the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2024) **120,** 505–525 doi: 10.1111/tpj.16996

Evolution and post-transcriptional regulation insights of m^{6}A writers, erasers, and readers in plant epitranscriptome

Jun Zhang[1,†], Lin Wu[2,†], Lele Mu[2,†], Yuhua Wang[1], Mengna Zhao[3], Huiyuan Wang[2], Xiangrong Li[3], Liangzhen Zhao[1], Chentao Lin[3], Hangxiao Zhang[3] and Lianfeng Gu[3,*]

[1]*Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China,*

[2]*College of Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, China, and*

[3]*Basic Forestry and Proteomics Research Center, School of Future Technology, Fujian Agriculture and Forestry University, Fuzhou 350002, China*

Received 14 May 2024; revised 30 June 2024; accepted 9 August 2024; published online 21 August 2024.
*For correspondence *(e-mail lfgu@fafu.edu.cn).
†These authors contributed equally to this work.

SUMMARY

As a dynamic and reversible post-transcriptional marker, *N*^{6}-methyladenosine (m^{6}A) plays an important role in the regulation of biological functions, which are mediated by m^{6}A pathway components including writers (MT-A70, FIP37, VIR and HAKAI family), erasers (ALKBH family) and readers (YTH family). There is an urgent need for a comprehensive analysis of m^{6}A pathway components across species at evolutionary levels. In this study, we identified 4062 m^{6}A pathway components from 154 plant species including green algae, utilizing large-scale phylogenetic to explore their origin and evolution. We discovered that the copy number of writers was conserved among different plant lineages, with notable expansions in the ALKBH and YTH families. Synteny network analysis revealed conserved genomic contexts and lineage-specific transpositions. Furthermore, we used Direct RNA Sequencing (DRS) to reveal the Poly(A) length (PAL) and m^{6}A ratio profiles in six angiosperms species, with a particular focus on the m^{6}A pathway components. The ECT1/2-Poeaece4 sub-branches (YTH family) with unique genomic contexts exhibited significantly higher expression level than genes of other ECT1/2 poeaece sub-branches (ECT1/2-Poeaece1-3), accompanied by lower m^{6}A modification and PAL. Besides, conserved m^{6}A sites distributed in CDS and 3′UTR were detected in the ECT1/2-Poaceae4, and the dual-luciferase assay further demonstrated that these conserved m^{6}A sites in the 3′UTR negatively regulated the expression of Firefly luciferase (LUC) gene. Finally, we developed transcription factor regulatory networks for m^{6}A pathway components, using yeast one-hybrid assay demonstrated that *PheBPC1* could interact with the *PheECT1/2-5* promoter. Overall, this study presents a comprehensive evolutionary and functional analysis of m^{6}A pathway components and their modifications in plants, providing a valuable resource for future functional analysis in this field.

Keywords: *N*^{6}-methyladenosine, m^{6}A pathway components, phylogenetic, synteny network, RNA methylation, poly(A) tail length.

INTRODUCTION

N^6-Methyladenosine (m^{6}A) is the most prevalent internal chemical modified nucleotide, widely existing in rRNA, mRNA, tRNA, miRNA, and lncRNA (Jia et al., 2013). As a dynamic and reversible post-transcriptional marker, m^{6}A plays a vital role in the regulation of various life activities (Arribas-Hernández & Brodersen, 2020; Yue et al., 2019). Most of the biological functions of m^{6}A are mediated by m^{6}A pathway components which include writers, erasers, and readers (Yue et al., 2019). Writers install the m^{6}A to mRNA, whereas erasers remove the m^{6}A from the mRNA. The readers bind specifically to m^{6}A-modified RNA to perform the biological function of RNA methylation. The writer, eraser, and reader form a complex regulatory system that guides the formation, deletion, and decoding of m^{6}A (Yue et al., 2019).

Deposition of m^{6}A to RNAs is controlled by the m^{6}A writer proteins (Yan et al., 2022). In Arabidopsis, the writer components include MTA, MTB, FIP37, VIR, and HAKAI. As the earliest writer component discovered in plants, the

MT-A70 family can be divided into three subfamilies in higher eukaryotes: MTA (human METTL3 homologous protein), MTB (human METTL14 homologous protein), and MTC (human METTL4 homologous protein) (Bujnicki et al., 2002; Iyer et al., 2016; Liang et al., 2020). METTL3 and METTL14 interact to form a stable core heterodimer that catalyzes the methylation of m^6A at specific locations in mRNA (Bokar et al., 1994, 1997; Wang et al., 2016). Up to now, there are few studies on the specific role of MTC in m^6A modification. It has been suggested that METTL4 may be a DNA methylase (Greer et al., 2015) or a U2 snRNA MTase (Luo et al., 2022). However, further investigations are needed to elucidate its precise role in m^6A processes of plants. In addition to the MT-A70 family, writer complexes also include FIP37 (human WTAP homologous protein), VIR (human VIRMA homologous protein), and HAKAI (human HAKAI homologous protein). A recent study showed that WTAP acts as a bridge facilitating the connecting between VIRMA and METTL3-METTL14. VIRMA serves as a platform for the binding of other regulatory proteins. METTL3-METTL14 operates as a catalytic pocket toward VIRMA, poised for the arrival of the substrate (Yan et al., 2022). HAKAI can interact with other m^6A methyltransferase complex members, and the absence of HAKAI leads to a significant reduction in m^6A, thus identifying it as another m^6A methyltransferase component (Růžička et al., 2017).

The first reported erasers were FTO (Jia et al., 2011) and ALKBH5 (Zheng et al., 2013) in animals, members of the ALKBH (ALKB homolog) subfamily in the Fe (II)/2-oxoglutarate (2OG) dioxygenase superfamily. The ALKBH family consists of 14 members in Arabidopsis: ALKBH1A-D, ALKBH2, ALKBH6, ALKBH8A-B, ALKBH9A-C, and ALKBH10A-C, which are divided into seven main clades (Liang et al., 2020). So far, m^6A demethylases identified in plants include ALKBH9B (Martínez-Pérez et al., 2017), ALKBH10B (Duan et al., 2017), and *SlALKBH2* (Zhou et al., 2019). Knockout of ALKBH9B has been reported to affect the infection capacity of Mosaic virus (Martínez-Pérez et al., 2017) and to cause sensitivity to ABA treatment during seed germination and early seedling development (Tang et al., 2022). ALKBH10B has been reported to regulate flowering transition (Duan et al., 2017), and its mutants were highly sensitive to ABA osmotic stress and salt stress during seed germination (Tang et al., 2021). *SlALKBH2* was reported to be involved in regulating tomato ripening (Zhou et al., 2019).

Proteins containing the YTH domain are reported to be m^6A readers. YTH family members are highly conserved and contain a YTH domain with an aromatic cage for m^6A recognition (Yue et al., 2019). In humans, there are five members of the YTH family (Meyer & Jaffrey, 2017), whereas in plants, the number of YTH family members has significantly expanded. Previous studies categorized YTH proteins into two main groups: YTHDF and YTHDC. Moreover, the YTHDF group is subdivided into YTHDFa, YTHDFb, and YTHDFc, whereas the YTHDC group is further categorized into YTHDCa and YTHDCb (Scutenaire et al., 2018). Arabidopsis has 13 members, including ECT1-12 and CPSF30 (Yue et al., 2019). ECT2/ECT3/ECT4 have been reported to be associated with the number of trichome branches, leaf formation time, and leaf morphogenesis (Arribas-Hernández et al., 2018; Scutenaire et al., 2018; Wei et al., 2018). Additionally, CPSF30-L is required for flower transformation and ABA response (Song et al., 2021).

The direct RNA sequencing (DRS) technique has the potential to detect RNA methylation (Liu et al., 2019; Parker et al., 2020), and the development of multiple computational tools (Zhong et al., 2023) allows us to identify m^6A sites at the single-base. In addition to RNA methylation, DRS can also identify poly(A) tail length (PAL) for poly(A)$^+$ RNA (Gao et al., 2022; Li et al., 2023; Liufu et al., 2023; Parker et al., 2020). Poly(A) tail plays an important role in maintaining mRNA stability, improving translation efficiency, and facilitating cytoplasmic localization (Beilharz & Preiss, 2007; Passmore & Coller, 2022; Piqué et al., 2008). Thus, DRS technique can provide a comprehensive profile of RNA methylation and PAL of m^6A pathway components.

In this study, we performed genome-wide identification of m^6A writers, erasers, and readers across 154 plant genomes, combining phylogenetic and synteny network analyses, to investigate their origin and evolution in plants. This enriches our understanding of m^6A pathway components in green plants, especially angiosperms. Additionally, we obtained PAL and RNA methylation profiles in six angiosperms based on DRS technology, further exploring the post-transcriptional regulation mode of m^6A pathway components. In conclusion, this study provides comprehensive post-transcriptional regulation and phylogenetic analysis for understanding the function of m^6A pathway components in plant.

RESULTS

Identification of m^6A writers, erasers, and readers in 154 genomes for comprehensive profiling of evolutionary trends and expansion patterns

To comprehensively and systematically identify the family members of m^6A writers, erasers, and readers in plants, we selected 154 species from 13 chlorophytes, 7 charophytes, 7 bryophytes, 1 lycophytes, 3 ferns, 9 gymnosperms, and 114 angiosperms for identification (Figure S1; Table S1). In summary, through homology searches and domain predictions, we identified 4062 m^6A pathway components candidate protein sequences from 154 species, including 512 MT-A70s, 200 FIP37s, 145 VIRs, 148 HAKAIs, 1615 ALKBHs, and 1496 YTHs (Table S2). The copy number of m^6A writer components, including MT-A70s, FIP37s, VIRs, and HAKAIs, was conserved among different land plant lineages (Figure 1; Table S3), suggesting that the

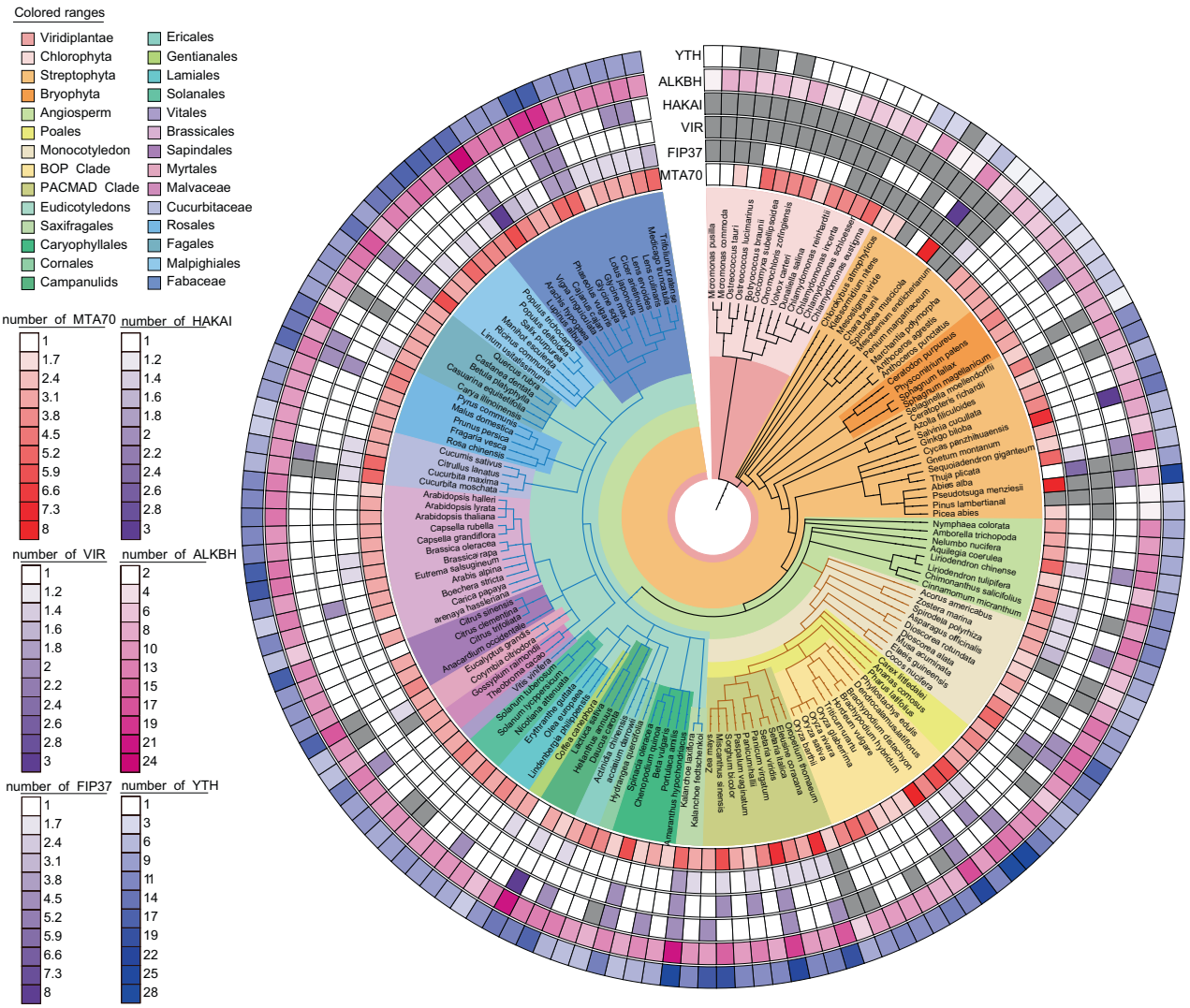


Figure 1. Number of writer, eraser and reader family members in 154 species.
The intensity of color represents the copy number, whereas the gray block indicates zero gene count. From the outermost circle to the inner circle, it represents YTH, ALKBH, HAKAI, VIR, FIP37, and MTA70, respectively. The species taxonomy relationship was depicted in the innermost circle.

mechanism for writing m^{6}A in land plants may be conserved. Notably, FIP37, VIR, and HAKAI were not found in mamiellophyceae (marine algae) (Figure 1). The copy number of ALKBHs showed a slight expansion, whereas the YTHs showed a significant expansion among the different plant lineages (Figure 1; Table S3). m^{6}A influences plant growth and metabolism by recruiting reader proteins (Yue et al., 2019), this phenomenon may be explained by the fact that higher plant growth metabolism is more complex and requires more m^{6}A readers to participate in regulation.

