotton genome: Challenge into the polyploidy. Sci. Bull. 2017, 62, 1622–1623. [CrossRef]

60. Meyer, K.D.; Saletore, Y.; Zumbo, P.; Elemento, O.; Mason, C.E.; Jaffrey, S.R. Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3’ UTRs and near Stop Codons. Cell 2012, 149, 1635–1646. [CrossRef]

61. Meyer, K.D.; Jaffrey, S.R. Rethinking m6A readers, writers, and erasers. Annu. Rev. Cell Dev. Biol. 2017, 33, 319–342. [CrossRef]
62. Luo, G-Z.; MacQueen, A.; Zheng, G.; Duan, H.; Dore, L.C.; Lu, Z.; Liu, J.; Chen, K.; Jia, G.; Bergelson, J.; et al. Unique features of the m³A methylome in Arabidopsis thaliana. Nat. Commun. 2014, 5, 1–8. [CrossRef] [PubMed]

63. Yu, J.; Jung, S.; Cheng, C.-H.; Ficklin, S.P.; Lee, Y.; Zheng, P.; Jones, D.; Percy, R.G.; Main, D. CottonGen: A genomics, genetics and breeding database for cotton research. Nucleic Acids Res. 2013, 42, D1229–D1236. [CrossRef] [PubMed]

64. Wong, D.C.J.; Schlechter, R.; Vannonzi, A.; Höll, J.; Himmam, I.; Bogs, J.; Tornielli, G.B.; Castellarin, S.D.; Matus, J.T. A systems-oriented analysis of the grapevine R2R3-MYB transcription factor family uncovers new insights into the regulation of stilbene accumulation. DNA Res. 2016, 23, 451–466. [CrossRef]

65. Mistry, J.; Finn, R.D.; Eddy, S.R.; Bateman, A.; Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 2013, 41, e121. [CrossRef] [PubMed]

66. Finn, R.D.; Mistry, J.; Schuster-Böckler, B.; Griffiths-Jones, S.; Hollich, V.; Lassmann, T.; Mo Xen, S.; Marshall, M.; Khanna, A.; Durbin, R.; et al. Pfam: Clans, web tools and services. Nucleic Acids Res. 2006, 34, D247–D251. [CrossRef] [PubMed]

67. Wheeler, T.J.; Eddy, S.R. Hmm: DNA homology search with profile HMMs. Bioinformatics 2013, 29, 2487–2489. [CrossRef]

68. Letunic, I.; Bork, P. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2017, 46, D493–D496. [CrossRef]

69. Finn, R.D.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam families database: Towards a more sustainable future. Nucleic Acids Res. 2016, 44, D279–D285. [CrossRef]

70. Chen, C.J.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.H.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol. Plant 2020, 13, 1194–1202. [CrossRef]

71. Hu, B.; Jin, J.; Guo, A.-Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: An upgraded gene feature visualization server. Bioinformatics 2015, 31, 1296–1297. [CrossRef]

72. Wilkins, M.R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.C.; Williams, K.L.; Appel, R.D.; Hochstrasser, D.F. Protein Identification and Analysis Tools in the ExPASy Server. In 2-D Proteome Analysis Protocols. Methods in Molecular Biology; Humana Press: Totowa, NJ, USA, 1999; Volume 112, pp. 331–352. [CrossRef]

73. Katoh, K.; Standley, D.M. MAFFT Multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 2013, 30, 772–780. [CrossRef]

74. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with cufflinks and SPLAT. Nat. Methods 2014, 11, 531–552. [CrossRef]

75. Zhao, B.; Cao, J.F.; Hu, G.J.; Chen, Z.W.; Wang, L.Y.; Shangguan, X.X.; Wang, L.J.; Mao, Y.B.; Zhang, T.Z.; Wendel, J.F.; et al. Core cis-element variation confers subgenome-biased expression of a transcription factor that functions in cotton fiber elongation. New Phytol. 2018, 218, 1061–1075. [CrossRef]

76. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. Nat. Methods 2015, 12, 357–360. [CrossRef]

77. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, Stringtie and Ballgown. Nat. Protoc. 2016, 11, 1650–1667. [CrossRef]

78. Risso, D.; Ngai, J.; Speed, T.P.; Dudoit, S. Normalization of RNA-seq data using factor analysis of control genes or samples. Nat. Biotechnol. 2014, 32, 896–902. [CrossRef]

79. Wang, M.; Wang, Q.; Zhang, B. Evaluation and selection of reliable reference genes for gene expression under abiotic stress in cotton (Gossypium hirsutum L.). Gene 2013, 530, 44–50. [CrossRef]

80. Zhang, Y.; Parmigiani, G.; Johnson, W.E. Combat-seq: Batch effect adjustment for RNA-seq count data. NAR Genom. Bioinform. 2020, 2, lqaa078. [CrossRef]

