HHS Public Access

Author manuscript

Nat Commun. Author manuscript; available in PMC 2015 May 28.

Published in final edited form as:

Nat Commun. ; 5: 5630. doi:10.1038/ncomms6630.

Unique Features of the m^{6}A Methylome in *Arabidopsis thaliana*

Guan-Zheng Luo[#1,2], **Alice MacQueen**[#3], **Guanqun Zheng**[#1,2], **Hongchao Duan**[4], **Louis C Dore**[1,2], **Zhike Lu**[1,2], **Jun Liu**[4], **Kai Chen**[1,2], **Guifang Jia**[4,*], **Joy Bergelson**[3,*], and **Chuan He**[1,2,4,*]

[1]Department of Chemistry and Institute for Biophysical Dynamics, The University of Chicago, 929 East 57th Street, Chicago, Illinois 60637, USA

[2]Howard Hughes Medical Institute, The University of Chicago, 929 East 57th Street, Chicago, Illinois 60637, USA

[3]Department of Ecology and Evolution, The University of Chicago, 1101 East 57th Street, Chicago, Illinois 60637, USA

[4]Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

[#] These authors contributed equally to this work.

Abstract

Recent discoveries of reversible *N*^{6}-methyladenosine (m^{6}A) methylation on messenger RNA (mRNA) and mapping of m^{6}A methylomes in mammals and yeast have revealed potential regulatory functions of this RNA modification. In plants, defects in m^{6}A methyltransferase cause an embryo-lethal phenotype, suggesting a critical role of m^{6}A in plant development. Here, we profile m^{6}A transcriptome-wide in two accessions of *Arabidopsis thaliana* and reveal that m^{6}A is a highly conserved modification of mRNA in plants. Distinct from mammals, m^{6}A in *A. thaliana* is enriched not only around the stop codon and within 3′ untranslated regions (3′ UTRs), but also around the start codon. Gene ontology analysis indicates that the unique distribution pattern of m^{6}A in *A. thaliana* is associated with plant-specific pathways involving the chloroplast. We also discover a positive correlation between m^{6}A deposition and the mRNA abundance, suggesting a regulatory role of m^{6}A in plant gene expression.

[*]To whom correspondence should be addressed. guifangjia@pku.edu.cn (G. J.) or jbergels@uchicago.edu (J. B.) or chuanhe@uchicago.edu (C. H.) .

AUTHOR CONTRIBUTIONS

C.H., J.B. and G.J. conceived the project; C.H., J.B and G.J. designed most experiments; GZ.L., A.M. and Z.L. performed data analyses; G.J, A.M., G.Z and H.D. performed the experiments; and GZ.L., G.J. and C.H. wrote the paper with suggestions from J.B.

ACCESSION CODES

The high throughput data used in this study are deposited in the NCBI GEO database with accession number GSE59154.

COMPETING FINANCIAL INTERESTS STATEMENT

The authors declare no conflict of interest.

Keywords

N^6-methyladenosine (m^6A); RNA methylation; plant mRNA methylome; chloroplast; gene expression

INTRODUCTION

N^6-methyladenosine (m^6A) is the most prevalent internal messenger RNA (mRNA) modification[1-3] in eukaryotes including mammals[4], plants[5,6], *Drosophila*[7] and yeast[8], as well as viruses with a nuclear phase[9]. This modification is installed by N^6-adenosine methyltransferase. A 70kD SAM (S-adenosylmethionine)-binding subunit methyltransferase like 3 (*METTL3*, also called *MT-A70*) was identified as one component of a methyltransferase complex in mammalian cells[10]. Recent studies have characterized this complex, which consists of *METTL3*, methyltransferase like 14 (*METTL14*) and Wilms tumor 1 associated protein (*WTAP*)[11-13]. *METTL14* and *METTL3* are two active methyltransferases that form a heterodimer to catalyze m^6A RNA methylation, while WTAP interacts with this complex and substantially affects the mRNA methylation inside cells but not *in vitro*[11]. Knockdowns of these methyltransferases affect mouse embryonic stem cell differentiation[12]. Since 2011, two m^6A RNA demethylases of *FTO* and *ALKBH5* have been discovered; these demethylases are involved in mammalian development, RNA metabolism and fertility[14,15]. These findings reveal the first examples of reversible RNA modification and indicate regulatory functions of reversible m^6A methylation on mRNA and certain non-coding RNAs that contain m^6A[16]. Subsequent profiling of m^6A distributions in mammalian transcriptomes[17,18] and the recent mapping of the yeast m^6A methylome in the meiotic state[19] further confirm the dynamic nature of m^6A modification. These studies revealed that m^6A is enriched around the stop codon and at 3′ UTRs, as well as in long internal exons and at the transcription start site (TSS)[17-19]. Cellular proteins have also been found to preferentially bind m^6A-containing RNA[17,20]. The human YTH domain family 2 (*YTHDF2*) has recently been characterized to specifically recognize m^6A-methylated mRNA and accelerate the decay of the bound mRNA[20].

In *Arabidopsis thaliana*, the m^6A content in mRNA varies across tissues with a high ratio of m^6A/A found in flower buds[6]. This variation correlates with the expression levels of the plant methyltransferase *MTA* (the plant homolog of human *METTL3*, encoded by *At4g10760*)[6]. Previous studies have also shown that m^6A predominantly locates at the 3′ end of transcripts in a region 100-150 bp before the poly(A) tail in *A. thaliana* mRNA[21]. Inactivation of *MTA* prevents the progression of the developing embryo from passing the globular stage; an embryo-lethal phenotype with seed arrestment has been observed[6]. Reduced expression of *MTA* in *A. thaliana* leads to decreased m^6A level in mRNA and abnormal growth with reduced apical dominance, abnormal organ definition, and increased trichome branching[21]. These data demonstrate that m^6A in mRNA plays functional roles in plant development.