Phylogenetic insights into m^{6}A pathway components: Clade divergence, expansion, and ancestral models

To better understand the evolutionary history of m^{6}A pathway components, we constructed phylogenetic trees using dataset1 (including 133 MT-A70s, 312 ALKBHs, and 227 YTHs from 40 non-angiosperms and 4 angiosperms) and dataset2 (including 393 MT-A70s, 1294 ALKBHs, and 1315 YTHs from 114 angiosperms), respectively (Figure 2a; Table S4). For MT-A70s, the phylogenetic results of MT-A70-dataset1 showed that the MT-A70 family had five main clades in plants, namely MTA, MTB, MTC, Chlorophytes-specific1, and Chlorophytes-specific2 (Figure 2b). The MTA, MTB, and MTC clades included not only land plants but also chlorophytes and charophytes (Figure 2b), suggesting that the divergence of these three clades already occurred in the common ancestor of algae and land plants. We reconstructed the phylogenetic tree by adding animal MT-A70s and the results showed that animal MT-A70s and plant MT-A70s were also classified into MTA, MTB, and

MTC (Figure S2), suggesting that the divergence of these three clades already occurred in the common ancestor of animals and plants. Chlorophytes-specific1 and Chlorophytes-specific2 contain only chlorophytes (Figure 2b; Figure S2) and there might be a loss of these two clades in charophytes and land plants.

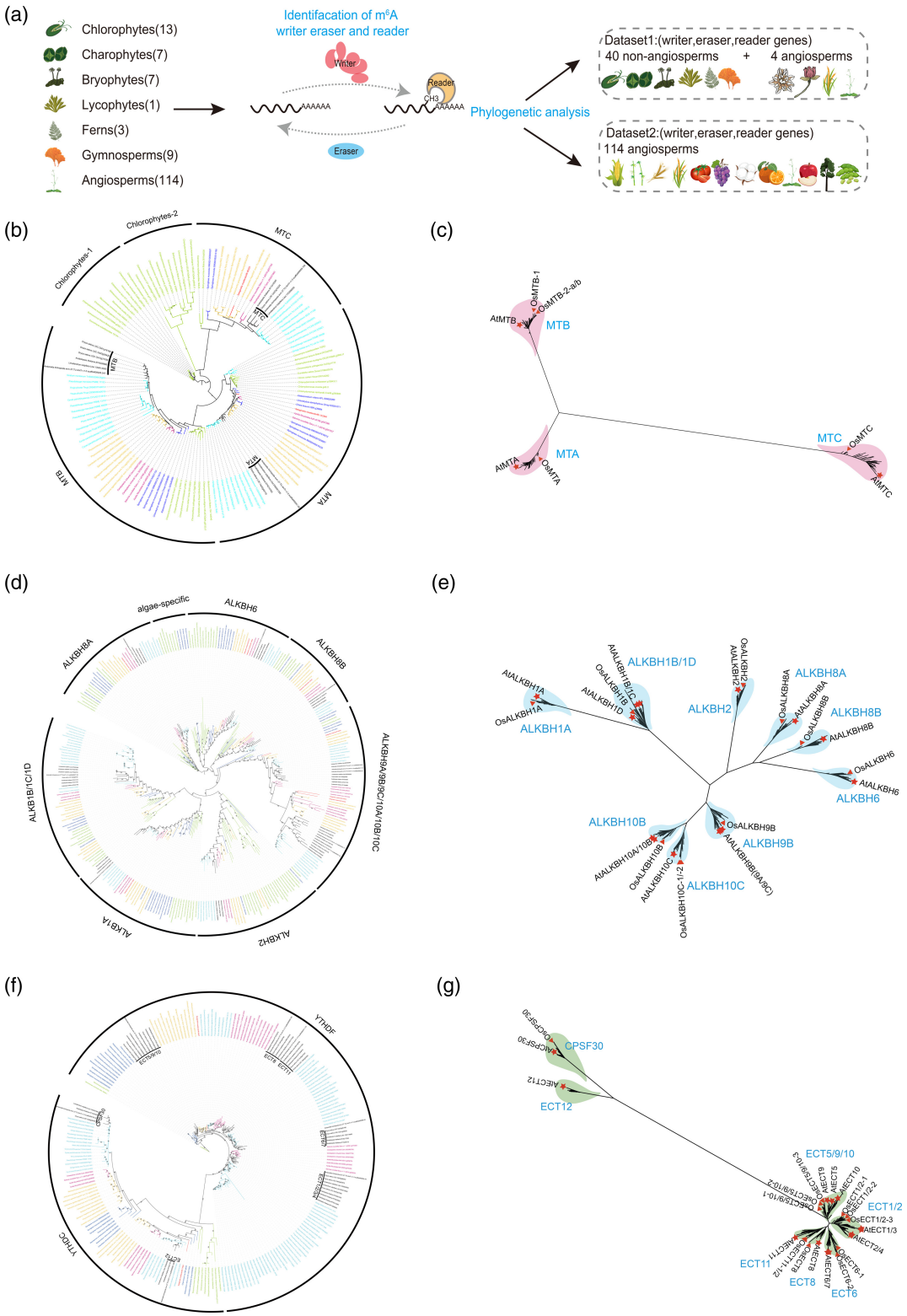


Figure 2. Phylogenetic analysis of m^6A pathway components.
(a) Flowchart for phylogenetic analysis of m^6A pathway components.
(b) The phylogenetic tree of MT-A70s in green plants was constructed using MT-A70-dataset1, which includes 133 MT-A70 proteins from 4 angiosperms (outlined in black) and 40 non-angiosperms including 13 chlorophytes (depicted in green), 7 charophytes (indicated in blue), 7 bryophytes (highlighted in yellow), 1 lycophyte (shown in red), 3 ferns (displayed purple) and 9 gymnosperms (illustrated in light blue).
(c) The phylogenetic classification of MT-A70s in angiosperms was constructed using MT-A70-dataset2, which includes 393 MT-A70 proteins from 114 angiosperms.
(d) The phylogenetic tree of ALKBHs in green plants was constructed using ALKBH-dataset1, which includes 212 ALKBH proteins from 4 angiosperms and 40 non-angiosperms.
(e) The phylogenetic classification of ALKBHs in angiosperms was constructed using ALKBH-dataset2, which includes 1294 ALKBH proteins from 114 angiosperms.
(f) The phylogenetic tree of YTHs in green plants was constructed using YTH-dataset1, which includes 227 YTH proteins from 4 angiosperms and 40 non-angiosperms. The YTHDF and YTHDC clases were named according to Scutenaire et al. (2018).
(g) Phylogenetic classification of YTHs in angiosperms was constructed using YTH-dataset2, which includes 1315 YTH proteins from 114 angiosperms. The ECT1/2, ECT5/9/10, ECT6, ECT8, ECT11, CPSF30 and ECT12 clades were named by 13 Arabidopsis YTHs to better demonstrate the expansion of YTH family in angiosperms. CPSF30 used to be named YTHDCa; ECT12 used to be named YTHDCb; ECT1/2 used to be named YTHDFa; ECT5/9/10 used to be named YTHDFb; ECT6, ECT8 and ECT11 used to be divided into the same branch, YTHDFc (Scutenaire et al., 2018).

The phylogenetic results of MT-A70-dataset2 revealed that angiosperms MT-A70s were clustered into three major groups: MTA, MTB, and MTC (Figure 2c). In the MTA and MTC clades, most diploid plants had only one copy, whereas in the MTB clade, we observed expansion within monocots (Figure S3). Further analysis of the MTB clade showed that this multi-copies phenomenon occurred only in Poaceae, while there was still only single copy in other members of the order Poales such as *Ananas comosus* (Figure S4), suggesting that the divergence of Poaceae-Group1/2 (MTB-Poaceae1/2) occurred after the emergence of the order Poales but at least before the emergence of Poaceae. The differentiation of *OsMTB1* and *OsMTB2-a/b* resulted from this gene duplication, followed by gene duplication within the genus Oryza resulting in *OsMTB2-a* and *OsMTB2-b* (Figure S4).

For ALKBHs, the phylogenetic analysis of ALKBH-dataset1 revealed eight major clades within plant: ALKBH1A, ALKBH1B/1C1D, ALKBH2, ALKBH6, ALKBH8A, ALKBH8B, ALKBH9A/9B/9C/10A/10B/10C, and algae-specific (Figure 2d). In this study, except for the algae-specific clade, which included only chlorophytes and charophytes, the remaining seven clades encompassed not only land plants but also chlorophytes and charophytes (Figure 2d), suggesting that divergence of this seven clades occurred in the common ancestor of algae and land plants. Within the ALKBH9A/9B/9C/10A/10B/10C clade, divergence into ALKBH9A/9B/9C and ALKBH10A/10B/10C was observed in seed plants (gymnosperms and angiosperms), but only one copy exists in chlorophytes and charophytes (Figure 2d), suggesting that differentiation of ALKBH9A/9B/9C and ALKBH10A/10B/10C occurred after the emergence of charophytes but before the emergence of seed plants. We also added animal ALKBHs to reconstruct the phylogenetic tree. The results showed that ALKBH2 and algae-specific branches correspond to two copies of animal ALKBHs. ALKBH1A, ALKBH6, ALKBH8A, and ALKBH9A/9B/9C/10A/10B/10C branches correspond to one copy of animal ALKBHs. No copy of animal ALKBHs was found in ALKBH1B/1C/1D and ALKBH8B branches (Figure S5).

The phylogenetic results of ALKBH-dataset2 revealed nine major clades in angiosperms: ALKBH1A, ALKBH1B/1D, ALKBH2, ALKBH6, ALKBH8A, ALKBH8B, ALKBH9B, ALKBH10B, and ALKBH10C (Figure 2e). ALKBH9A/9B/9C/10A/10B/10C can be further divided into ALKBH9A/9B/9C, ALKBH10A/10B, and ALKBH10C. ALKBH10A and 10B were lineage-specific pairs within Brassicaceae (Figure S6a), so ALKBH10A/10B was recorded as ALKBH10B. It can be concluded that most diploid plants have only one gene corresponding to Arabidopsis ALKBH10A and 10B. A similar situation also occurred in ALKBH1B and 1C, ALKBH9A, 9B and 9C, all of which were lineage-specific paralog pairs within Brassicaceae, so their corresponding clades were denoted as ALKBH1B and ALKBH9B, respectively (Figure S6b,c). ALKBH1A, ALKBH2, ALKBH6, ALKBH8A, and ALKBH8B correspond to only one copy in most angiosperms species, indicating that they did not undergo further differentiation (Figure S7). However, ALKBH1B/1D, ALKBH9B, ALKBH10B, and ALKBH10C were observed to have multiple copies (Figure S7). Among them, multiple copies of ALKBH10B (Figure S8) and ALKBH10C (Figure S9) were due to a large number of lineage-specific paralog pairs of different orders, families, or genera, such as Fabales and Cucurbitales.

For YTHs, the phylogenetic results of YTH-dataset1 indicated that the YTH family was divided into two clades: YTHDF and YTHDC (Figure 2f). To further reconstruct the phylogenetic tree, we included animal YTHs (YTHDF1-3 and YTHDC1-2) and observed that the human YTHDF1-3 were clustered together with YTHDF branches, and the human YTHDC1-2 were clustered together with YTHDC branches (Figure S10). The YTHDF and YTHDC clades encompassed not only land plants but also chlorophytes and charophytes (Figure 2f), suggesting that the divergence of two clades was already occurred in the common ancestor of algae and land plants. In the YTHDC clade, the

divergence of CPSF30 (formerly named YTHDCa) and ECT12 (formerly named YTHDCb) was observed in charophytes and land plants, while only one copy exists in chlorophytes (Figure 2f), suggesting that differentiation of CPSF30 and ECT12 occurred in the common ancestor of charophytes and land plants. In the YTHDF clade, we observed the formation of five clades in angiosperms: ECT1/2/3/4 (formerly named YTHDFa), ECT5/9/10 (formerly named YTHDFb), ECT6/7 (part of YTHDFc), ECT8 (part of YTHDFc), and ECT11 (part of YTHDFc) (Figure 2f,g). However, chlorophytes and charophytes have only one copy of the YTHDF clade (Figure 2f), suggesting that differentiation of the five branches occurred after the appearance of charophytes but at least before the emergence of angiosperms.

The phylogenetic results of YTH-dataset2 showed that angiosperms YTHs were mainly divided into seven clades: ECT1/2, ECT5/9/10, ECT6, ECT8, ECT11, ECT12, and CPSF30 (Figure 2g). The YTHDC clade can be further divided into ECT12 and CPSF30, and the YTHDF clade can be further divided into ECT1/2, ECT5/9/10, ECT6, ECT8, and ECT11. Similar to the ALKBH family, ECT1 and ECT3 (Figure S11a), ECT2 and ECT4 (Figure S11b), ECT6 and ECT7 (Figure S11c) in the YTH family were also found to be lineage-specific paralog pairs within the Brassicaceae, so their corresponding clades were denoted as ECT1, ECT2, and ECT6, respectively.

In this study, we found that ECT1/2 and ECT5/9/10 clades have multiple copies in many species (Figure S12). Our study showed that ECT1/2 forms two sub-branches: Eudicots-Group1 (ECT1/2-Eudicots1) and Eudicots-Group2 (ECT1/2-Eudicots 2) in Eudicots (Figure S13). Further analysis of these two sub-branches showed that genes of the Brassicales were presented in both ECT1/2 branches. Among them, ECT1 and ECT3 differentiated from gene duplication in the order Brassicales (Figure S11a), whereas ECT2/4 differentiated from gene duplication in the Brassicaceae (Figure S11b). In addition to Brassicales, some dicotyledonous plants, such as Fagales, Sapindales, Fabales, and Malpighiales were clustered in both ECT1/2 branches. However, Cucurbitales and Rosales only contain one ECT1/2 branches (Figure S13). In monocotyledonous plants, we observed that all the gene members of the Poaceae family form four subclades: Poaceae-Group1-4 (ECT1/2-Poaceae1-4), with the loss of *Oryza sativa* gene in the ECT1/2-Poaceae2 clade (Figure S13). ECT5/9/10 forms three sub-branches in dicotyledonous plants: Eudicots-Group 1 (ECT5/9/10-Eudicots1), Eudicots-Group 2 (ECT5/9/10-Eudicots2), and Eudicots-Group 3 (ECT5/9/10-Eudicots3), each corresponding to an Arabidopsis gene (Figure S14). In the ECT5/9/10-Eudicots1-3 sub-clades, there are orders with genes present in all three sub-clades, such as Brassicales and Malpighiales; in only two sub-clades, such as Rosales and Fagales; or in only one subclade, such as Fabales and Cucurbitales (Figure S14). In monocotyledonous plants, a significant expansion of the gene members of the Poaceae was also observed, dividing them into three sub-clades: Poaceae group1-3 (ECT5/9/10-Poaceae1-3) (Figure S14). ECT6 also forms three subclades in the Poaceae including Poaceae-Group1-3 (ECT6-Poaceae1-3) (Figure S15). Additionally, ECT12 was identified in basal angiosperms and eudicots but not in monocots (Figure S12), and we speculate that ECT12 was lost in monocots.