81. You, Q.; Xu, W.Y.; Zhang, K.; Zhang, L.W.; Yin, X.; Yao, D.X.; Wang, C.C.; Zhang, X.Y.; Zhao, X.H.; Provart, N.J.; et al. ccNET: Database of co-expression networks with functional modules for diploid and polyploid Gossypium. Nucleic Acids Res. 2017, 45, 5625–5626. [CrossRef]

82. Huang, C.; Li, P.; Cao, J.; Zheng, Z.; Huang, J.; Zhang, X.; Shangguan, X.; Wang, L.; Chen, Z. Comprehensive identification and expression analysis of CRY gene family in Gossypium. BMC Genom. 2022, 23, 1–17. [CrossRef]

83. Zhu, J.; Chen, J.; Gao, F.; Xu, C.; Wu, H.; Chen, K.; Si, Z.; Yan, H.; Zhang, T. Rapid mapping and cloning of the virescent-1 gene in cotton by bulked segregant analysis–next generation sequencing and virus-induced gene silencing techniques. J. Exp. Bot. 2017, 68, 4125–4135. [CrossRef] [PubMed]

84. Tian, X.; Ruan, J.-X.; Huang, J.-Q.; Yang, C.-Q.; Fang, X.; Chen, Z.-W.; Hong, H.; Wang, L.-J.; Mao, Y.-B.; Lu, S.; et al. Characterization of gossypol biosynthetic pathway. Proc. Natl. Acad. Sci. USA 2018, 115, E5410–E5418. [CrossRef] [PubMed]

85. Huang, J.-Q.; Fang, X.; Tian, X.; Chen, P.; Lin, J.-L.; Guo, X.-X.; Li, J.-X.; Fan, Z.; Song, W.-M.; Chen, F.-Y.; et al. Aromatization of natural products by a specialized detoxification enzyme. Nat. Chem. Biol. 2020, 16, 250–256. [CrossRef] [PubMed]

86. Cao, J.-F.; Huang, J.-Q.; Liu, X.; Huang, C.-C.; Zheng, Z.-S.; Zhang, X.-F.; Shangguan, X.-X.; Wang, L.-J.; Zhang, Y.-G.; Wendel, J.F.; et al. Genome-wide characterization of the GRF family and their roles in response to salt stress in Gossypium. BMC Genom. 2020, 21, 1–16. [CrossRef] [PubMed]

87. Zhang, X.; Cao, J.; Huang, C.; Zheng, Z.; Liu, X.; Shangguan, X.; Wang, L.; Zhang, Y.; Chen, Z. Characterization of cotton ARF factors and the role of GhARF2 in fiber development. BMC Genom. 2021, 22, 1–15. [CrossRef]

88. Cui, W.; Chen, Z.; Shangguan, X.; Li, T.; Wang, L.; Xue, X.; Cao, J. TRY intron2 determined its expression in inflorescence activated by SPL9 and MADS-box genes in Arabidopsis. Plant Sci. 2022, 321, 111311. [CrossRef]
89. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* 2001, 25, 402–408. [CrossRef]

90. Cao, J.-F.; Zhao, B.; Huang, C.-C.; Chen, Z.-W.; Zhao, T.; Liu, H.-R.; Hu, G.-J.; Shangguan, X.-X.; Shan, C.-M.; Wang, L.-J.; et al. The miR319-targeted GhTCP4 promotes the transition from cell elongation to wall thickening in cotton fiber. *Mol. Plant* 2020, 13, 1063–1077. [CrossRef]

91. Bent, A. *Arabidopsis thaliana* Floral Dip Transformation Method. *Methods Mol. Biol.* 2006, 343, 87–104. [CrossRef]

92. Shangguan, X.; Yang, Q.; Wu, X.; Cao, J. Function analysis of a cotton R2R3 MYB transcription factor GhMYB3 in regulating plant trichome development. *Plant Biol.* 2021, 23, 1118–1127. [CrossRef]

93. Chen, Z.; Zhao, J.; Song, J.; Han, S.; Du, Y.; Qiao, Y.; Liu, Z.; Qiao, J.; Li, W.; Li, J.; et al. Influence of graphene on the multiple metabolic pathways of Zea mays roots based on transcriptome analysis. *PLoS ONE* 2021, 16, e0244856. [CrossRef]

94. Chen, Z.; Zhao, J.; Qiao, J.; Li, W.; Guan, Z.; Liu, Z.; Bai, X.; Xing, B.; Zhang, J.; Li, J.; et al. Graphene-Mediated Antioxidant Enzyme Activity and Respiration in Plant Roots. *ACS Agric. Sci. Technol.* 2022, 2, 646–660. [CrossRef]

95. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014, 15, 550. [CrossRef]

96. Mao, X.; Cai, T.; Olyarchuk, J.G.; Wei, L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 2005, 21, 3787–3793. [CrossRef]