In order to further investigate the functions of m^6A and to facilitate future studies of m^6A in plants, we report here transcriptome-wide m^6A profiling in two accessions of *A. thaliana*,

Can-0 and Hen-16. These accessions are wild-collected natural lines from the two extremes of the natural range of photosynthetically active radiation (PAR) in the spring[22]. We show that m^{6}A is a highly conserved RNA modification in mRNA across these two accessions. Intriguingly, m^{6}A in *A. thaliana* is enriched not only around the stop codon and within 3′ UTRs, as in yeast and mammalian systems, but also around the start codon, a property distinct from other known m^{6}A methylomes[17,18]. A positive correlation between m^{6}A deposition and mRNA levels indicates a regulatory role of m^{6}A in plant gene expression.

RESULTS

m^{6}A is abundant and conserved in A. thaliana mRNA

m^{6}A is known to be a relatively abundant internal modification in *A. thaliana* mRNA[6]. We selected ten geographically diverse accessions of *A. thaliana* to grow in a common laboratory environment in order to measure the m^{6}A/A ratio of purified mRNA (Supplementary Table 1). These wild-collected natural lines were collected from sites that vary widely in PAR values[22]. We observed that the ratio of m^{6}A/A in total mRNA from these ten accessions varied within the range of 0.45-0.65% (Fig. 1a), although not directly related to PAR, suggesting that the m^{6}A methylation level in mRNA is relatively stable but potentially affected by complex environmental factors.

To obtain the transcriptome-wide m^{6}A map of whole *A. thaliana* plants, we interrogated two accessions (Can-0 and Hen-16) using the m^{6}A-targeted antibody coupled with high-throughput sequencing[17,18]. Can-0 was originally collected from the Canary Islands where PAR in spring is 123.74, the highest seen for 1,191 accessions of *A. thaliana*[22]. Hen-16 was obtained from northern Sweden where spring PAR is 55.29[22], the lowest end of the range. The m^{6}A level in mRNA isolated from Can-0 is higher than that from Hen-16 as measured by LC-MS/MS (Fig. 1a).

More than 70% of the m^{6}A peaks of Can-0 and Hen-16 were consistently detected in two biological replicates for each accession. We used these recurrent peaks as high-confidence m^{6}A sites for further analysis. In total, we identified 7,489 m^{6}A peaks representing the transcripts of 6,289 genes in Can-0, and 6,094 m^{6}A peaks representing transcripts of 5,416 genes in Hen-16 (Supplementary Data 1). Among them, 4,317 m^{6}A peaks were detected within both Can-0 and Hen-16 ($P<1e-5$, Chi-squared test), indicating that m^{6}A is highly conserved across *A. thaliana* accessions (Fig. 1b-c, Supplementary Data 1). We validated 11 m^{6}A peak containing genes by RT-qPCR and all of them showed significant enrichment in IP-pulldown samples (Supplementary Fig. 1). Based on these results, we estimated that the *A. thaliana* transcriptome contains 0.5-0.7 m^{6}A peaks per 1,000 nucleotides, or 0.7-1.0 m^{6}A peaks per actively expressed transcript (Supplementary Table 2, Supplementary Fig. 2). These levels are comparable to those obtained in mammals[17]. The relative abundance of m^{6}A peaks in mRNAs from Can-0 and Hen-16 is consistent with total m^{6}A levels measured in mRNAs isolated from these two accessions (Fig. 1a-b).

To determine if the m^{6}A peaks that we identified contained the m^{6}A consensus sequence of RRACH (where R represents purine, A is m^{6}A, and H is a non-guanine base)[23,24], we analyzed the top 1,000 most significant peaks (Supplementary Fig. 3). We found at least one

such motif in 934 peaks (Fig. 1d). The same sequence motif appears to be necessary for m^{6}A methylation in plant mRNA as has been observed in mammals and yeast mRNA (Supplementary Fig. 4)[17-19].

m^{6}A distribution exhibits a distinct topology in *A. thaliana*

We next analyzed the distribution of m^{6}A in the whole transcriptome for both strains of *A. thaliana*. We determined the distribution of m^{6}A reads along transcripts in the m^{6}A-IP and non-IP (input) samples, respectively. Intriguingly, we found that reads from m^{6}A-IP are highly enriched around the start codon, stop codon and within 3′ UTRs in both strains (Fig. 2a and Supplementary Fig. 5). The prevalence of m^{6}A-IP reads around the start codon has not been observed in mammals or yeast.

To further confirm the preferential locations of m^{6}A on transcripts, we investigated the metagene profiles of m^{6}A peaks. Consistent with the distribution of reads, m^{6}A peaks are abundant near the stop codon (61%) and start codon (16%), followed by the coding regions (CDs, 13%) and then 3′ UTRs (9%) (Fig. 2b). After segment normalization by the total length of each gene portion, we observed that m^{6}A is exclusively enriched around the start codon and stop codon (Fig. 2c). We mapped the number of m^{6}A peaks around the start codon, the stop codon and upstream of transcription termination site (TTS) (Supplementary Fig. 6a-c). The m^{6}A peaks are enriched at three locations: the region 60 bp downstream and 50 bp upstream of the start codon, the region 80 bp downstream and 100 bp upstream of the stop codon, and the region 80-220 bp before the poly(A) tail; the last location of enrichment is consistent with the previous finding that m^{6}A preferentially locates at the 3′ end of transcripts in a region 100-150 bp before the poly(A) tail in *Arabidopsis*[21]. We focused on the m^{6}A peaks near the start codon and identified several new sequence motifs by comparing peak regions with scrambled background sequences, suggesting the possibility that distinct sequence motifs may exist for m^{6}A methylation in plant mRNA (Supplementary Fig. 7).

Using strict criteria of peptide similarity, we identified 1,684 m^{6}A-containing transcripts/genes in *A. thaliana* that are conserved in humans. Through comparison with the published data[17], we found that 813 of these transcripts (48%) are also methylated in humans (Supplementary Data 2). Interestingly, almost all of the conserved methylated sites are located at the stop codon (74%) and 3′ UTRs (14%) of transcripts (Fig. 2b, 2d), thus highlighting that m^{6}A peaks around the start codon are specific to *A. thaliana*. We depicted the m^{6}A peak distribution along transcripts in available datasets of human and mouse, confirming the observation that m^{6}A is exclusively enriched at the stop codon and 3′ UTRs in these systems (Supplementary Fig. 8). Clearly, the enrichment of m^{6}A near the start codon in *A. thaliana* is distinct. This feature of the m^{6}A distribution may play a unique role in plant-specific pathways.

m^{6}A-containing mRNAs in important biological pathways

The presence of m^{6}A is critical for normal plant development[6,21]. In order to uncover further functional insights about m^{6}A in *A. thaliana*, we selected genes containing m^{6}A in both Can-0 and Hen-16, and identified the enriched gene ontology (GO) terms using the

DAVID tool. We found that these genes are highly enriched in chloroplast/plastid and protein transport/localization categories (Fig. 3a).