Synteny network analysis of m^{6}A pathway components

Conserved genomic context information provides crucial insights for comparative evolutionary analysis (Zhao et al., 2017). Through clustering and phylogenetic analysis, synteny network method can identify lineage-specific transposition events and provide synteny conserved evidence for these genes, thereby suggesting important conserved gene functions (Zhao et al., 2017). Given the great evolutionary distance between non-angiosperms and angiosperms, we analyzed 114 angiosperms to perform the synteny network analysis of MTA70, FIP37, VIR, HAKAI, ALKBH, and YTH gene family, shedding light on the synteny of m^{6}A pathway components in angiosperms. In these networks, nodes represent genes and edges represent a syntenic relationship between them, genes with more syntenic relationships are more likely to be clustered into the same cluster. We identified 20, 12, 9, 4, 48, and 34 clusters in the MTA70, FIP37, VIR, HAKAI (Figure 3a–d; Figures S16 and S17), ALKBH (Figure 3e; Figures S18 and S19), and YTH families (Figure 3f; Figures S20 and S21), respectively (Table S5). In the MTA70 family, clusters1 (MTA) and 2 (MTC) comprised genes from monocots, eudicots, and basal angiosperms, representing angiosperm-conserved clusters (Figure 3a; Figures S16 and S17a). Cluster3 (MTB) was primarily composed of monocots genes, making it a monocot-conserved cluster (Figure 3a; Figures S16 and S17a). Cluster4 (MTB) had a prevalence of genes from eudicots, categorizing it as a eudicot-conserved cluster (Figure 3a; Figures S16 and S17a). Other clusters in the MTA70 family tended to encompass genes from specific order, family, or genus, designating them as lineage-specific clusters (Figure 3a; Figures S16 and S17a). For FIP37, VIR, and HAKAI families, FIP37-cluster2 (Figure 3b; Figures S16 and S17b) was an angiosperm-conserved cluster, whereas VIR-cluster7 (Figure 3c; Figures S16 and S17c) and HAKAI-cluster1 (Figure 3d; Figures S16 and S17d) were monocot-conserved cluster, VIR-cluster2 (Figure 3c; Figures S16 and S17c) and HAKAI-cluster2 (Figure 3d; Figures S16 and S17d) was a eudicot-conserved cluster. In the ALKBH family, angiosperm-conserved clusters included cluster3 (ALKBH8A), 6 (ALKBH10B), 9 (ALKBH1B/1D), 10 (ALKBH10C), 11 (ALKBH1A), 13 (ALKBH8B), and 15 (ALKBH2). Monocot-conserved clusters included cluster 20 (ALKBH9B), 25 (ALKBH9B), whereas eudicot-conserved clusters included cluster1 (ALKBH9B), 4 (ALKBH1B/1D), and

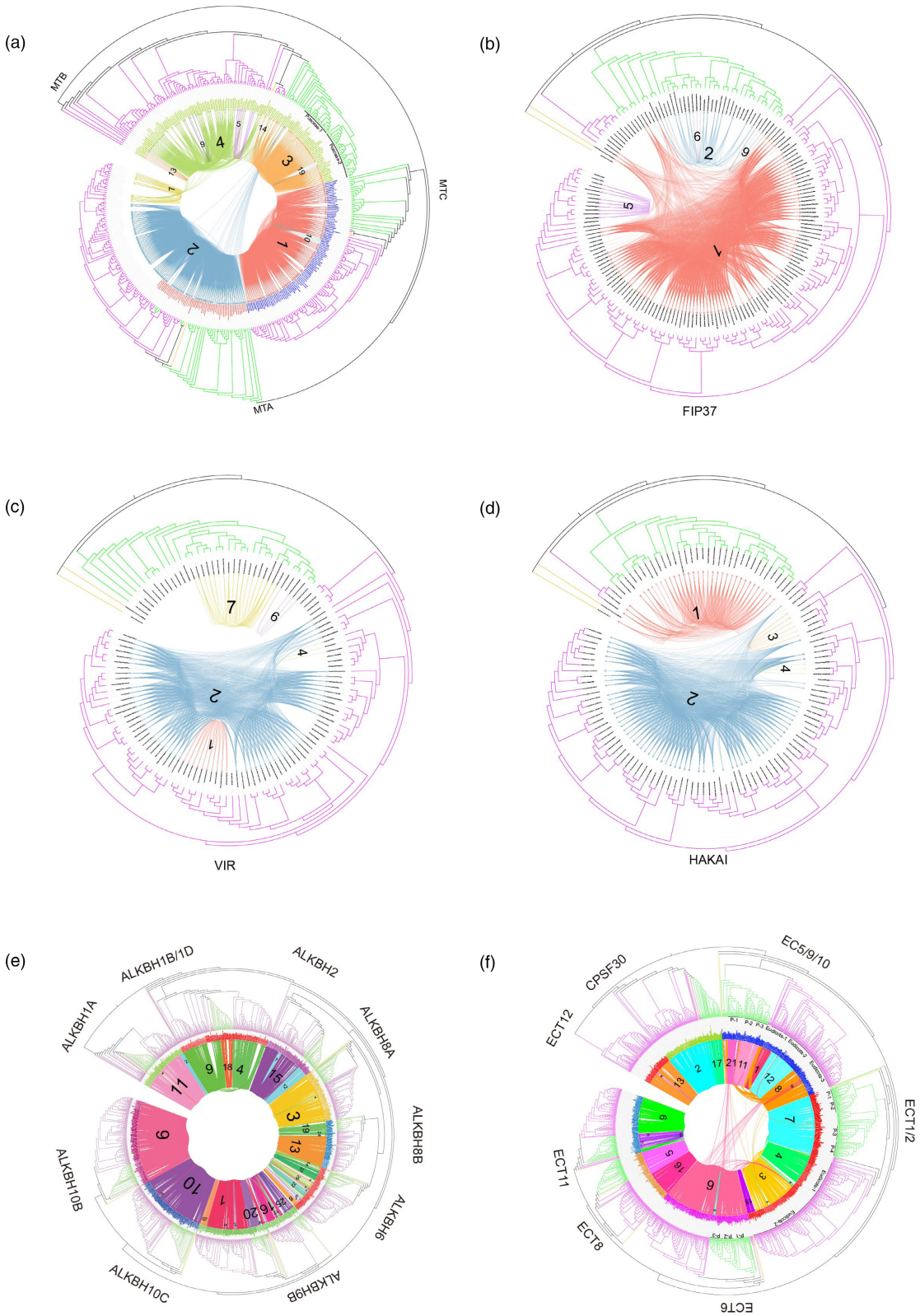


Figure 3. Phylogenetic and syntenic relationships of (a) MT-A70s, (b) FIP37s, (c) VIRs, (d) HAKAIs, (e) ALKBHs, and (f) YTHs in angiosperms.
The lines represent syntenic relationship between two genes, with different colors representing different clusters. Cluster ids were marked on the corresponding line.

16 (ALKBH9) (Figure 3e; Figures S18 and S19). Within the YTH family, angiosperm-conserved cluster included cluster6 (ECT6). Monocot-conserved clusters encompassed cluster17 (CPSF30), 5 (ECT8), 10 (ECT11), 7 (ECT1/2), 11 (ECT5/9/10), 21 (ECT5/9/10), whereas eudicot-conserved clusters included cluster2 (CPSF30), 13 (ECT12), 16 (ECT8), 9 (ECT11), 3 (ECT1/2), 4 (ECT1/2), 1 (ECT5/9/10), 8 (ECT5/9/10), and 12 (ECT5/9/10) (Figure 3f; Figures S20 and S21).

In this study, FIP37 genes of eudicots, excluding the Brassicaceae family, clustered together in the FIP37-cluster1, whereas Brassicaceae FIP37 genes formed a distinct cluster, FIP37-cluster5 (Figure 3b; Figures S16 and S17b). Genomicus parallel coordinate plot (Nguyen et al., 2022) was utilized to confirm this pattern by studying the synteny relationship between *Vitis vinifera* (VIT_17s0000g08940) and *Arabidopsis thaliana* (AT3G54170). No synteny was observed when searching for any Brassicaceae FIP37 homologous gene using *V. vinifera* FIP37 homologous gene (VIT_17s0000g08940) (Figure S22a). However, a unique synteny pattern between FIP37 genes in Brassicaceae and other eudicots was detected when using the *A. thaliana* FIP37 gene (AT5G20240) for the search (Figure S22b). The unique genomic context of the Brassicaceae FIP37 gene may be due to gene transposition, genomic rearrangement event, or extreme genome fractionation (Zhao et al., 2017). This suggests that Brassicaceae FIP37 might have a unique gene function that was different from other FIP37 homologous genes. Similar observations were made in VIR (Figure 3c; Figures S16 and S17, VIR cluster1 and 2), ALKBH1A (Figure 3d; Figures S18 and S19, cluster2 and 11), ALKBH2 (Figure 3d; Figures S18 and S19, cluster7 and 15).

The synteny networks method can also be used to identify ancient tandem duplications (Zhao et al., 2017). However, we did not identified ancient tandem duplications similar to those found in the MADS-box (Zhao et al., 2017), CLKB (Zhang et al., 2020), and LEA (Artur et al., 2019) families. We did detect tandem duplications in the MTA70, FIP37, ALKBH, and YTH families. However, these tandem duplications occurred after emergence of angiosperms, resulting in an increase in gene copy number within each order, family, or genus (Figure S23a–d; Table S6). This indicates that although there were no ancient tandem duplications in m^{6}A pathway components, tandem duplications still play a certain role in the expansion of m^{6}A pathway components gene family in some species.

Our phylogenetic analysis of the MTA70 and YTH families showed that the Poaceae gene family in the MTB branch expanded, forming two branches: MTB-Poaceae1-2. Similarly, in the ECT1/2 branch, the genes expanded to form four branches: ECT1/2-Poaceae1-4. However, in synteny network analysis, MTB (Figure 3a; Figures S16 and S17, MT-A70-cluster3) and ECT1/2 (Figure 3F; Figures S20 and S21, cluster7) of all monocots were found all in the same cluster, respectively. To further clarify the synteny between the branches of MTB-Poaceae1-2 and between the branches of ECT1/2-Poaceae1-4, the synteny between MT-A70-cluster3 and ECT1/2-cluster7 are shown in Figure S24 according to the phylogenetic classification. We observed that the collinearity between MTB-Poaceae1 and MTB-Poaceae2 was very tight (Figure S24a). We speculate the expansion of Poaceae MTB gene may be due to an ancient τ-whole genome duplication event in monocots (Jiao et al., 2014). For the ECT1/2 branch, ECT1/2-Poaceae1 and ECT1/2-Poaceae2 were tightly connected. ECT1/2-Poaceae3 showed moderately connectivity to ECT1/2-Poaceae1 and 2, and ECT1/2-Poaceae4 was loosely connected to the other three sub-branches (Figure S24b), which may be due to the fact that after the formation of ECT1/2-Poaceae1-4 sub-branches, the surrounding regions of each sub-branch gene experienced varying degrees of gene loss, resulting in changes in their synteny relationships.

The profile of poly(A) tail length and m^{6}A modifications via direct RNA sequencing in *Zea mays*, *Oryza* sativa, *Dendrocalamus latiflorus*, and three other angiosperms

To explore the PAL and m^{6}A profiles using DRS, we selected six angiosperms: two eudicots, *A. thaliana* (Brassicales) and *Populus trichocarpa* (Malpighiales); and four monocots, *Zea mays* (Poaceae, PACMAD clade), *O. sativa* (Poaceae, BOP clade), *Phyllostachys edulis* (Poaceae, BOP clade), and *Dendrocalamus latiflorus* (Poaceae, BOP clade). DRS libraries from *Z. mays*, *O. sativa*, and *D. latiflorus* were sequenced in this study (Table S7), and the other three DRS libraries were obtained from previously published studies (Fu et al., 2014; Gao et al., 2022; Li et al., 2023; Liufu et al., 2023; Miao et al., 2022; Parker et al., 2020) (Table S8). The PAL distribution of transcripts exhibited a peak at ~40 nt, and the median distribution of gene PAL concentrated between 50 and 100 nt (Figure 4a). In addition, the median PAL transcripts of *Z. mays* (56.84 nt) and *O. sativa* (60.1 nt) measured in this study was slightly shorter than that of other species (*A. thaliana*: 67.97 nt, *P. trichocarpa*: 72.72 nt, *P. edulis*: 70.21 nt, and *D. latiflorus*: 74.48 nt) (Figure S25a), and these two species also exhibited relatively shorter median gene PAL (Figure 4a).