We next sought to determine if the unique m^{6}A position patterns are related to plant-specific GO categories. We classified genes into four subgroups according to the distribution of m^{6}A peaks: PeakStart (m^{6}A peaks around start codon), PeakStop (m^{6}A peaks around stop codon), PeakBoth (m^{6}A peaks around both start and stop codons) and others (Fig. 3b and Supplementary Data 1). Then we performed GO enrichment analysis for each subgroup. All subgroups of m^{6}A-containing genes exhibit high enrichment of chloroplast-related cellular components, indicating that a large proportion of these genes produce proteins localized in chloroplast though encoded by nuclear genome (Supplementary Fig. 9). Strikingly, more than 60% of genes belonging to the PeakBoth subgroup could be attributed to chloroplast components (Fig. 3c). Similarly, about 40% of the genes belonging to the PeakStart subgroup are associated with chloroplast. Photosynthesis is one of the most important functions in plants; 10%-20% of nuclear genes encode proteins that are imported to chloroplast[25]. Apparently, genes with m^{6}A enrichment around the start codon are highly enriched in chloroplast components.

The significant enrichment of chloroplast-related GO categories among transcripts with m^{6}A peaks around the start codon prompted us to examine photosynthesis-related genes in PeakStart and PeakBoth subgroups. Indeed, we identified dozens of well-studied photosynthesis-related genes that carry m^{6}A peaks (Supplementary Data 3). For instance, AT5G01920 (or STN8) is an important chloroplast thylakoid protein kinase that is specific to the phosphorylation of N-terminal threonine residues in D1, D2 and CP43 proteins and Thr-4 in PsbH of the photosystem II[26]. In our m^{6}A-IP data, the transcript of STN8 contains two clear m^{6}A peaks around the start and stop codons (Fig. 3d). The large fraction of m^{6}A-containing genes associated with chloroplast suggests a relationship between m^{6}A mRNA methylation and photosynthesis, one of the defining processes of plants.

Strain-specific m^{6}A marking of mRNAs

Although most m^{6}A peaks are shared between the two *A. thaliana* strains, we could detect a proportion of strain-specific peaks, even using strict criteria (Materials and Methods). In total, we identified 1,319 Can-0-specific peaks and 546 Hen-16-specific peaks (Supplementary Data 4). Furthermore, some of the common m^{6}A peaks in these two strains showed altered intensity (Fig. 4). GO analysis indicated that the genes with dynamic m^{6}A peaks are enriched in several categories of fundamental biological functions including mRNA metabolic process, response to stimulus, and regulation of translational elongation (Supplementary Data 5). For instance, *AT5G06290* (or *2-Cys Prx B*) has been shown to be sensitive to light intensities and oxidative stress[27]. In our m^{6}A-IP data, we observed significant m^{6}A peaks around the start codon of *AT5G06290* in Can-0 but not in Hen-16 (Fig. 4).

As an initial exploration into the functional implications of m^{6}A methylation differences across genomes, we asked whether m^{6}A methylation could underlie expression differences. Using the RNA-Seq data, we calculated gene expression to assign differentially expressed genes (DE genes) of the two strains. Indeed, we found 801 more highly expressed genes in

Can-0 (Can-high), and 1,011 more highly expressed genes in Hen-16 (Hen-high) (Fig. 5a and Supplementary Data 6), perhaps reflecting, at least in part, the substantial differences in PAR in the native habitat of these accessions. Within the Can-high list, we detected many more genes that contained m^{6}A peaks in Can-0 than in Hen-16 (195/81, P<0.01, Fisher's exact test). Correspondingly, in the list of Hen-high genes, we detected more genes containing m^{6}A peaks in Hen-16 than in Can-0 (221/140, P<0.01, Fisher's exact test) (Fig. 5a). This observation suggests that each strain possesses its own characteristic m^{6}A methylation sites that appear to be associated with gene activation. This hypothesis seems to contradict the recent discovery that a main function of m^{6}A is to mediate mRNA degradation in mammalian cells[11-13,20]. However, when we checked positions of the m^{6}A peaks involved in up-regulated genes, we found that a large portion of these peaks are located at the 5′ end of the corresponding DE genes (Supplementary Data 6). To confirm this observation, we divided strain specific m^{6}A-containing genes into two groups and examined how their expression levels correlate with the locations of m^{6}A peaks. Our analysis showed that m^{6}A peaks at the 5′ end of transcripts correlate with higher expression levels of each strain. Genes in the PeakStart category possesses higher overall expression levels and correlate well with strain-specific m^{6}A peaks (Fig. 5b, note that the gene expression ratios of Can-0/Hen-16 are shown). We further examined the fraction of each subgroup of genes based on their expression levels. Genes with both 5′ and 3′ m^{6}A peaks (PeakBoth) are enriched in the high-expression fraction, while genes with m^{6}A peaks at other locations (PeakOther) tend to be lower-expressed (Fig. 5c). Thus, m^{6}A enrichment at the 5′ end of the plant transcripts correlates with higher expression level in *A. thaliana* in general, perhaps by stabilizing the transcripts via reader proteins or interaction with translation machineries.

In a recent study, Bodi and colleagues generated an m^{6}A reduction mutant of *A. thaliana* and profiled gene expression[21]. They identified 883 up-regulated and 654 down-regulated genes in the mutant plant[21]. By comparing the differentially expressed genes with our list of m^{6}A peaks, we found that 116 among 883 up-regulated transcripts are modified by m^{6}A whereas 142 among 654 down-regulated transcripts contain m^{6}A (Supplementary Data 7). These data suggest that transcripts carrying m^{6}A tend to be down-regulated in the m^{6}A reduction mutant (P<0.01, Chi-squared test). We extended this analysis to the whole transcriptome. Indeed, among genes with m^{6}A modification, more are down-regulated than up-regulated in the m^{6}A reduction mutant plant. This trend did not exist for genes without m^{6}A modification (Fig. 5d). Together, their data supported our hypothesis that m^{6}A tends to positively correlate with gene expression in a large fraction of transcripts in *A. thaliana*.

Mammalian microRNA (miRNA)-binding sites are mostly found within 3′ UTRs[28]; however, in plants they occur typically in the coding regions of target transcripts[29]. Unlike mammals, miRNAs in plants recognize the target sites by near-perfect sequence complementarity. By comparing the transcripts containing m^{6}A with the public miRNA targets database (TAIR)[30], we found that only 5% of known miRNA-targeted transcripts contain m^{6}A. Furthermore, we employed a *de novo* miRNA target prediction method[31] to inspect m^{6}A peak regions. Among the top 1,000 most significant m^{6}A peak regions, only 11 regions could be potentially targeted by miRNA (Supplementary Data 8). Taken together,

our results indicate that m^{6}A is unlikely to directly impact miRNA binding sites in *A. thaliana* although such an association has been proposed in animals[18]. However, we could not exclude the possibilities that m^{6}A may affect miRNA maturation, or impact miRNA targeting sites through reader proteins of m^{6}A or structural changes induced by methylation.

DISCUSSION

The discovery of m^{6}A demethylases and the mapping of the m^{6}A methylomes in mammalian systems indicate that m^{6}A methylation of mRNA is a reversible and dynamic process with regulatory functions[14-18,32,33]. The importance of m^{6}A in post-transcriptional regulation of gene expression is further reinforced by the discovery and characterization of mammalian reader proteins that recognize m^{6}A modifications of mRNA and subsequently affect the stability of the target transcripts[20]. Previous studies have shown that m^{6}A plays a critical role in plant development[6,21]. Here, we report the transcriptome-wide m^{6}A distributions in two accessions of *A. thaliana*, Can-0 and Hen-16. We found that m^{6}A is highly conserved across *A. thaliana* accessions. The methylation sites in a portion of transcripts are also conserved in corresponding transcripts of humans, indicating a fundamental functional role of m^{6}A in eukaryotes. Nevertheless, there are differences between the two *A. thaliana* accessions with a higher total m^{6}A level in Can-0 than Hen-16, and with ~1,400 more m^{6}A peaks identified in Can-0. The m^{6}A distribution could be influenced by differences in the accession's climates of origin and genetic backgrounds. The modification sites could be conserved but the modification fraction at each site could vary depending on environmental factors.

Importantly, we discovered features of the m^{6}A distribution in *A. thaliana* mRNA that are distinct from those of mammals: i) m^{6}A in both accessions is enriched around the stop codon and at 3′ UTRs, as has been found in mammals, but also around the start codon; ii) the m^{6}A methylation around the start codon is heavily associated with the chloroplast, a photosynthesis organelle in plants. In addition, m^{6}A is less likely to be directly associated with microRNA recognition sites in *A. thaliana*. These distinct differences, namely the start codon enrichment of m^{6}A and its association with chloroplast, strongly suggest additional, plant-specific functions of this mRNA methylation. Previous studies indicated that the currently available m^{6}A antibody could also recognize *N*^{6}-,2′-O-dimethyladenosine (m^{6}Am)[17,19], which might lead to enrichment peaks at TSS. However, in our results, the start codon enrichment is clearly distinct from the m^{6}Am peak observed previously[17] (Fig. 1). More accurate, base-resolution methods are highly desirable to determine the exact sites and modification fractions in the future.

Noticeably, the distinct enrichment of m^{6}A around the start codon correlates with the overall up-regulation of mRNA expression level. This relationship contradicts observations in mammalian systems, in which m^{6}A methylation around the stop codon and at 3′ UTRs is negatively correlated with gene expression[11,12,20]. Published microarray data using a different accession of *Arabidopsis* (Col-0) also supports our findings with high correlations with our RNA-seq data (Supplementary Fig. 10)[21]. Although our analysis demonstrates significant overlap of m^{6}A sites between the two ecotypes (Can-0 and Hen-16), more precise analysis would benefit from future studies on coincident accessions. Our results

suggest that m^{6}A reader proteins may exist in *A. thaliana* to recognize m^{6}A at the 5′ end and subsequently affect the stability of the target mRNA. This function could directly impact translation through the methylation itself or through the reader protein(s). It should be noted that several RNA binding proteins have been pulled down as potential "m^{6}A readers" with a few biochemically confirmed to specifically recognize methylated mRNAs[17,20]. In particular, *YTHDF2* has been shown to bind and accelerate the decay of m^{6}A-modified mRNA[20]. However, the functions of other m^{6}A readers are still unknown; some of them could promote translation of a specific set of transcripts. Our discovery suggests the versatile roles of m^{6}A in plants beyond mediating mRNA decay.

It has been demonstrated that m^{6}A methyltransferase in *A. thaliana* is critical for normal plant development[6,21]. We found that transcripts with m^{6}A are highly enriched in chloroplast/plastid and protein transport/localization categories, indicating a mechanism by which m^{6}A affects plant-specific metabolism. Past studies have demonstrated that light and circadian cycles impact the stability and translation of specific plant transcripts[34-37]. However, the mechanisms of post-transcriptional gene regulation in response to light availability have not been clearly understood. The stability and translation of chloroplast mRNAs are known to be regulated by RNA-binding proteins that reside at 5′ UTRs and 3′ UTRs[38].