We identified 46 833, 40 782, 27 124, 33 930, 42 092, and 28 583 m^{6}A modification sites from 7735, 8410, 6872, 9014, 8303 and 8274 genes in *A. thaliana*, *P. trichocarpa*, *Z. mays*, *O. sativa*, *P. edulis*, and *D. latiflorus*, respectively. The m^{6}A ratio for each modification site was concentrated in the range of 0.2–0.4, whereas the median m^{6}A ratio for genes was prominent around 0.5 (Figure 4a). The median modification rates varied slightly across, with values of 36.3% (*A. thaliana*), 37.8% (*P. trichocarpa*), 38% (*Z. mays*), 39.2% (*O. sativa*), 41.2% (*P. edulis*), and 43.8% (*D. latiflorus*) (Figure S25b). We further explore the potential correlation

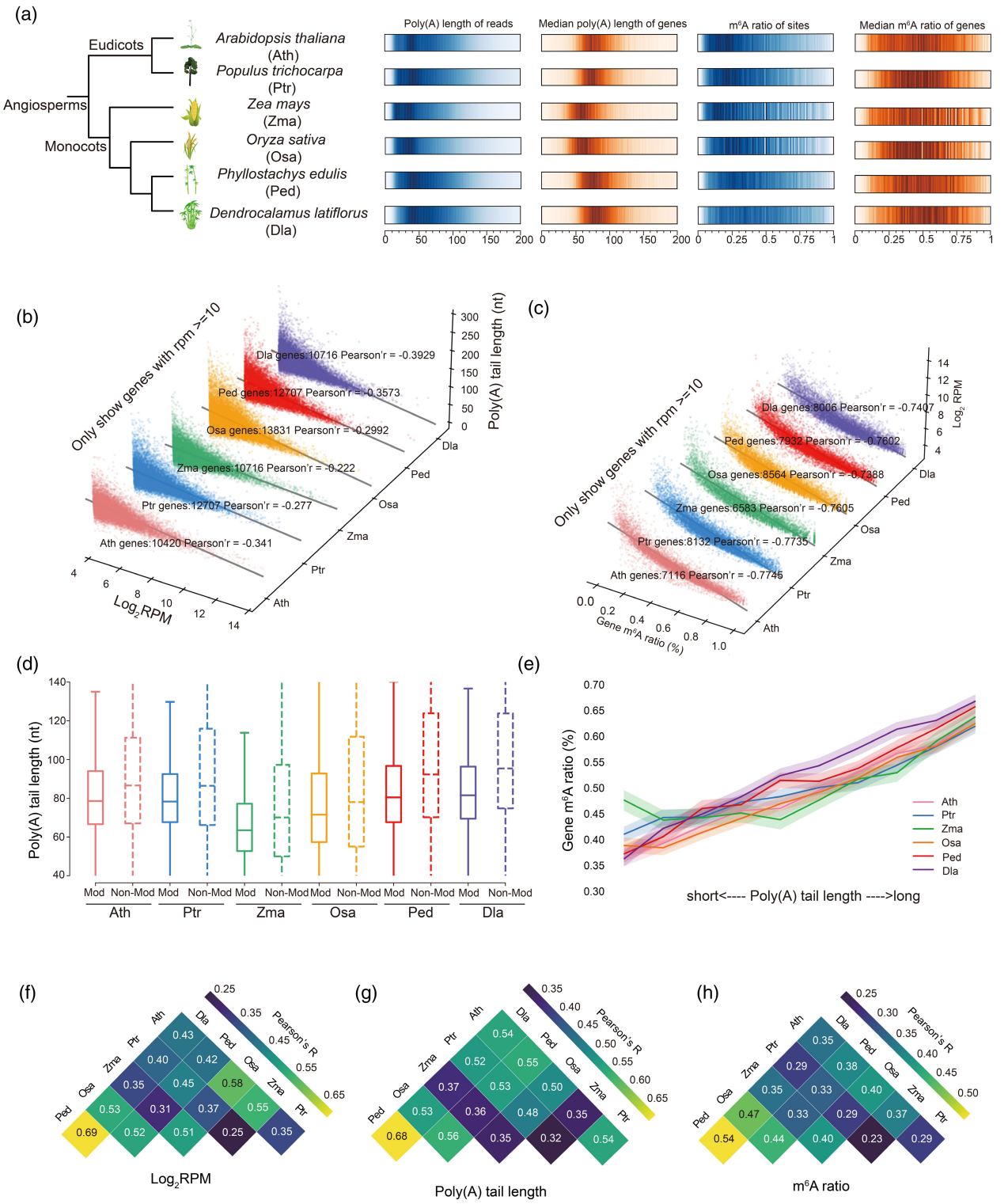


Figure 4. The profile of PAL and m^{6}A in six angiosperms.
(a) Distribution of global poly(A) tail lengths of long reads, median poly(A) tail length of genes, tm^{6}A ratio, and median m^{6}A ratio of genes in different species.
(b) Scatter plot showing the correlation between gene expression (RPM) and PAL.
(c) Scatter plot illustrating the correlation between gene expression (RPM) and m^{6}A ratio.
(d) Box plots showing PAL for m^{6}A-modified and non-m^{6}A-modified genes in different species.
(e) Line chart showing m^{6}A ratio distribution in genes sorted by PAL in different species.
(f–h) Heatmap showing the Pearson correlation of gene expression level (f), PALs (g), and m^{6}A (h) for orthologous genes coefficient between different species.

between the copy number of m^{6}A writers and the m^{6}A ratio across six species. Species with a higher number of writers exhibit higher m^{6}A ratio (Figure S25c). Several research reports that MTC components of the MT-A70 family are not m^{6}A writers (Greer et al., 2015; Luo et al., 2022). Thus, we also excluded MTC components of the MT-A70 family from m^{6}A writers. After recalculating and replotting, the results remained consistent (Figure S25d). Our findings indicate a clear dose–response relationship between the copy number of writers and m^{6}A modifications.

In most species, the majority of m^{6}A modifications occurred within the CDS, followed by the 3′UTR and 5′UTR, except for of *O. sativa*, where modifications were more prevalent in the 3′UTR (Figure S26a). The 3′UTR exhibited the highest m^{6}A ratio, followed by the 5′UTR and CDS (Figure S26b). The normalized distribution of m^{6}A sites on transcripts revealed enrichment near the gene stop codon and 3′UTR region for each species analyzed (Figure S27), consistent with previous study observed in both animal and plant species (Fu et al., 2014; Gao et al., 2022; Li et al., 2023; Liufu et al., 2023; Miao et al., 2022; Parker et al., 2020).

Previous studies have shown a close relationship between m^{6}A ratio and PAL with gene expression (Lim et al., 2016; Lima et al., 2017; Liu et al., 2021). In this study, we observed a gradual shortening of PAL (Figure 4b) and decrease in m^{6}A ratio (Figure 4c) with increasing gene expression, which was consistent with the previously reported negative correlation between gene expression with modification (Gao et al., 2022; Li et al., 2023; Liufu et al., 2023; Shen et al., 2016; Zhou et al., 2019) and PAL (Gao et al., 2022; Li et al., 2023; Lima et al., 2017; Liufu et al., 2023). We recalculated the relationship between m^{6}A ratio and PAL with gene expression using TPM derived from RNA-seq data, which was align well with the above results (Figure S28). Since m^{6}A is enriched around 3′UTR, near the poly(A) tail addition site, we also investigated whether there is correlation between m^{6}A ratio and PAL. We found that the PAL of the m^{6}A-modified gene was shorter than that of the non-m^{6}A-modified gene (Figure 4d). However, in m^{6}A-modified genes, m^{6}A ratio gradually increased as PAL increased, suggesting a positive correlation between m^{6}A and PAL in genes with m^{6}A (Figure 4e). These trends were consistent across the species tested (Figure 4b–e), suggesting that the correlation of gene expressive, m^{6}A ratio, and PAL is conserved in angiosperms.

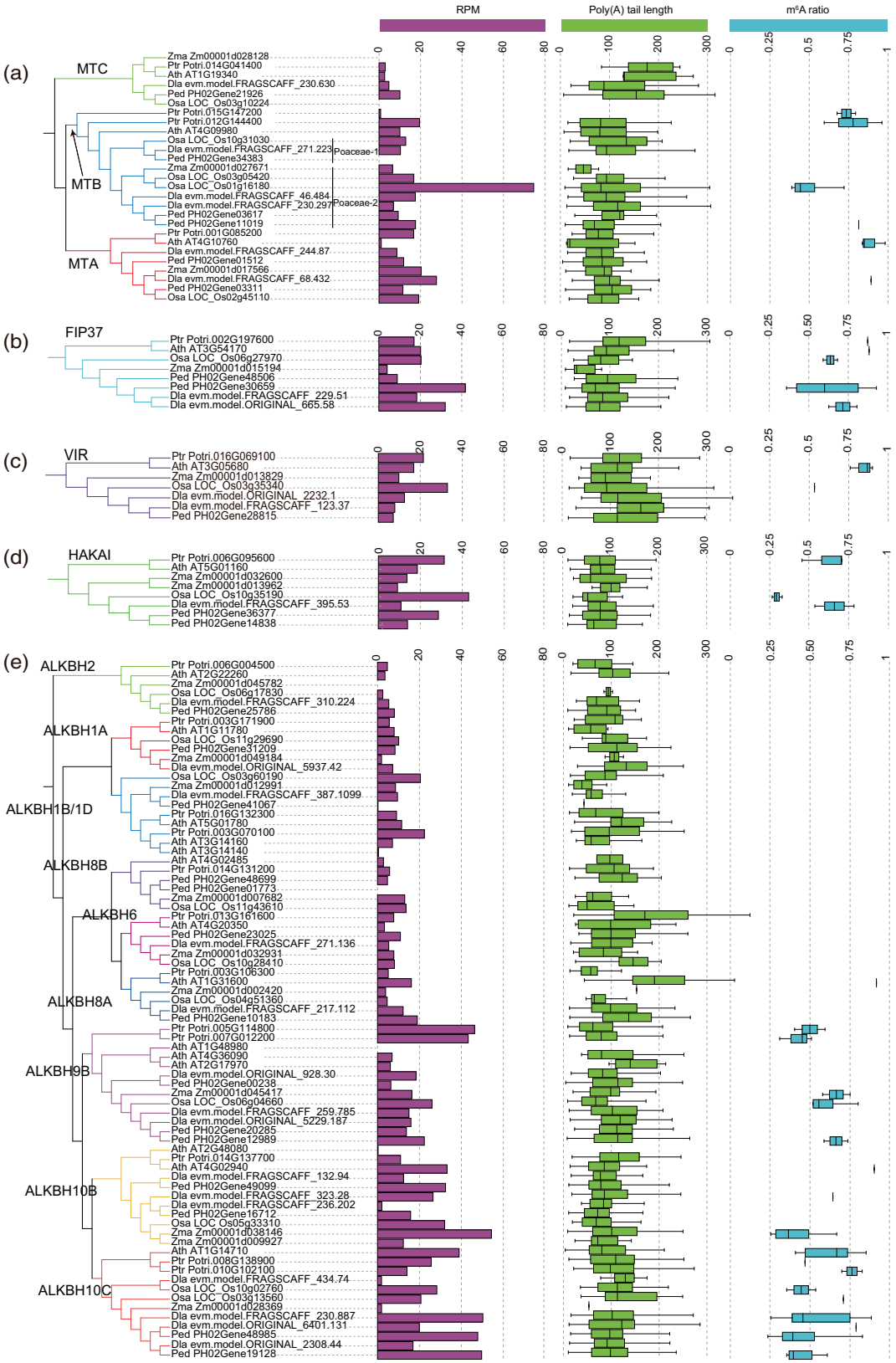
Orthologous often exhibit similar gene functions, and we speculated that they might have similar transcriptional and post-transcriptional regulation patterns. Indeed, we found positively correlated between any two species tested in expression (Figure 4f; Figure S29a), PAL (Figure 4g; Figure S29b) and m^{6}A ratio of orthologous genes (Figure 4H; Figure S29c). We noticed that the correlation of expression level, PAL and m^{6}A ratio of orthologous genes between *P. edulis* and *D. latiflorus* were also the highest (Figure 4f–h; Figure S29), possibly due to their membership in the Bambusoideae subfamily of the Poaceae family.

Poly(A) tail length, and m^{6}A ratio of m^{6}A pathway components: Insights from phylogenetic classification in six species

Based on the phylogenetic classification of m^{6}A pathway components in six species, we showed the expression level, PAL and m^{6}A ratio of m^{6}A writers, erasers, and readers (Figures 5 and 6). The gene expression of m^{6}A pathway components using TPM was consistent with the RPM trends (Figure S30). Almost all m^{6}A pathway components were expressed and PAL could be detected. Among the 50, 75, and 115 writers, erasers, and readers in six species, 37, 56, and 422 modification sites could be detected in 16, 18, and 42 genes, respectively. The presence of m^{6}A modification sites on m^{6}A pathway components indicates that writers, erasers, and readers may also be regulated by m^{6}A. In this study, we observed similar transcriptional and post-transcriptional patterns tend to exist within the same phylogenetic branch of the m^{6}A pathway components genes (Figures 5 and 6). For example, in the MTA-70 family, the expression level of genes in the MTC branch were lower than that of genes in the MTA and MTB branches, whereas the PAL was relatively higher (Figure 5a). The PAL of VIR genes in all species was also higher than that of other writers (Figure 5c). In the ALKBH family, we also observed that the expression levels of most genes in ALKBH10A/10B and ALKBH10C branches were higher than those of genes in other branches (Figure 5e).

In our synteny network analysis, ECT1/2-Eudicots1-2 were clustered into two distinct cluster (Figure 3f, cluster3 and 4), ECT1/2-Poeaceae4 was loosely connected to ECT1/2-Poeaceae1-3 (Figure S24b). Notably, in the ECT1/2 branch of the YTH family, the expression levels of genes in sub-clades ECT1/2-Poeaceae4 and ECT1/2-Eudicots2 (except AT1G55500) were significantly higher than those of genes in other ECT1/2 sub-branches. Along with the higher gene expression, a lower m^{6}A ratio is observed in these

Figure 5. Transcriptional and post-transcriptional regulation analysis of writers and erasers in six specious.
Presented from left to right are the phylogenetic tree, expression level histogram, PALs box plot, and m^{6}A ratio box plot of MTA-70 (a), FIP37 (b), VIR (c), HAKAI (d) and ALKBH (e) family genes of six species, including *Arabidopsis thaliana*, *Populus trichocarpa*, *Zea mays*, *Oryza sativa*, *Phyllostachys edulis*, and *Dendrocalamus latiflorus*.



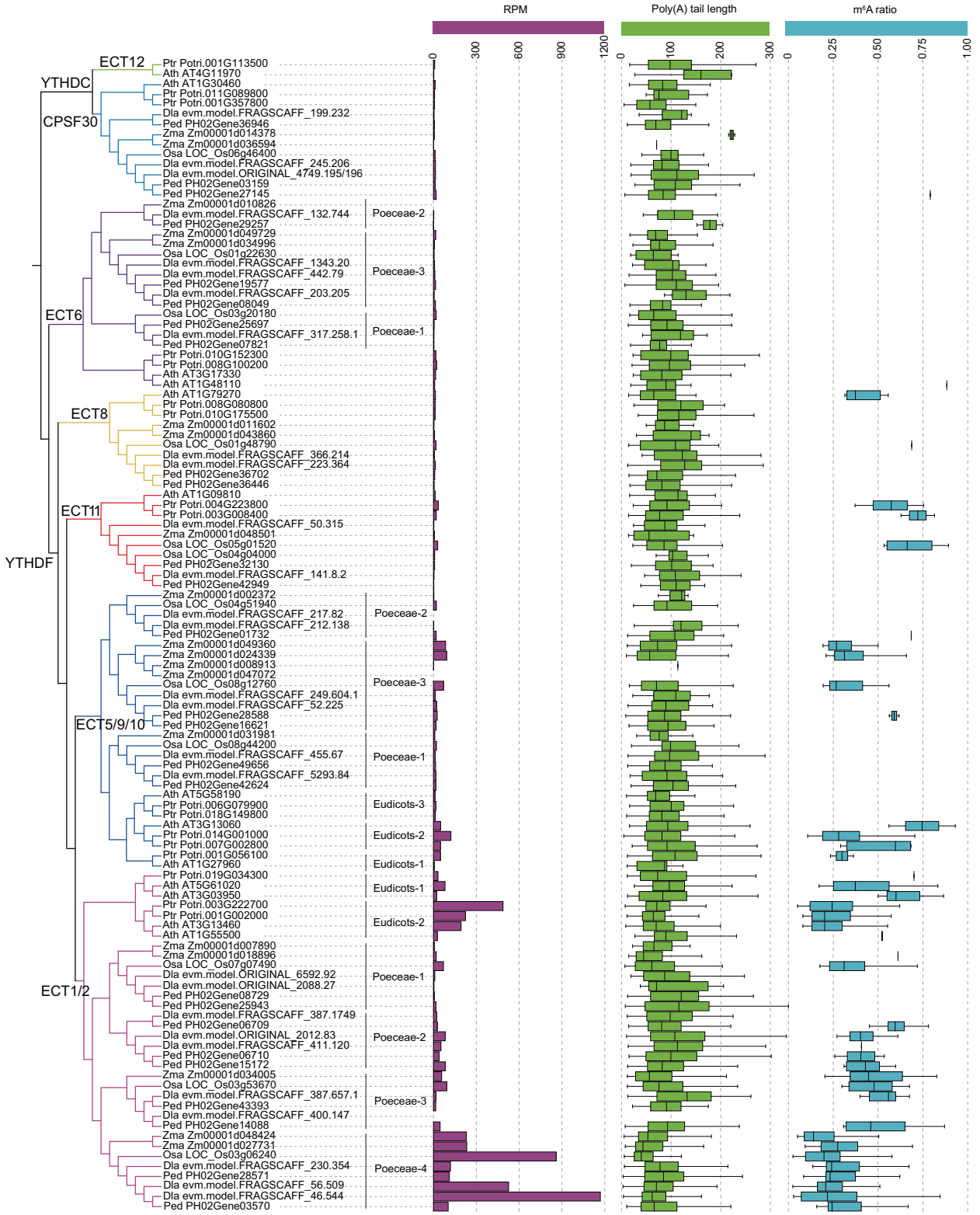


Figure 6. Transcriptional and post-transcriptional regulation analysis of readers in six specious.
Presented from left to right are the phylogenetic tree, expression level histogram, PALs box plot, and m^{6}A ratio box plot of YTH family genes of six species, including *Arabidopsis thaliana*, *Populus trichocarpa*, *Zea mays*, *Oryza sativa*, *Phyllostachys edulis* and *Dendrocalamus latiflorus*.

two branches (Figure 6). Meanwhile, the PAL of genes in these two branches were slightly lower than that of genes in other ECT1/2 sub-branches (Figure 6).