Intriguingly, our data show that more than 60% of mRNAs containing m^{6}A at both the start and stop codons encode proteins that could be components of the chloroplast (Fig. 3c). Given that 10%-20% of nuclear-encoded genes are associated with the chloroplast and photosynthesis in plant[25], this percentage represents a significant enrichment. It is possible that the dynamic m^{6}A methylation at 5′ and 3′ ends modulate the RNA affinities of RNA-binding proteins, thereby controlling mRNA transport and localized expression. Technologies that can provide more accurate measurements of the m^{6}A sites are required in the future to gain deeper insights (e.g., splicing), the first m^{6}A transcriptome-wide map of a plant species *A. thaliana* presented here provides a starting roadmap for uncovering m^{6}A functions that may affect/control plant metabolism in the future.

METHODS

Plant Material

Seeds from the Can-0 and Hen-16 accessions of *Arabidopsis thaliana* were sown in 50:50 Metromix 200:Farfad C2 soil in 48-cell flats and stratified for 5 days at 4 °C to synchronize germination. Seedlings were germinated in controlled-environment growth chambers at the University of Chicago greenhouses on a 16-hour light, 8-hour dark cycle. Plants were thinned between days 5-7 of growth. Above-ground tissue was harvested between the fifth and seventh hour of the light cycle on day 21. The tissue was flash frozen in liquid nitrogen, ground using a mortar and pestle, and stored at −80 °C.

High-throughput m^{6}A sequencing

To obtain sufficient (10 mg) total RNA for immunoprecipitation of m^{6}A-containing mRNA, approximately 200 plants from each accession were harvested and pooled in 0.6 and 2.1 g

quantities. RNA was extracted using a protocol modified from a published procedure[39]; reactions took place in 15 ml or 50 ml conicals, and reagents were scaled up linearly with respect to the increased tissue mass, with 20 or 70 times the amount of each reagent, respectively. RNA was tested for quality via Nanodrop and gel electrophoresis. Polyadenylated RNA was extracted using FastTrack MAG Maxi mRNA isolation kit (Invitrogen). RNA was randomly fragmented to ~200 nt by RNA Fragmentation Reagents (Ambion). Fragmented RNA was incubated for 2 h at 4 °C with m^{6}A antibody (Synaptic Systems Cat. No. 202003, diluted to 0.5 μg μl^{-1}) in IP buffer (50 mM Tris-HCl, 750 mM NaCl and 0.5% Igepal CA-630) supplemented with BSA (0.5 μg μl^{-1}). The mixture was then incubated with protein-A beads and eluted with elution buffer (1× IP buffer and 6.7 mM m^{6}A). Eluted RNA was precipitated by 75% ethanol. The eluted RNA was treated with RNasin (Ambion Cat No. AM2694) according to the manufacturer's instructions. TruSeq Stranded mRNA Sample Prep Kit (Illumina) was used to construct the library from immunoprecipitated RNA and input RNA according to a published protocol[40]. Sequencing was done on an Illumina HiSeq machine with 2x100 cycles Solexa paired-end sequencing.

RT-qPCR validation for m^{6}A enriched genes

Eleven genes enriched in the m^{6}A IP and six genes not differentially expressed between the IP and non-IP samples were tested by RT-qPCR. Data cleanup and analysis proceeded as described by previous protocol[41]. The dissociation curves for each reaction were plotted and those with irregular features were removed. RNA passed through a beads only column, which should not bind any RNA, was treated as the input control for the IP step. Ct values from qPCR on the flow-through from the m^{6}A IP and the m^{6}A IP were expressed as the percent input of the beads only sample. Primer sequences are listed in Supplementary Data 9. Detailed information for plotting the qPCR figure can be found in an online manual at: http://www.lifetechnologies.com/us/en/home/life-science/epigenetics-noncoding-rna-research/chromatin-remodeling/chromatin-immunoprecipitation-chip/chip-analysis.html

Data Analysis

Sequence data were analyzed according to the procedure described by Meyer *et al* [18]. Briefly, Tophat[42] with Bowtie[43] was run in order to align the input and IP sequenced samples to the Columbia reference genome and annotation file (Tair10)[30]. The BEDTools tool[44] was used to divide the aligned accepted hits into 1 bp intervals. The read depth was then averaged for each 25 nt discrete, non-overlapping genomic window using an ad hoc R program (Supplementary Data 10). To identify 25 nt windows enriched for m^{6}A, the number of reads that mapped to each window for the IP and input sample, and the total reads for each, were compared using Fisher's exact tests and corrected for multiple testing using Benjamini-Hochberg to reduce FDR to 0.05. To determine which of these windows cluster to form distinct peaks, we concatenated adjacent significant windows together and filtered out peaks less than 100 nt in length. Significant peaks with FDR<0.05 in Can-0, Hen-16, or both were annotated using an ad hoc R script (Supplementary Data 10). IntersectBED with the Columbia reference genome and annotation file was used to further annotate this set of significant peaks[44]. Peaks that shared more than 50% overlapping length were defined as recurrent peaks. For a peak to be classified as strain-specific, it should not overlap (1 nucleotide) any peak in any two replicates of the other strain. Sequence motifs were

identified by using Homer[45]. Gene expression was calculated by Cufflinks[42] using the input sequencing reads. Cuffdiff[42] was used to find the differentially expressed genes (DE genes) between Can-0 and Hen-16. Gene function analysis (GO enrichment) was performed with the DAVID tool[46]. Plant miRNA targets were predicted by psRobot[31]. Microarray data were downloaded from GEO (accession ID: GSE349243) and RMA method[47] was introduced to calculate gene expression and differentially expressed genes between *mta* mutant and wide type.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work is supported by NIH GM083068 (J. B.), Howard Hughes Medical Institute (C. H.) and National Natural Science Foundation of China 21210003 (C. H. & G. J) and 21372022 (G. J). We thank S. F. Reichard and Ian Roundtree for editing the manuscript.

REFERENCES

1. Fu Y, Dominissini D, Rechavi G, He C. Gene expression regulation mediated through reversible m(6)A RNA methylation. Nature reviews. Genetics. 2014; 15:293–306. doi:10.1038/nrg3724.
2. Machnicka MA, et al. MODOMICS: a database of RNA modification pathways--2013 update. Nucleic acids research. 2013; 41:D262–267. doi:10.1093/nar/gks1007. [PubMed: 23118484]
3. Meyer KD, Jaffrey SR. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. Nature reviews. Molecular cell biology. 2014; 15:313–326. doi:10.1038/nrm3785. [PubMed: 24713629]
4. Wei CM, Gershowitz A, Moss B. Methylated nucleotides block 5′ terminus of HeLa cell messenger RNA. Cell. 1975; 4:379–386. [PubMed: 164293]
5. Nichols JL. 'Cap' structures in maize poly(A)-containing RNA. Biochimica et biophysica acta. 1979; 563:490–495. [PubMed: 223642]
6. Zhong S, et al. MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. The Plant cell. 2008; 20:1278–1288. doi:10.1105/tpc.108.058883. [PubMed: 18505803]
7. Levis R, Penman S. 5′-terminal structures of poly(A)+ cytoplasmic messenger RNA and of poly(A)+ and poly(A)-heterogeneous nuclear RNA of cells of the dipteran Drosophila melanogaster. Journal of molecular biology. 1978; 120:487–515. [PubMed: 418182]
8. Clancy MJ, Shambaugh ME, Timpte CS, Bokar JA. Induction of sporulation in Saccharomyces cerevisiae leads to the formation of N6-methyladenosine in mRNA: a potential mechanism for the activity of the IME4 gene. Nucleic acids research. 2002; 30:4509–4518. [PubMed: 12384598]
9. Krug RM, Morgan MA, Shatkin AJ. Influenza viral mRNA contains internal N6-methyladenosine and 5′-terminal 7-methylguanosine in cap structures. Journal of virology. 1976; 20:45–53. [PubMed: 1086370]
10. Bokar JA, Rath-Shambaugh ME, Ludwiczak R, Narayan P, Rottman F. Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. The Journal of biological chemistry. 1994; 269:17697–17704. [PubMed: 8021282]
11. Liu J, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nature chemical biology. 2014; 10:93–95. doi:10.1038/nchembio.1432. [PubMed: 24316715]
12. Wang Y, et al. N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nature cell biology. 2014; 16:191–198. doi:10.1038/ncb2902. [PubMed: 24394384]

13. Ping XL, et al. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell research. 2014; 24:177–189. doi:10.1038/cr.2014.3. [PubMed: 24407421]
14. Jia G, et al. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nature chemical biology. 2011; 7:885–887. doi:10.1038/nchembio.687. [PubMed: 22002720]
15. Zheng G, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Molecular cell. 2013; 49:18–29. doi:10.1016/j.molcel.2012.10.015. [PubMed: 23177736]
16. Jia G, Fu Y, He C. Reversible RNA adenosine methylation in biological regulation. Trends in genetics: TIG. 2013; 29:108–115. doi:10.1016/j.tig.2012.11.003. [PubMed: 23218460]
17. Dominissini D, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature. 2012; 485:201–206. doi:10.1038/nature11112. [PubMed: 22575960]
18. Meyer KD, et al. Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell. 2012; 149:1635–1646. doi:10.1016/j.cell.2012.05.003. [PubMed: 22608085]
19. Schwartz S, et al. High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell. 2013; 155:1409–1421. doi:10.1016/j.cell.2013.10.047. [PubMed: 24269006]
20. Wang X, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature. 2014; 505:117–120. doi:10.1038/nature12730. [PubMed: 24284625]
21. Bodi Z, et al. Adenosine Methylation in Arabidopsis mRNA is Associated with the 3′ End and Reduced Levels Cause Developmental Defects. Frontiers in plant science. 2012; 3:48. doi: 10.3389/fpls.2012.00048. [PubMed: 22639649]
22. Hancock AM, et al. Adaptation to climate across the Arabidopsis thaliana genome. Science. 2011; 334:83–86. doi:10.1126/science.1209244. [PubMed: 21980108]
23. Wei CM, Gershowitz A, Moss B. 5′-Terminal and internal methylated nucleotide sequences in HeLa cell mRNA. Biochemistry. 1976; 15:397–401. [PubMed: 174715]
24. Schibler U, Kelley DE, Perry RP. Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. Journal of molecular biology. 1977; 115:695–714. [PubMed: 592376]
25. Soll J, Schleiff E. Protein import into chloroplasts. Nature reviews. Molecular cell biology. 2004; 5:198–208. doi:10.1038/nrm1333. [PubMed: 14991000]
26. Vainonen JP, Hansson M, Vener AV. STN8 protein kinase in Arabidopsis thaliana is specific in phosphorylation of photosystem II core proteins. The Journal of biological chemistry. 2005; 280:33679–33686. doi:10.1074/jbc.M505729200. [PubMed: 16040609]
27. Horling F, et al. Divergent light-, ascorbate-, and oxidative stress-dependent regulation of expression of the peroxiredoxin gene family in Arabidopsis. Plant physiology. 2003; 131:317–325. doi:10.1104/pp.010017. [PubMed: 12529539]
28. Stark A, Brennecke J, Bushati N, Russell RB, Cohen SM. Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3′UTR evolution. Cell. 2005; 123:1133–1146. doi:10.1016/j.cell.2005.11.023. [PubMed: 16337999]
29. Axtell MJ, Bowman JL. Evolution of plant microRNAs and their targets. Trends in plant science. 2008; 13:343–349. doi:10.1016/j.tplants.2008.03.009. [PubMed: 18502167]
30. Lamesch P, et al. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic acids research. 2012; 40:D1202–1210. doi:10.1093/nar/gkr1090. [PubMed: 22140109]
31. Wu HJ, Ma YK, Chen T, Wang M, Wang XJ. PsRobot: a web-based plant small RNA meta-analysis toolbox. Nucleic acids research. 2012; 40:W22–28. doi:10.1093/nar/gks554. [PubMed: 22693224]
32. Nilsen TW. Molecular biology. Internal mRNA methylation finally finds functions. Science. 2014; 343:1207–1208. doi:10.1126/science.1249340. [PubMed: 24626918]
33. Fu Y, Dominissini D, Rechavi G, He C. Gene expression regulation mediated through reversible mA RNA methylation. Nature reviews. Genetics. 2014 doi:10.1038/nrg3724.