We further explored whether the m^{6}A modification sites of the gene in the same phylogenetic clade are conserved by performing multiple sequence alignment according to phylogenetic classification. Among the eight genes in all the tested monocots in this study, belonging to ECT1/2- Poaceae4, we found four and one completely conserved sites in the CDS and 3′UTR, respectively (Figure S31, red box). In the 3′UTR region, we also found two potentially conserved m^{6}A sites with slight differences in RRACH motif (Figure S31, blue box).

To investigation the role of conserved m^{6}A sites in gene expression regulation, we introduced mutations into three conserved m^{6}A sites located in the 3′UTR of *PheECT1/2-6* (PH02Gene28571) from the ECT1/2-Poeaceae4 sub-branch, substituting G for the m^{6}A sites. Additionally, we fused the 300 bp region of the 3′UTR sequence containing the three mutated sites with firefly luciferase (LUC) to generate the *proPheECT1/2*:LUC:3′UTR-mut construct (Figure S32a), whereas the native sequence served as a control (*proPheECT1/2*:LUC:3′UTR-WT). These constructs were then transformed into *Agrobacterium tumefaciens* for transient Agrobacterium-mediated infiltration of *Nicotiana benthamiana*. The LUC activity of *proPheECT1/2*:LUC:3′-UTR-mut was stronger than that of *proPheECT1/2*:LUC:3′-UTR-WT (Figure S32b). Using Renilla luciferase (REN) activity levels as a reference, we observed a 178% increase in LUC activity levels for the *proPheECT1/2*:LUC:3′UTR-mut compared with the control (Figure S32c). These results indicate that conserved m^{6}A sites in the 3′UTR region negatively regulate LUC gene expression, consistent with the negative correlation between gene expression and m^{6}A ratio reported by previous studies (Fu et al., 2014; Gao et al., 2022; Li et al., 2023; Liufu et al., 2023; Miao et al., 2022; Parker et al., 2020) and including our own findings (Figure 4c).

We also investigated the presence of alternative polyadenylation (APA) in m^{6}A pathway components. Among the 50 writers, 75 erasers and 115 readers across six species, 7 writers, 11 erasers, and 26 readers exhibited two or more poly(A) sites (Figure S33), indicating potential regulation by APA. However, it is not ruled out that as the sequencing depth increases, more APA events exist in the m^{6}A pathway components.

Core transcription factor regulation of m^{6}A pathway components across five plant species

Transcription factors (TFs) can either promote or inhibit gene transcription by binding to promoter region, thereby affecting the expression level (Jin et al., 2017). To further understand the type and number of TFs that may regulate transcription of m^{6}A pathway components, we used the PlantTFDB to predict TFs in the promoter of *A. thaliana*, *P. trichocarpa*, *Z. mays*, *O. sativa*, and *P. edulis*. Among the TFs, the common families in five specious were BCR-BPC, bZIP, C2H2, Dof, ERF, LBD, MADS, and MYB (Figure 7; Table S9). TFs that are orthologous in most species and capable of regulating multiple m^{6}A pathway components were considered core transcription factors (Figure 7; Table S9). The quadruple mutant of *BPC1*/*BPC2*/*BPC4*/*BPC6* are dwarfed and displays small curled leaves (Monfared et al., 2011). In this study, *AtBPC1*(AT2G01930) and its homologous genes acted on a number of m^{6}A pathway components, including members of the ECT1/2 clade (Figure 7a–e, green connecting lines; Table S9). Members of the ECT1/2 clade of Arabidopsis, such as *AtECT2*/*AtECT3*/*AtECT4*, can affect leaf formation time and leaf morphogenesis (Arribas-Hernández et al., 2018; Scutenaire et al., 2018; Wei et al., 2018), aligning with the biological function of *AtBPC1*. We further investigated whether *BPC1* directly interacts with gene of ECT1/2 clade using yeast one-hybrid (Y1H). The blue color of X-Gal was observed when the vectors carrying pro*PheECT1/2–5* and *PheBPC1* were co-transformed into EGY48 competent cells (Figure S34). Y1H showed that *PheBPC1* could interact with the *PheECT1/2-5* promoter in *P. edulis*.

It has been reported that *AtAS1* and *AtAS2* may play an important role in the development of plant virus infection symptoms (Machida et al., 2022). *AtAS2*(AT1G65620) and its homologous genes also acted on m^{6}A pathway components, including the *ALKBH9B* homologous gene in four other species except *P. trichocarpa* (Figure 7a,c–e, red connecting lines; Table S9), consistent with *AtALKBH9B*'s reported influence on the infection ability of Mosaic virus (Martínez-Pérez et al., 2017). *AS2* has also been reported to be a key regulator of flat symmetrical leaf development (Machida et al., 2022). In our study, *AS2* also acts on genes of the ECT1/2 clade in *O. sativa* and *P. edulis* (Figure 7d,e, blue connecting lines; Table S9). This study provides potential TFs regulatory information for further yeast one-hybrid assays to detect the binding of these TFs to m^{6}A pathway components.

DISCUSSION

In this study, we identified 4062 components of the m^{6}A pathway across 154 plant species, utilizing large-scale phylogenetic approach to explore their evolution, and proposed hypothetical evolutionary models of MTA-70, ALKBH and YTH in green plants (Figure 8a). Subsequently, we further explored conserved genomic contexts and lineage-specific transpositions of m^{6}A pathway components in 114 angiosperms using synteny network analysis (Figure 8b). Furthermore, we used DRS technology to reveal the PAL and m^{6}A ratio profiles in six angiosperms species, with a particular focus on the m^{6}A pathway components (Figure 8c).

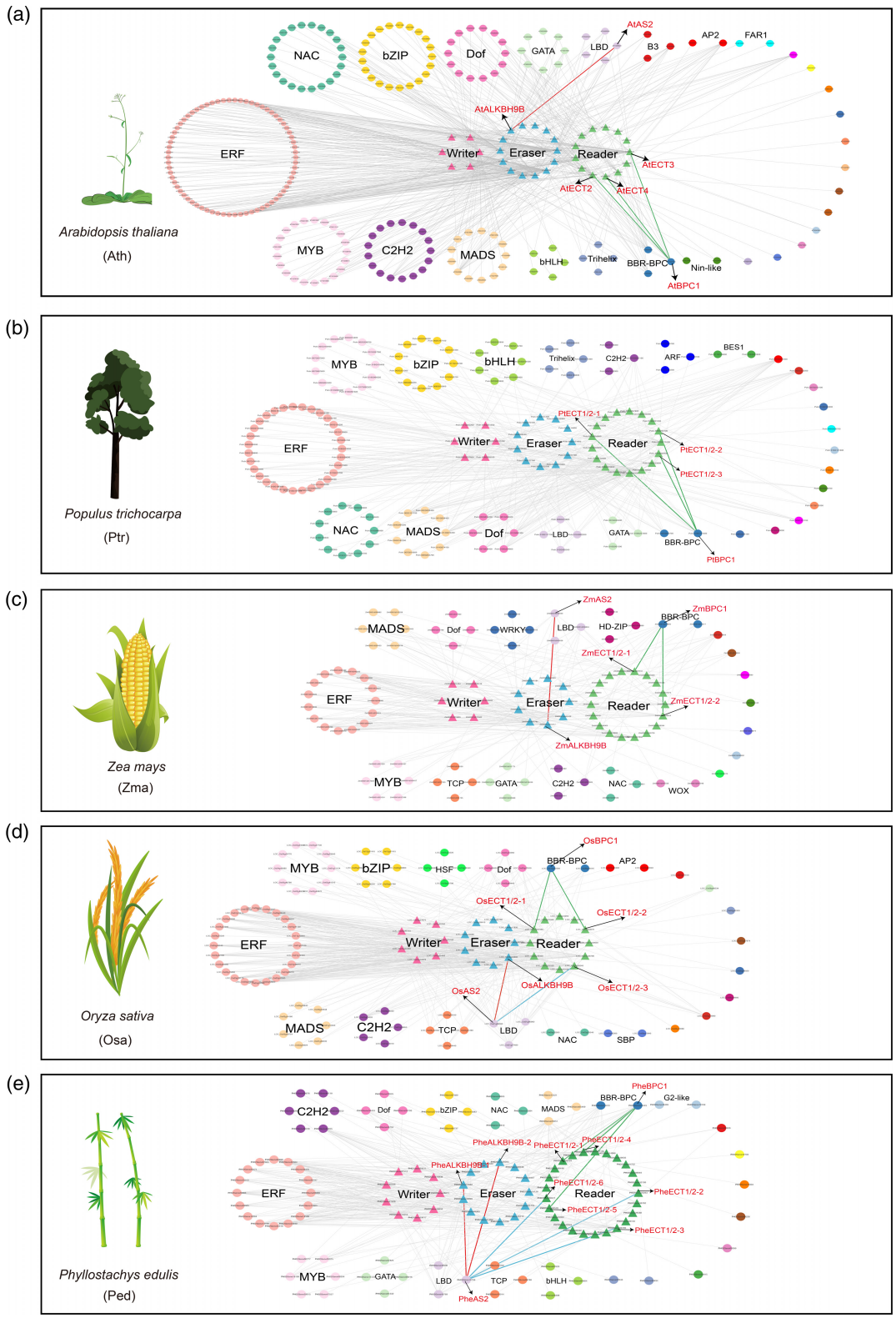


Figure 7. Transcription factor networks of m^{6}A regulatory proteins in five species. Core transcription factor regulation of m^{6}A pathway components in *Arabidopsis thaliana* (a), *Populus trichocarpa* (b), *Zea mays* (c), *Oryza sativa* (d), and *Phyllostachys edulis* (e) is illustrated. Different colored circles represent different TF families, whereas different colored triangles represent writers, eraser, and readers. Different colored lines represent the regulatory relationships between TF families and m^{6}A pathway components.

Due to the quality of genome assembly and the accuracy of genome annotation, genome-wide gene family identification often results in the omission of some genes. Thus, we selected a sufficient number of species for the identification of MTA70, FIP37, VIR, HAKAI, ALKBH, and YTH gene families. Our study revealed widespread distribution of m^6A pathway components in green plants. The copy numbers of MTA70, FIP37, VIR, and HAKAI remained stable in land plants, whereas the copy numbers of ALKBH and YTH varied among different species (Figure 1). MTA70, ALKBH, and YTH family members were detected in unicellular algae (Figure 1), suggesting that m^6A regulatory networks were already established in the common ancestor of green plants. FTO, one of the first mammalian demethylases to be discovered, was found only in vertebrates and marine algae (Robbens et al., 2008). Interestingly, writer components FIP37, VIR and HAKAI were not identified in mamiellophyceae (marine algae) (Figure 1). This suggests that marine algae may possess different m^6A regulatory mechanisms compared with land plants.

Previous studies have divided the MT-A70 family into three subfamilies-MTA, MTB, and MTC in higher eukaryotes (Bujnicki et al., 2002; Iyer et al., 2016; Liang et al., 2020; Yue et al., 2019), and categorized ALKBH family into seven main clades (Liang et al., 2020). Regarding YTH proteins, previous studies have primarily categorized them into YTHDF and YTHDC, with subdivisions such as YTHDFa, YTHDFb, YTHDFc, YTHDCa, and YTHDCb (Scutenaire et al., 2018; Yue et al., 2019). In this study, based on the comprehensive and systematic phylogenetic analysis of the m^6A pathway components in land plants and angiosperms, we inferred the possible evolutionary model of the MT-A70, ALKBH and YTH family in plants (Figure 8a). MT-A70 family gave rise to five clades prior to the emergence of hydrobiontic algae. Notably, Chlorophytes-1 and Chlorophytes-2 were lost in land plants and charophytes. Within angiosperms, three main clades eventually evolved, with the MTB branch expanding in Poaceae family to form MTB-Poaceae1 and MTB-Poaceae2 sub-branches (Figure 8a, Upper left panel). The plant ALKBH family gave rise to eight clades before the emergence of hydrobiontic algae, and Algae-specific clade were lost in land plants and charophytes. Subsequently, ALKBH9B and ALKBH10B/10C differentiated before gymnosperms emerged and after charophytes appeared. ALKBHs eventually expanded into nine major clades in angiosperms, with ALKBH1B/1D, ALKBH9B, ALKBH10B, and ALKBH10C clades exhibiting an increased copy number in some angiosperms (Figure 8a, Upper right panel). YTH family gave rise to two clades (YTHDF and YTHDC) before the emergence of hydrobiontic algae. Following the emergence of chlorophytes and preceding the appearance of charophytes, YTHDC were divided into CPSF30 and ECT12. After charophytes differentiated, YTHDF eventually expanded into ECT1/2, ECT5/9/10, ECT6, ECT8, ECT11 in angiosperms. The angiosperm YTHs eventually expanded into seven major clades. Notably, YTH family was further expanded or lose in angiosperms. In ECT1/2 and ECT5/9/10 branches, significant expansion occurred in eudicots and monocots, eventually forming two and three eudicots sub-branches, and four and three poaceae sub-branches, respectively. In ECT6, monocots also expanded, forming three poaceae sub-branches. Additionally, various lineage-specific gene duplication and loss events occurred at different taxonomic levels. Notably, ECT12 was lost in monocots (Figure 8a, Lower panel).

Generally, genes within the same phylogenetic branch often exhibit similar function, whereas genes in different phylogenetic branches tend to have specific function. The m^6A demethylase is a member of the ALKBH family, but not all ALKBH members appear to function as m^6A demethylase. In *A. thaliana*, identified m^6A demethylases include *AtALKBH9B* (Martínez-Pérez et al., 2017), *AtALKBH10B* (Duan et al., 2017). In tomato, the demethylase *SLALKBH2* (Solyc01g104130) (Zhou et al., 2019) belongs to the clade ALKBH9A/9B/9C (Table S4). Other members of the ALKBH family, such as *AtALKBH2*, functions as a repair enzyme to demethylate 1-methyladenine (1-mA) and 3-methylcytidine (3-mC) on DNA or RNA (Duncan et al., 2002; Fu & Samson, 2012; Meza et al., 2012). *AtALKBH8A* is responsible for the final step of mcm5U formation in tRNAs (Leihne et al., 2011). Additionally, it has also been demonstrated that reducing the transcription levels of *AtALKBH6* does not alter m^6A methylation levels (Huong et al., 2020). Phylogenetic analysis showed that human ALKBH5 was clustered together with ALKBH9A/9B/9C/10A/10B/10C clade (Figure S5). We hypothesize that genes belonging to the ALKBH9A/9B/9C/10A/10B/10C branches are more likely to be involved in m^6A demethylation. Moreover, ALKBH10B and ALKBH10C formed an angiosperm conserved cluster (Figure 3E, cluster6 and 10), suggesting highly conserved functions as demethylases in angiosperms.

We emphasized a novel insight in m^6A reader (Figure 8d). In the ECT1/2 clade, all gene members of the Poaceae family further formed four sub-clades: ECT1/2-Poeaceae1-4. Synteny network results shows ECT1/2-Poaceae4 was loosely connected to the other three sub-branches (Figure S24b). We observed high expression levels of genes in the ECT1/2-Poaceae4 sub-branch, significantly different from ECT1/2-Poeaceae1-3 (Figure 6). The expression patterns of *P. edulis* in different tissues also showed differences between ECT1/2-Poaceae4 and ECT1/2-Poeaceae1-3: ECT1/2-Poaceae4 genes were highly expressed in the flowering tissues, whereas ECT1/2-Poeaceae1-3 genes were highly expressed in bamboo shoots (Figure S35). Thus, we speculate that differences in synteny between ECT1/2-Poaceae4 and ECT1/2-Poaceae1-3

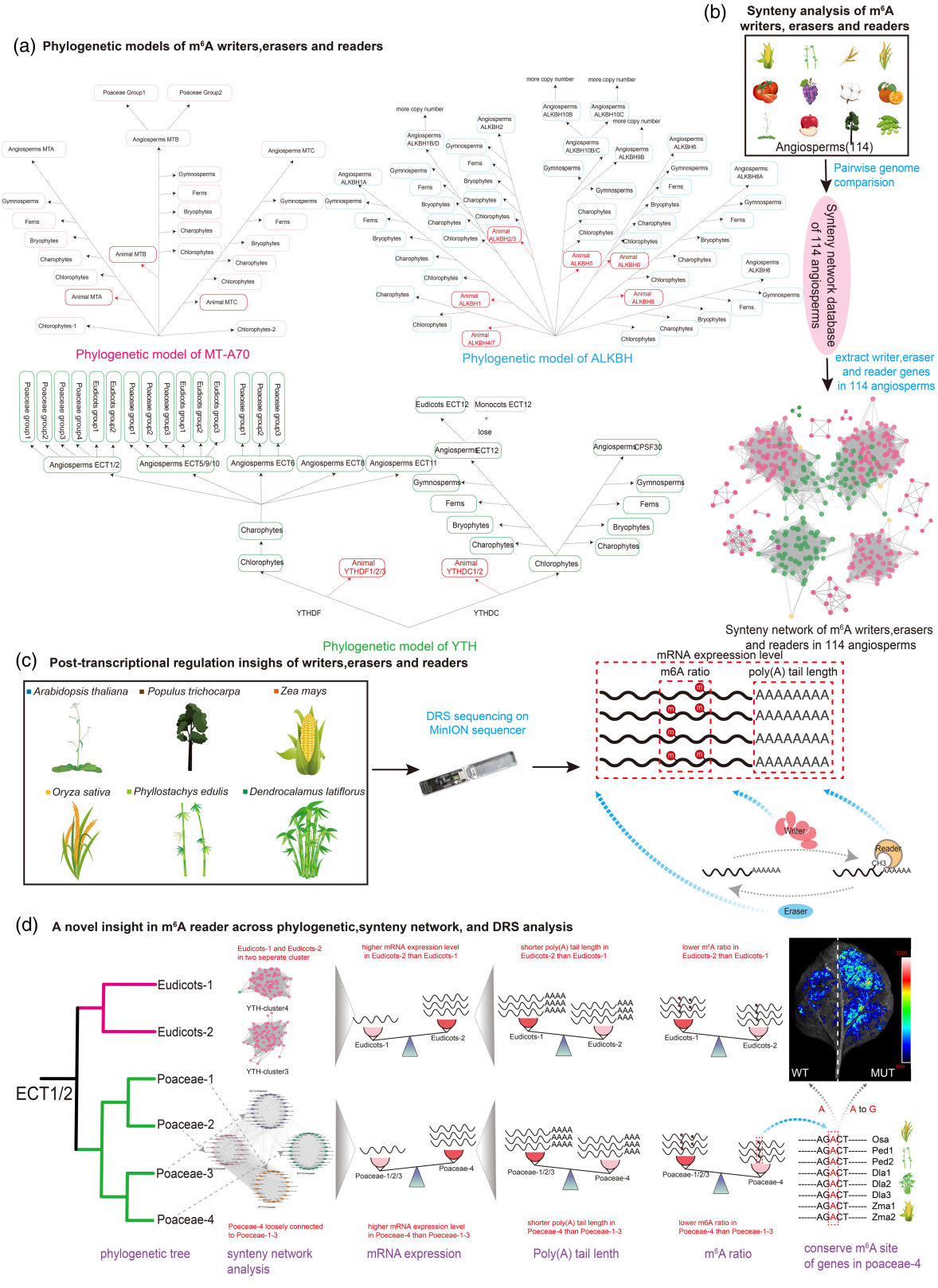


Figure 8. A schematic overview and summary of novel insights from this study.
(a) Phylogenetic revelations. Within the ECT1/2 clade (formerly named YTHDFa), further subdivision reveals ECT1/2 branching into two divisions (Eudicots1-2) in eudicots and four divisions (Poaceae1-4) in monocots.
(b) Synteny network analysis. Synteny network analysis indicates that ECT1/2-Eudicots1 and ECT1/2-Eudicots2 cluster separately, whereas ECT1/2-Poaceae4 exhibits a looser connection to ECT1/2-Poaceae1-3. Disparities in phylogenetic and synteny patterns may contribute to the heightened gene expression observed in ECT1/2-Eudicots2 and ECT1/2-Poaceae4.
(c) Post-transcriptional regulation insights of m^{6}A writers, erasers, and readers. We note significantly lower m^{6}A ratios and slightly shorter PALs of genes in the ECT1/2-Eudicots2 and ECT1/2-Poaceae4 branches compared with other branches, potentially offering an additional explanation for the elevated gene expression in these branches. Notably, conserved m^{6}A sites were observed in the ECT1/2-Poaceae4 branch.
(d) Novel perspective on m^{6}A readers across phylogenetic, synteny network, and DRS analyses.

may lead to differences in expression patterns. In this study, we found a negatively correlation between PAL and m^{6}A ratio with gene expression, whereas PAL and m^{6}A ratio were positively correlated (Figure 4b,c,e). Additionally, we found these trends were consistent across the species, suggesting that the correlation of gene expression, m^{6}A ratio and PAL is conserved in angiosperms. The m^{6}A ratio and PAL of genes in ECT1/2-Poaceae4 subclade were lower than genes in ECT1/2-Poaceae1-3 sub-clades (Figure 6), suggesting that the "lower m^{6}A ratio" and "shorter PAL" of genes in ECT1/2-Poaceae4 sub-clade may be another reason for their higher expression.

We also observed a similar pattern in the ECT1/2 Eudicots branches. ECT1/2 branches divide into sub-branches within Eudicots: ECT1/2-Eudicots1 and ECT1/2-Eudicots2 (Figure S13). Synteny network analysis further segregated ECT1/2-Eudicots1 and 2 into two separate clusters. As suspected, these two branches also exhibit distinct pattern at the transcriptional and post-transcriptional levels. Specifically, the ECT1/2-Eudicots2 (except AT1G55500) sub-branches exhibited significantly higher expression level compared with genes in the other ECT1/2-Eudicots1 sub-branches, accompanied with lower m^{6}A modification and PAL (Figure 6).

RNA modifications were detected in most of the m^{6}A pathway components in the two eudicots and four monocots tested, suggesting that these m^{6}A pathway components have potential self-regulatory circuits. Specifically, conserved modification sites distributed in the CDS and 3′UTR were identified in the ECT1/2-Poaceae4 sub-branch of the YTH family (Figure S31). We accessed single nucleotide polymorphism (SNP) data from the Rice Functional Genome Breeding database (RFGB) for the rice LOC_Os03g06240 gene, which contains conserved m^{6}A modification sites. Out of the 44 SNP sites identified, we observed none overlapped with the 28 m^{6}A modification sites. Therefore, this finding suggests a low frequency of SNPs at the gene including conserved m^{6}A modification sites. In the future, a comprehensive SNP analysis of all modification sites can be performed to obtain more general conclusions. Dual luciferase experiments further demonstrate that conserved m^{6}A sites in the 3′UTR region negatively regulate LUC gene expression, consistent with the conclusion that gene expression is negatively correlated with m^{6}A modification (Figure S32). An intriguing research avenue would be to mutate these conserved modification sites at the 3′UTR using gene editing techniques to explore resulting phenotypes.

MATERIALS AND METHODS

Genome-wide identification of m^{6}A writers, erasers, and readers in 154 genomes

Genomic sequence and taxonomy relationship of 154 plants were retrieved from PHYTOZOME v.13, Ensembl Plants, NCBI, and Gigascience (http://gigadb.org), and other public databases. To identify members of m^{6}A writers, erasers, and readers gene families, the corresponding domain sequences were downloaded from the Pfam website (http://pfam.xfam.org/) including MTA70 (PF05063), WTAP (PF17098), virilizer (PF15912), clavaminate synthase-like (PF13532) and YTH (PF04146). Meanwhile, writers, erasers, and readers identified in *A. thaliana*, *O. sativa*, *Ginkgo biloba*, *Selaginella moellendorffii*, *Physcomitrium patens*, and *Chlamydomonas reinhardtii* were used as query sequences for BLASTP (evalue $<1 \times 10^{-5}$). The obtained sequences were manually screened according to the corresponding family characteristics. For the MT-A70 family of writers, the MT-A70 domain was a requirement. The FIP37 component had to contain the WTAP domain, the virilizer had to contain the VIR_N domain, and HAKAI had to contain the HAKAI domain. Erasers and readers were identified based on the presence of the 2OG-FeII_Oxy superfamily domain and the YTH domain, respectively. In cases where genes contain multiple transcripts, the longest transcript was selected for subsequent analysis.

Multiple sequence alignment and phylogenetic analysis

Above identified protein sequences were aligned using mafft v7.455 (Katoh & Standley, 2013) with parameter "L-INS-I." To enhance the alignment quality, only the regions encoding functional domains of m^{6}A writers, erasers, and readers were retained to avoid the varying lengths of m^{6}A pathway components from different species. The corresponding coding nucleotide sequence were identified using PAL2NAL (Suyama et al., 2006) based on the aligned protein sequence. The nucleotide alignment was then used for subsequent phylogenetic analysis.

For exploring phylogenetic relationships among different species, the sequences of each family were divided into two datasets based on taxonomy: which were used to construct phylogenetic trees: dataset I, consisting of all non-angiosperms (including 13 chlorophytes, 7 charophytes, 7 bryophytes, 1 lycophyte, 3 ferns, and 9 gymnosperms, a total of 40 species) and 4 angiosperms (*Amborella trichopoda*, *Liriodendron tulipifera*, *O. sativa*, *A. thaliana* represent basal angiosperms, magnoliidea, monocots, and eudicots, respectively), and Data set II, consisting of all

angiosperms (114 species) (details in Table S4). IQ-TREE2 (Minh et al., 2020) was used to construct maximum likelihood phylogenetic tree using parameter "-m MFP -bb 3000 –bnni." The resulting phylogenetic trees were visualized using iTOL (Letunic & Bork, 2016) (https://itol.embl.de).

Synteny network analyses

The synteny network analysis of m^{6}A writers, erasers, and readers was carried out following the synteny network pipeline (Zhao et al., 2017). This involved synteny comparisons intra- and interspecies species. All possible syntenic gene pairs were extracted and combined into a synteny network dataset. Subsequently, gene families of interest from synteny networks were extracted and clustered. Initially, MCScanX with default parameter (Wang et al., 2012) was utilized to identify intra- and intergenomic synteny in 114 angiosperms genome, followed by extracting synteny blocks including m^{6}A writers, erasers, and readers for network clustering algorithm using infomap (https://pypi.org/project/infomap/). The clusters generated by infomap were visualized using CYTOSCAPE v.3.5.1 (Shannon et al., 2003). Genes derived from the synteny clusters were mapped back to the phylogenetic tree for visualization using iTOL (Letunic & Bork, 2016). Parallel coordinate synteny plot was generated using Genomicus (http://genomicus.biologie.ens.fr/genomicus-plants).

Plant materials and RNA extraction

The experimental materials were consisted of *O. sativa* cv. Nipponbare, *D. latiflorus* Munro, and *Z. mays* (B73) (including public data from three other species, details in Tables S7 and S8). Total RNA was extracted using an RNAprep pure Plant Kit (Tiangen, cat. no. DP441). Residual genomic DNA in the RNA sample was removed using DNase I. Quality assessment of the RNA was conducted through 1% agarose-gel electrophoresis, Nanodrop 2000 spectrophotometry (Thermo Scientific) for purity, and Agilent 2200 Bioanalyzer for integrity and concentration.