34. Dickey LF, Petracek ME, Nguyen TT, Hansen ER, Thompson WF. Light regulation of Fed-1 mRNA requires an element in the 5′ untranslated region and correlates with differential polyribosome association. The Plant cell. 1998; 10:475–484. [PubMed: 9501119]
35. Gutierrez RA, Ewing RM, Cherry JM, Green PJ. Identification of unstable transcripts in Arabidopsis by cDNA microarray analysis: rapid decay is associated with a group of touch- and specific clock-controlled genes. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99:11513–11518. doi:10.1073/pnas.152204099. [PubMed: 12167669]
36. Tang L, Bhat S, Petracek ME. Light control of nuclear gene mRNA abundance and translation in tobacco. Plant physiology. 2003; 133:1979–1990. doi:10.1104/pp.103.029686. [PubMed: 14681536]
37. Juntawong P, Bailey-Serres J. Dynamic Light Regulation of Translation Status in Arabidopsis thaliana. Frontiers in plant science. 2012; 3:66. doi:10.3389/fpls.2012.00066. [PubMed: 22645595]
38. Monde RA, Schuster G, Stern DB. Processing and degradation of chloroplast mRNA. Biochimie. 2000; 82:573–582. [PubMed: 10946108]
39. Onate-Sanchez L, Vicente-Carbajosa J. DNA-free RNA isolation protocols for Arabidopsis thaliana, including seeds and siliques. BMC research notes. 2008; 1:93. doi: 10.1186/1756-0500-1-93. [PubMed: 18937828]
40. Dominissini D, Moshitch-Moshkovitz S, Salmon-Divon M, Amariglio N, Rechavi G. Transcriptome-wide mapping of N(6)-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing. Nature protocols. 2013; 8:176–189. doi:10.1038/nprot. 2012.148. [PubMed: 23288318]
41. Rieu I, Powers SJ. Real-time quantitative RT-PCR: design, calculations, and statistics. The Plant cell. 2009; 21:1031–1033. doi:10.1105/tpc.109.066001. [PubMed: 19395682]
42. Trapnell C, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols. 2012; 7:562–578. doi:10.1038/nprot.2012.016. [PubMed: 22383036]
43. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature methods. 2012; 9:357–359. doi:10.1038/nmeth.1923. [PubMed: 22388286]
44. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–842. doi:10.1093/bioinformatics/btq033. [PubMed: 20110278]
45. Heinz S, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell. 2010; 38:576–589. doi:10.1016/j.molcel.2010.05.004. [PubMed: 20513432]
46. Huang da W, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome biology. 2007; 8:R183. doi: 10.1186/gb-2007-8-9-r183. [PubMed: 17784955]
47. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004; 20:307–315. doi:10.1093/bioinformatics/btg405. [PubMed: 14960456]

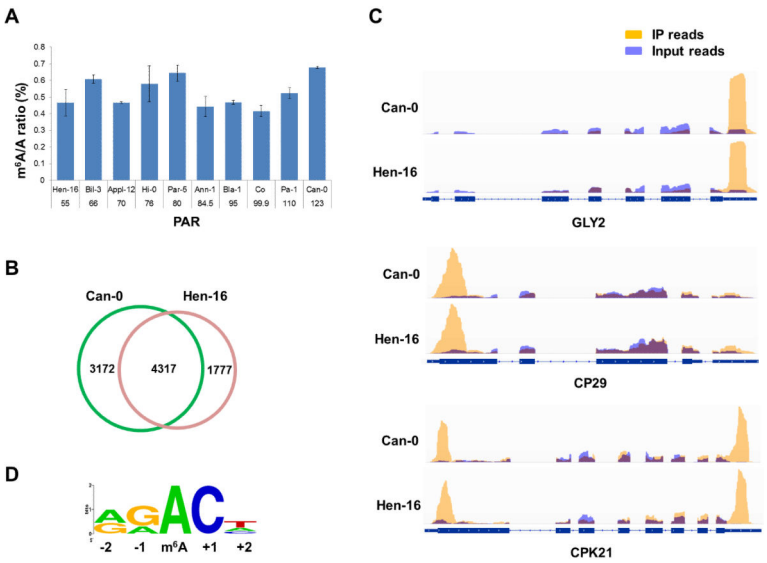

Figure 1. Overview of m^{6}A methylome in *A. thaliana*

(a) The m^{6}A/A ratio of mRNA isolated from each *A. thaliana* strain. Error bars are calculated as the standard deviation from three replicates. PAR values are displayed below the strain names. (b) Numbers of strain-specific and common m^{6}A peaks. (c) Examples of m^{6}A peaks conserved between Can-0 and Hen-16. Orange color represents IP reads while blue color represents input reads. The purple color comes from mixing orange with blue. (d) The RRACH conserved sequence motif for m^{6}A-containing peak regions.

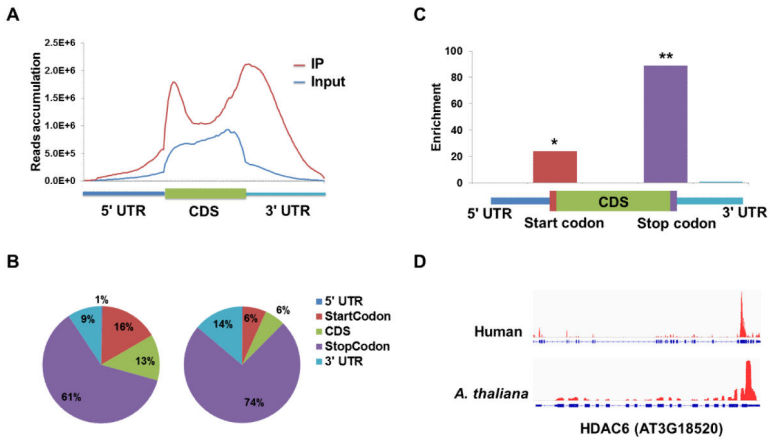

Figure 2. Distribution pattern of m^{6}A peaks along transcripts

(a) Accumulation of m^{6}A-IP reads along transcripts. Each transcript is divided into 3 parts: 5′ UTRs, CDs and 3′ UTRs. (b) The m^{6}A peak distribution within different gene contexts. Left panel: total genes with m^{6}A peaks; right panel: genes conserved in human and *Arabidopsis*. (c) The m^{6}A peak distribution along a metagene. Enrichment scores are calculated as. *n*: number of peaks belonging to each category; *N*: number of total peaks; *p*:proportion of each category within the genome by length. *P<2.2e-16, **P<1e-30. P-values are determined by Chi-squared test. (d) An example of homologous genes with m^{6}A peaks conserved in human and *A. thaliana*.