Transcriptome sequencing (RNA-seq) and bioinformatics analysis

Strand-specific RNA-seq libraries of *O. sativa*, *D. latiflorus* Munro, and *Z. mays* were constructed using the dUTP method and sequenced on an Illumina Novaseq 6000 instrument. Raw sequencing reads from six species were processed for quality-controlled using fastp with default parameter (Chen et al., 2018). Adapter sequences were removed, and then the quality-controlled FASTQ files were aligned to the corresponding genomes using HISAT2 using parameter "--dta --rna-strandness RF" (Kim et al., 2019). The resulting BAM files were filtered to retain reads aligned to a single region in the genome and were used for read counting with featureCounts using parameter "-p -t exon" (Liao et al., 2014). The expression level of each gene was calculated using Transcripts per Kilobase per Million mapped reads (TPM).

Nanopore DRS and bioinformatics analysis

The DRS library was constructed using the Nanopore DRS kit (SQK-RNA002, Oxford Nanopore Technologies). Briefly, the poly(A) + RNAs were ligated to the ONT reverse transcription adapter (RTA) using concentrated T4 DNA Ligase (M0202, NEB) and then reverse transcribed using SuperScript III Reverse Transcriptase (18080093, Thermo Fisher Scientific). The resulting products were purified using 1.8X VAHTS RNA Clean Beads (N412, Vazyme) with subsequent washing using 80% freshly prepared ethanol. An RNA Adapter (RMX) was ligated onto the RNA:DNA hybrid, and the mixture was purified again using 1.8X VAHTS RNA Clean Beads and washed with Wash Buffer (WSB) twice. The sample was then eluted in Elution Buffer (ELB) and mixed with RNA Running Buffer (RRB). The DRS was carried out using a Spot-ON flow cell FLO-MIN106D (Oxford Nanopore Technologies) on the MinION MK 1B device (Oxford Nanopore Technologies). After obtaining the first strand of cDNA using reverse transcription PCR, sequencing adapter was ligated to initiate sequencing. Raw current signals from six species were base-called using Guppy (v3.6.1) with parameter "--flowcell FLO-MIN106 --kit SQK-RNA002." Short reads obtained from Illumina RNA-seq data were tcombined to correct DRS reads using LoRDEC using parameter "-k 19 -s 3" (Salmela & Rivals, 2014). The corrected long read segment was mapped to the corresponding genome using minimap2 with parameter "-ax map-ont -uf" (Li, 2018). Read counts were obtained using featurecount using parameter "-L -R CORE -t transcript" (Liao et al., 2014), and the expression level of each gene was calculated using Reads per Million mapped reads (RPM).

The polya module of nanopolish with default parameter (v0.13.2) (Workman et al., 2019) was used to estimate the PAL of each read. Only the reads with a QC label "PASS" in the nanopolish report was used for downstream analysis. The PAL of each gene was defined as the median of all reads mapped to the gene and labeled "PASS" by nanopolish QC.

m^{6}A modification sites were identified at the single base level using Nanom^{6}A (Gao et al., 2021). Briefly, multi_to_single_fast5 was used to split each fast5 file into single-read levels. Raw signals were assigned to corresponding transcriptome reference sequences using Re-squiggle algorithm from Tombo. Modification sites were identified using Nanom^{6}A (Gao et al., 2021) with a cut-off of 10, based on the parameter "--support 10," allowing normalization for sequencing depth across species.

Poly(A) site with >5% usage rate was remained as candidate poly(A) site. The distance between two distinct poly(A) site should be more than 30 nt. The reads with termination site located 13 nt immediately upstream and downstream a poly(A) site was assigned to the poly(A) site.

Upstream promoter region transcription factor prediction

The 2 kb upstream of the start codon of each m^{6}A pathway components gene was designated as the proximal promoter region and submitted to the PlantTFDB database (Jin et al., 2017) for predicting TFs binding site with an *E*-value of 10^{-6}. The action network diagram of trans-acting factors and m^{6}A pathway components was constructed using CYTOSCAPE v.3.5.1 (Shannon et al., 2003). OrthoFinder (Emms & Kelly, 2019) is used to identify orthologous TFs in *A. thaliana*, *P. trichocarpa*, *Z. mays*, *O. sativa*, and *P. edulis*.

Dual-luciferase assay

The 2 kb upstream promoter region of PheECT1/2-6 and the first 300 bp of the 3′UTR region were amplified from *P. edulis* genomic DNA to serve as the control. All primer sequences, promoter and 3′UTR sequences, are listed in Table S10. In the experimental group, three conserved m^{6}A sites in the 3′-UTR region were replaced by A to G. For the dual-luciferase assay, PheECT1/2-6-3′UTR-WT or PheECT1/2-6-3′UTR-MUT were inserted into the pGreenII 0800-LUC vector to obtain recombinant plasmids. The Dual-luciferase assay were conducted using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI). LUC and REN activity levels were measured with a Cytation 3 Cell Imaging Multi-Mode Reader (BioTek, Santa Barbara, CA).

Yeast one-hybrid assay

The 2 kb promoter of *PheECT1/2-5* was cloned and ligated into the pLaczi-2μ vector. *PheBPC1* was cloned and expressed by fusion with the pB42AD vector. pLaczi-2μ and pB42AD, *proPheECT1/2-5*: pLaczi-2μ and pB42AD, pLaczi-2μ and *PheBPC1*:pB42AD, *proPheECT1/2-5*:pLaczi-2μ and *PheBPC1*:pB42AD was co-transformed into Y1H competent cells (EGY48). The mixes were incubated at 30°C for 30 min and resuspended every 10 min, followed by incubation at 42°C for 15 min with resuspension every 7.5 min. The Y1H yeast cells collected by centrifugation were evenly cultured on an SD/-Trp/-Ura solid medium and incubated at 30°C for 2–4 days. After growing single colonies, the cells were spread and coated on an SD/-Trp/-Ura + Gal + Raf + X-Gal color plate for observation.

AUTHOR CONTRIBUTIONS

JZ and LG conceived and designed the research. JZ, YW, MZ, HW, XL, and HZ performed the bioinformatics analyses. LW, LM, and LZ performed experiments. CL revised the manuscript. JZ and LG wrote the article.

ACKNOWLEDGEMENTS

This research was funded by the National Key R&D Program of China (2023YFD2200201 and 2021YFD2200505), the National Natural Science Foundation of China (3237141305), the Natural Science Foundation of Fujian Province (2021J02027), and the S&T Innovation (KFB23180) and the Forestry Peak Discipline Construction Project (72202200205) of Fujian Agriculture and Forestry University.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests associated with this work.

DATA AVAILABILITY STATEMENT

The Nanopore DRS and Illumina RNA-seq raw data in this study are available in NGDC (https://ngdc.cncb.ac.cn/) under accession CRA017191 of PRJCA020476.

SUPPORTING INFORMATION

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

Figure S1. Phylogeny of green plants showing the species included in this study.

Figure S2. Phylogenetic tree of plant MT-A70s constructed using MT-A70s from MT-A70-dataset I and 6 animals.

Figure S3. Number of members in different clades of the MT-A70 family from 114 angiosperms.

Figure S4. Phylogenetic tree of the MTB clade of the MT-A70 family in monocots.

Figure S5. Phylogenetic tree of plant ALKBHs constructed with ALKBHs from ALKBH-dataset I and 6 animals.

Figure S6. Phylogenetic tree of the ALKBH family in Brassicales. (a) ALKBH1B/1C in Brassicales. (b) ALKBH9A/9B/9C in Brassicales. (c) ALKBH10A/10B in Brassicales.

Figure S7. Number of members of different clades of the ALKBH family from 114 angiosperms.

Figure S8. Phylogenetic tree of ALKBH10B clade of the ALKBH family in eudicots.

Figure S9. Phylogenetic tree of ALKBH10C clade of the ALKBH family in eudicots.

Figure S10. Phylogenetic tree of plant YTHs constructed with YTHs from YTH-dataset I and 6 animals.

Figure S11. Phylogenetic tree of the YTH family in Brassicales. (a) ECT1/3 in Brassicales. (b) ECT2/4 in Brassicales. (c) ECT6/7 in Brassicales.

Figure S12. The heatmap represents the number of members of different clades of the YTH family from 114 angiosperms.

Figure S13. Phylogenetic tree of the ECT1/2 clade of the YTH family.

Figure S14. Phylogenetic tree of the ECT5/9/10 clade of the YTH family.

Figure S15. Phylogenetic tree of the ECT6 clade of the YTH family.

Figure S16. Synteny network clusters of MT-A70, FIP37, VIR, and HAKAI. Phylogenetic profiling of 20, 12, 9, and 4 clusters of MT-A70, FIP37, VIR, and HAKAI, respectively.

Figure S17. The network of MT-A70, FIP37, VIR, and HAKAI clusters. Pink circles for Eudicot genes, green circles for monocot genes, and yellow circles for basal angiosperm genes. Nodes represent genes and edges represent syntenic relationship between them.

Figure S18. Synteny network clusters of ALKBH and phylogenetic profiling of 48 clusters of ALKBH.

Figure S19. The network of ALKBH clusters. Pink circles for Eudicot genes, green circles for monocot genes, and yellow circles for basal angiosperm genes. Nodes represent genes and edges represent a syntenic relationship between them.

Figure S20. Synteny network clusters of YTH and phylogenetic profiling of 34 YTH clusters.

Figure S21. Network of YTH clusters. Pink circles represent Eudicot genes, green circles represent monocot genes, and yellow circles represent basal angiosperm genes. Nodes depict genes, and edges indicate syntenic relationship between them.

Figure S22. Parallel coordinate synteny plots of FIP37 derived from Genomicus. (a) Synteny relationship of the grape homolog of FIP37 (VIT_17s0000g08940) across multiple lineages. (b) Synteny relationship of the Arabidopsis FIP37 gene (AT3G54170) across multiple lineages.

Figure S23. Phylogenetic and tandem relationships of MT-A70s (a), FIP37s (b), ALKBHs (c), and YTHs (d) in angiosperms. The two ends of the red line represent pairs of genes for which tandem duplication occurs.

Figure S24. Synteny network of MT-A70 cluster3 (a) and YTH cluster7 (b).

Figure S25. Violin plot of PAL (a) and box plot of m^6A ratio (b) across different species. Scatter plot shows the correlation between writer copy numbers including MTC (C) and excluding MTC (D), and the m^6A ratio.

Figure S26. Percentage (a) and ratio (b) of m^6A modification sites within CDSs and UTRs across different species.

Figure S27. Metagene plot of methylation sites along the gene body across different species.

Figure S28. (a) Scatter plot showing the correlation between gene expression (TPM) and PAL. (b) Scatter plot illustrating the correlation between gene expression (TPM) and m^6A ratio.

Figure S29. Correlation of the expression level (a), median PAL (b), and median m^6A ratio (c) of orthologous gene pairs across different species.

Figure S30. Scatter plot (left panel) and bar plot (right panel) showing the correlation between RPM and TPM of m^{6}A pathway components in *Arabidopsis thaliana* (a), *Populus trichocarpa* (b), *Zea mays* (c), *Oryza sativa* (d), *Phyllostachys edulis* (e), and *Dendrocalamus latiflorus* (f).

Figure S31. Conserved modification site of ECT1/2-Poaceae4 branch gene. "M" represents modified A base. The red box represents fully conserved modification site, and the blue box represent potentially conserved modification site with slight differences in RRACH motif.

Figure S32. The dual-luciferase activity assay demonstrated that conserved m^{6}A sites in the 3′UTR region negatively regulated the expression of LUC. (a) Schematic diagram of the dual-luciferase activity assay vector; (b) Live imaging of *Nicotiana benthamiana* leaves; (c) The dual-luciferase activity assay. Error bars represent the standard error of mean (SEM). *P*-values were determined using a two-tailed Student's *t*-test (***$P < 0.001$).

Figure S33. Phylogenetic tree and poly(A) site numbers histogram of writers (a) erasers (b) and readers (c) in six angiosperms.

Figure S34. Yeast one-hybrid (Y1H) analysis demonstrating PheBPC1 binding to the PheECT1/2–5 promoter.

Figure S35. Gene expression patterns of ECT1/2 in different subbranches of *Phyllostachys edulis* in different tissue. The raw RNA-seq data was downloaded from the SRA database (PRJEB2956).

Table S1. Sources of sequence and genome version information used in this study.

Table S2. Gene ID, CDSs, and protein sequences of the identified m^{6}A pathway components.

Table S3. The number of m^{6}A pathway components in green plants.

Table S4. Classification of MT-A70s, ALKBHs, and YTHs used in this study.

Table S5. The list of node and edge of all communities from synteny networks of MTA70s FIP37s VIRs HAKAIs ALKBHs, and YTHs.

Table S6. The list of tandem duplication genes in MTA70s, FIP37s, VIRs, HAKAIs, ALKBHs, and YTHs.

Table S7. Summary of sequenced and mapped reads in DRS samples generated in this study.

Table S8. Summary of previously published DRS datasets used in this study.

Table S9. The family, orthogroups ID, and bound m^{6}A pathway components of predicted TFs in this study.

Table S10. Primer, promoter and 3′UTR sequences used in dual-luciferase assay and yeast one-hybrid assay.

REFERENCES

Arribas-Hernández, L., **Bressendorff, S.**, **Hansen, M.H.**, **Poulsen, C.**, **Erdmann, S.** & **Brodersen, P.** (2018) An m(6)A-YTH module controls developmental timing and morphogenesis in Arabidopsis. *Plant Cell*, **30**, 952–967.

Arribas-Hernández, L. & **Brodersen, P.** (2020) Occurrence and functions of m(6)A and other covalent modifications in plant mRNA. *Plant Physiology*, **182**, 79–96.

Artur, M.A.S., **Zhao, T.**, **Ligterink, W.**, **Schranz, E.** & **Hilhorst, H.W.M.** (2019) Dissecting the genomic diversification of late embryogenesis abundant (LEA) protein gene families in plants. *Genome Biology and Evolution*, **11**, 459–471.

Beilharz, T.H. & **Preiss, T.** (2007) Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. *RNA*, **13**, 982–997.

Bokar, J.A., **Rath-Shambaugh, M.E.**, **Ludwiczak, R.**, **Narayan, P.** & **Rottman, F.** (1994) Characterization and partial purification of mRNA N^6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. *The Journal of Biological Chemistry*, **269**, 17697–17704.

Bokar, J.A., **Shambaugh, M.E.**, **Polayes, D.**, **Matera, A.G.** & **Rottman, F.M.** (1997) Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N^6-adenosine)-methyltransferase. *RNA*, **3**, 1233–1247.

Bujnicki, J.M., **Feder, M.**, **Radlinska, M.** & **Blumenthal, R.M.** (2002) 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. *Journal of Molecular Evolution*, **55**, 431–444.

Chen, S., **Zhou, Y.**, **Chen, Y.** & **Gu, J.** (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**, i884–i890.

Duan, H.C., **Wei, L.H.**, **Zhang, C.**, **Wang, Y.**, **Chen, L.**, **Lu, Z.** ***et al.*** (2017) ALKBH10B is an RNA N(6)-methyladenosine demethylase affecting Arabidopsis floral transition. *Plant Cell*, **29**, 2995–3011.

Duncan, T., **Trewick, S.C.**, **Koivisto, P.**, **Bates, P.A.**, **Lindahl, T.** & **Sedgwick, B.** (2002) Reversal of DNA alkylation damage by two human dioxygenases. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 16660–16665.

Emms, D.M. & **Kelly, S.** (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, **20**, 238.

Fu, D. & **Samson, L.D.** (2012) Direct repair of 3,N^4-ethenocytosine by the human ALKBH2 dioxygenase is blocked by the AAG/MPG glycosylase. *DNA Repair (Amst)*, **11**, 46–52.

Fu, Y., **Dominissini, D.**, **Rechavi, G.** & **He, C.** (2014) Gene expression regulation mediated through reversible m^{6}A RNA methylation. *Nature Reviews Genetics*, **15**, 293–306.

Gao, Y., **Liu, X.**, **Jin, Y.**, **Wu, J.**, **Li, S.**, **Li, Y.** ***et al.*** (2022) Drought induces epitranscriptome and proteome changes in stem-differentiating xylem of *Populus trichocarpa*. *Plant Physiology*, **190**, 459–479.

Gao, Y., **Liu, X.**, **Wu, B.**, **Wang, H.**, **Xi, F.**, **Kohnen, M.V.** ***et al.*** (2021) Quantitative profiling of N(6)-methyladenosine at single-base resolution in stem-differentiating xylem of *Populus trichocarpa* using Nanopore direct RNA sequencing. *Genome Biology*, **22**, 1–17.

Greer, E.L., **Blanco, M.A.**, **Gu, L.**, **Sendinc, E.**, **Liu, J.**, **Aristizábal-Corrales, D.** ***et al.*** (2015) DNA methylation on N^6-adenine in *C. elegans*. *Cell*, **161**, 868–878.

Huong, T.T., **Ngoc, L.N.T.** & **Kang, H.** (2020) Functional characterization of a putative RNA demethylase ALKBH6 in Arabidopsis growth and abiotic stress responses. *International Journal of Molecular Sciences*, **21**, 6707.

Iyer, L.M., **Zhang, D.** & **Aravind, L.** (2016) Adenine methylation in eukaryotes: apprehending the complex evolutionary history and functional potential of an epigenetic modification. *BioEssays*, **38**, 27–40.

Jia, G., **Fu, Y.** & **He, C.** (2013) Reversible RNA adenosine methylation in biological regulation. *Trends in Genetics*, **29**, 108–115.

Jia, G., **Fu, Y.**, **Zhao, X.**, **Dai, Q.**, **Zheng, G.**, **Yang, Y.** ***et al.*** (2011) N^6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature Chemical Biology*, **7**, 885–887.

Jiao, Y., **Li, J.**, **Tang, H.** & **Paterson, A.H.** (2014) Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. *Plant Cell*, **26**, 2792–2802.

Jin, J., **Tian, F.**, **Yang, D.C.**, **Meng, Y.Q.**, **Kong, L.**, **Luo, J.** ***et al.*** (2017) PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research*, **45**, D1040–D1045.

Katoh, K. & **Standley, D.M.** (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.

Kim, D., **Paggi, J.M.**, **Park, C.**, **Bennett, C.** & **Salzberg, S.L.** (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, **37**, 907–915.

Leihne, V., **Kirpekar, F.**, **Vågbø, C.B.**, **van den Born, E.**, **Krokan, H.E.**, **Grini, P.E.** ***et al.*** (2011) Roles of Trm9- and ALKBH8-like proteins in the formation of modified wobble uridines in Arabidopsis tRNA. *Nucleic Acids Research*, **39**, 7688–7701.

Letunic, I. & **Bork, P.** (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, **44**, W242–W245.

Li, H. (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, **34**, 3094–3100.

Li, T., **Wang, H.**, **Zhang, Y.**, **Wang, H.**, **Zhang, Z.**, **Liu, X.** ***et al.*** (2023) Comprehensive profiling of epigenetic modifications in fast-growing moso bamboo shoots. *Plant Physiology*, **191**, 1017–1035.

Liang, Z., **Riaz, A.**, **Chachar, S.**, **Ding, Y.**, **Du, H.** & **Gu, X.** (2020) Epigenetic modifications of mRNA and DNA in plants. *Molecular Plant*, **13**, 14–30.

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

Lim, J., **Lee, M.**, **Son, A.**, **Chang, H.** & **Kim, V.N.** (2016) mTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. *Genes & Development*, **30**, 1671–1682.

Lima, S.A., **Chipman, L.B.**, **Nicholson, A.L.**, **Chen, Y.H.**, **Yee, B.A.**, **Yeo, G.W.** ***et al.*** (2017) Short poly(A) tails are a conserved feature of highly expressed genes. *Nature Structural & Molecular Biology*, **24**, 1057–1063.

Liu, H., **Begik, O.**, **Lucas, M.C.**, **Ramirez, J.M.**, **Mason, C.E.**, **Wiener, D.** ***et al.*** (2019) Accurate detection of m(6)A RNA modifications in native RNA sequences. *Nature Communications*, **10**, 4079.

Liu, Y., **Zhang, Y.**, **Lu, F.** & **Wang, J.** (2021) Interactions between RNA m^6A modification, alternative splicing, and poly(A) tail revealed by MePAIso-seq2. *bioRxiv*, https://doi.org/10.1101/2021.08.29.458071

Liufu, Y., **Xi, F.**, **Wu, L.**, **Zhang, Z.**, **Wang, H.**, **Wang, H.** ***et al.*** (2023) Inhibition of DNA and RNA methylation disturbs root development of moso bamboo. *Tree Physiology*, **43**, 1653–1674.

Luo, Q., **Mo, J.**, **Chen, H.**, **Hu, Z.**, **Wang, B.**, **Wu, J.** ***et al.*** (2022) Structural insights into molecular mechanism for N(6)-adenosine methylation by MT-A70 family methyltransferase METTL4. *Nature Communications*, **13**, 5636.

Machida, Y., **Suzuki, T.**, **Sasabe, M.**, **Iwakawa, H.**, **Kojima, S.** & **Machida, C.** (2022) Arabidopsis ASYMMETRIC LEAVES2 (AS2): roles in plant morphogenesis, cell division, and pathogenesis. *Journal of Plant Research*, **135**, 3–14.

Martínez-Pérez, M., **Aparicio, F.**, **López-Gresa, M.P.**, **Bellés, J.M.**, **Sánchez-Navarro, J.A.** & **Pallás, V.** (2017) Arabidopsis m(6)A demethylase activity modulates viral infection of a plant virus and the m(6)A abundance in its genomic RNAs. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, 10755–10760.

Meyer, K.D. & **Jaffrey, S.R.** (2017) Rethinking m(6)A readers, writers, and erasers. *Annual Review of Cell and Developmental Biology*, **33**, 319–342.

Meza, T.J., **Moen, M.N.**, **Vågbø, C.B.**, **Krokan, H.E.**, **Klungland, A.**, **Grini, P.E.** ***et al.*** (2012) The DNA dioxygenase ALKBH2 protects *Arabidopsis thaliana* against methylation damage. *Nucleic Acids Research*, **40**, 6620–6631.

Miao, Z., **Zhang, T.**, **Xie, B.**, **Qi, Y.** & **Ma, C.** (2022) Evolutionary implications of the RNA N^6-methyladenosine methylome in plants. *Molecular Biology and Evolution*, **39**, msab299.

Minh, B.Q., **Schmidt, H.A.**, **Chernomor, O.**, **Schrempf, D.**, **Woodhams, M.D.**, **von Haeseler, A.** ***et al.*** (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, **37**, 1530–1534.

Monfared, M.M., **Simon, M.K.**, **Meister, R.J.**, **Roig-Villanova, I.**, **Kooiker, M.**, **Colombo, L.** ***et al.*** (2011) Overlapping and antagonistic activities of BASIC PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. *The Plant Journal*, **66**, 1020–1031.

Nguyen, N.T.T., **Vincens, P.**, **Dufayard, J.F.**, **Roest Crollius, H.** & **Louis, A.** (2022) Genomicus in 2022: comparative tools for thousands of genomes and reconstructed ancestors. *Nucleic Acids Research*, **50**, D1025–D1031.

Parker, M.T., **Knop, K.**, **Sherwood, A.V.**, **Schurch, N.J.**, **Mackinnon, K.**, **Gould, P.D.** ***et al.*** (2020) Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m(6)A modification. *eLife*, **9**, e49658.

Passmore, L.A. & **Coller, J.** (2022) Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nature Reviews Molecular Cell Biology*, **23**, 93–106.

Piqué, M., **López, J.M.**, **Foissac, S.**, **Guigó, R.** & **Méndez, R.** (2008) A combinatorial code for CPE-mediated translational control. *Cell*, **132**, 434–448.

Robbens, S., **Rouzé, P.**, **Cock, J.M.**, **Spring, J.**, **Worden, A.Z.** & **van de Peer, Y.** (2008) The FTO gene, implicated in human obesity, is found only in vertebrates and marine algae. *Journal of Molecular Evolution*, **66**, 80–84.

Růžička, K., **Zhang, M.**, **Campilho, A.**, **Bodi, Z.**, **Kashif, M.**, **Saleh, M.** ***et al.*** (2017) Identification of factors required for m(6)A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. *The New Phytologist*, **215**, 157–172.

Salmela, L. & **Rivals, E.** (2014) LoRDEC: accurate and efficient long read error correction. *Bioinformatics*, **30**, 3506–3514.

Scutenaire, J., **Deragon, J.M.**, **Jean, V.**, **Benhamed, M.**, **Raynaud, C.**, **Favory, J.J.** ***et al.*** (2018) The YTH domain protein ECT2 is an m(6)A reader required for normal trichome branching in Arabidopsis. *Plant Cell*, **30**, 986–1005.

Shannon, P., **Markiel, A.**, **Ozier, O.**, **Baliga, N.S.**, **Wang, J.T.**, **Ramage, D.** ***et al.*** (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, **13**, 2498–2504.

Shen, L., **Liang, Z.**, **Gu, X.**, **Chen, Y.**, **Teo, Z.W.**, **Hou, X.** ***et al.*** (2016) N(6)-methyladenosine RNA modification regulates shoot stem cell fate in Arabidopsis. *Developmental Cell*, **38**, 186–200.

Song, P., **Yang, J.**, **Wang, C.**, **Lu, Q.**, **Shi, L.**, **Tayier, S.** ***et al.*** (2021) Arabidopsis N(6)-methyladenosine reader CPSF30-L recognizes FUE signals to control polyadenylation site choice in liquid-like nuclear bodies. *Molecular Plant*, **14**, 571–587.

Suyama, M., **Torrents, D.** & **Bork, P.** (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, **34**, W609–W612.

Tang, J., **Yang, J.**, **Duan, H.** & **Jia, G.** (2021) ALKBH10B, an mRNA m(6)A demethylase, modulates ABA response during seed germination in Arabidopsis. *Frontiers in Plant Science*, **12**, 712713.

Tang, J., **Yang, J.**, **Lu, Q.**, **Tang, Q.**, **Chen, S.** & **Jia, G.** (2022) The RNA N(6) -methyladenosine demethylase ALKBH9B modulates ABA responses in Arabidopsis. *Journal of Integrative Plant Biology*, **64**, 2361–2373.

Wang, P., **Doxtader, K.A.** & **Nam, Y.** (2016) Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Molecular Cell*, **63**, 306–317.

Wang, Y., **Tang, H.**, **Debarry, J.D.**, **Tan, X.**, **Li, J.**, **Wang, X.** ***et al.*** (2012) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, **40**, e49.

Wei, L.H., **Song, P.**, **Wang, Y.**, **Lu, Z.**, **Tang, Q.**, **Yu, Q.** ***et al.*** (2018) The m(6)A reader ECT2 controls trichome morphology by affecting mRNA stability in Arabidopsis. *Plant Cell*, **30**, 968–985.

Workman, R.E., **Tang, A.D.**, **Tang, P.S.**, **Jain, M.**, **Tyson, J.R.**, **Razaghi, R.** ***et al.*** (2019) Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nature Methods*, **16**, 1297–1305.

Yan, X., **Pei, K.**, **Guan, Z.**, **Liu, F.**, **Yan, J.**, **Jin, X.** ***et al.*** (2022) AI-empowered integrative structural characterization of m(6)A methyltransferase complex. *Cell Research*, **32**, 1124–1127.

Yue, H., **Nie, X.**, **Yan, Z.** & **Weining, S.** (2019) N^6-methyladenosine regulatory machinery in plants: composition, function and evolution. *Plant Biotechnology Journal*, **17**, 1194–1208.

Zhang, X., **Li, X.**, **Zhao, R.**, **Zhou, Y.** & **Jiao, Y.** (2020) Evolutionary strategies drive a balance of the interacting gene products for the CBL and CIPK gene families. *The New Phytologist*, **226**, 1506–1516.

Zhao, T., **Holmer, R.**, **de Bruijn, S.**, **Angenent, G.C.**, **van den Burg, H.A.** & **Schranz, M.E.** (2017) Phylogenomic synteny network analysis of MADS-box transcription factor genes reveals lineage-specific transpositions, ancient tandem duplications, and deep positional conservation. *Plant Cell*, **29**, 1278–1292.

Zheng, G., **Dahl, J.A.**, **Niu, Y.**, **Fedorcsak, P.**, **Huang, C.M.**, **Li, C.J.** ***et al.*** (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Molecular Cell*, **49**, 18–29.

Zhong, Z.-D., **Xie, Y.-Y.**, **Chen, H.-X.**, **Lan, Y.-L.**, **Liu, X.-H.**, **Ji, J.-Y.** ***et al.*** (2023) Systematic comparison of tools used for m6A mapping from nanopore direct RNA sequencing. *Nature Communications*, **14**, 1906.

Zhou, L., **Tian, S.** & **Qin, G.** (2019) RNA methylomes reveal the m(6)A-mediated regulation of DNA demethylase gene SlDML2 in tomato fruit ripening. *Genome Biology*, **20**, 156.