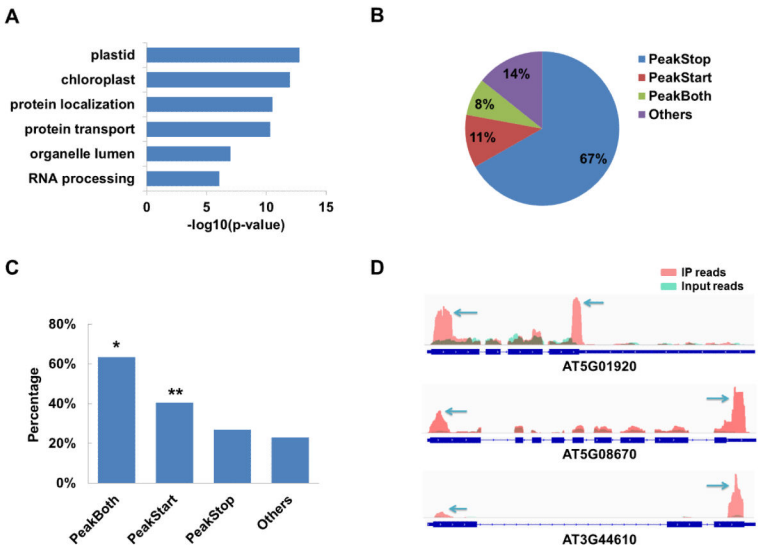

Figure 3. Functional annotation of genes with m^{6}A

(a) GO enrichment analysis of all the genes with m^{6}A peaks. GO categories are maintained by Gene Ontology Consortium. P-values are calculated using the DAVID tool. (b) Percentages of subgroups of genes divided by the position pattern of m^{6}A peaks. (c) Percentages of genes characterized as chloroplast-related for each subgroup. *P<6.0e-25, **P<5.2e-18. P-values are calculated using the DAVID tool. (d) Examples of chloroplast genes with m^{6}A peaks at both the start and stop codon. The m^{6}A-IP peaks are indicated by arrows.

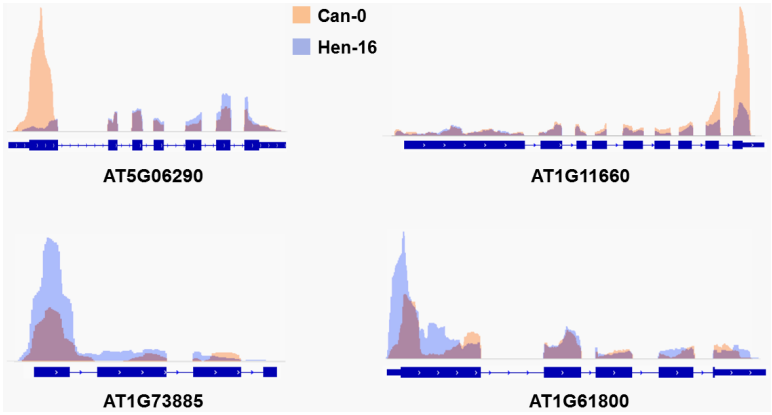

Figure 4. Dynamic m^{6}A peaks in two *Arabidopsis* strains

Different colors illustrate the accumulation of m^{6}A-IP reads from two accessions.

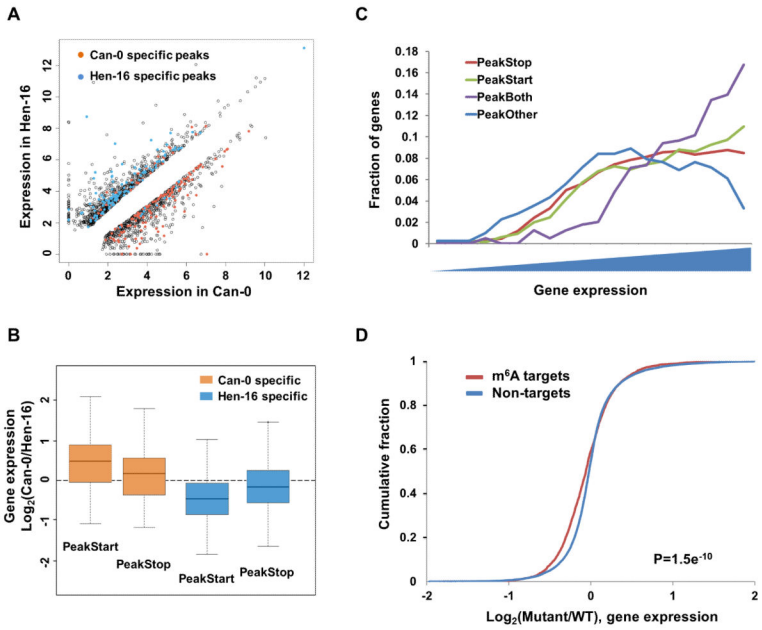

Figure 5. Relationship between m^{6}A peaks and mRNA level

(a) Differentially expressed mRNAs in Can-0 and Hen-16. Genes with Can-0-specific m^{6}A peaks are highlighted in orange, and genes with Hen-0-specific m^{6}A peaks are highlighted in blue. (b) The ratio of mRNA expression levels in two samples containing strain-specific m^{6}A peaks. Genes are divided to two categories (PeakStart and PeakStop) according to the peak positions. (c) Fraction of genes belonging to each subgroup defined by the m^{6}A distribution pattern. Genes are sorted by expression levels. (d) Cumulative distribution of mRNA expression changes between the *mta* mutant and WT for m^{6}A modified genes (red) and non-target genes (blue). P-values are calculated by two-sided Mann-Whitney test.