RNA $N^6$-methyladenosine methylation in post-transcriptional gene expression regulation

Yanan Yue,$^{1,2}$ Jianzhao Liu,$^{1,2}$ and Chuan He$^{1,2}$

$^1$Department of Chemistry, Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, USA; $^2$Howard Hughes Medical Institute, The University of Chicago, Chicago, Illinois 60637, USA

$N^6$-methyladenosine (m$^6$A) is the most prevalent and internal modification that occurs in the messenger RNAs (mRNA) of most eukaryotes, although its functional relevance remained a mystery for decades. This modification is installed by the m$^6$A methylation “writers” and can be reversed by demethylases that serve as “erasers.” In this review, we mainly summarize recent progress in the study of the m$^6$A mRNA methylation machineries across eukaryotes and discuss their newly uncovered biological functions. The broad roles of m$^6$A in regulating cell fates and embryonic development highlight the existence of another layer of epigenetic regulation at the RNA level, where mRNA is subjected to chemical modifications that affect protein expression.

Both DNA and histone proteins undergo dynamic and reversible chemical modifications to control gene expression (Strahl and Allis 2000; Bird 2001; Suzuki and Bird 2008; Bhutani et al. 2011; Jones 2012; Kohli and Zhang 2013). Although post-transcriptional modifications are known to occur to RNAs, the impact of these modifications on gene expression regulation has only recently begun to be explored (He 2010). To date, more than a hundred structurally distinct chemical modifications have been found in eukaryotic RNAs (Cantara et al. 2011; Machnicka et al. 2013), however, the enzymes responsible for each modification and the biological consequences of these modified RNAs are largely unknown. RNA modifications were once considered to be static, but a flurry of recent discoveries has demonstrated that some chemical modifications can be dynamic and participate in the regulation of diverse physiological processes (Motorin and Helm 2011; Yi and Pan 2011; Chan et al. 2012; Fu et al. 2014; Meyer and Jaffrey 2014; Kirchner and Ignatova 2015). The presence of $N^6$-methyladenosine (m$^6$A) in pol-yadenylated mRNA was first discovered in the 1970s (DesRosiers et al. 1974; Perry and Kelley 1974; Lavi and Shatkin 1975; Wei et al. 1975; Schibler et al. 1977; Wei and Moss 1977) by researchers who were characterizing the $5'$ cap structure of messenger RNA (mRNA) in mammalian cells. Since then, m$^6$A has been identified as the most prevalent internal modification in mRNA and long noncoding RNA (lncRNA) in higher eukaryotes. It is widely conserved among eukaryotic species that range from yeast, plants, and flies to mammals as well as among viral mRNAs that replicate inside host nuclei (Krug et al. 1976; Beemon and Keith 1977; Horowitz et al. 1984; Bokar 2005). In addition to its occurrence in mRNA, m$^6$A also exists in various classes of RNA in eukaryotes, bacteria, and archaea, including ribosomal RNAs, small nuclear RNAs, and transfer RNAs (Bjork et al. 1987; Maden 1990; Shimba et al. 1995; Gu et al. 1996; Agris et al. 2007; Piekna-Przybylska et al. 2008). Despite its widespread distribution in the mammalian transcriptome (on average, approximately three m$^6$A sites per mRNA), functional insight has been lacking, possibly due to the low abundance of m$^6$A mRNA and technical difficulties in global detection.

Interest in the biological relevance of m$^6$A in mRNA resurgence after the discovery of two mammalian RNA demethylases, FTO [fat mass and obesity-associated protein] (Jia et al. 2011) and its homolog, ALKBH5 (Zheng et al. 2013), which selectively reverse m$^6$A to adenosine in nuclear RNA. FTO is associated with human obesity (Dina et al. 2007; Frayling et al. 2007; Loos and Yeo 2014) and mental development (Hess et al. 2013), while ALKBH5 is shown to affect mouse spermatogenesis in a demethylation-dependent manner (Zheng et al. 2013), suggesting broad roles of m$^6$A in various physiological processes. Shortly after these findings, YTHDF2 (YTH domain-containing family protein 2) was identified as the first m$^6$A reader protein that preferentially recognizes...
m^6^A-containing mRNA (Dominissini et al. 2012; Wang et al. 2014a) and mediates mRNA decay (Wang et al. 2014a), thereby suggesting a role for m^6^A RNA as a negative regulator of gene expression. On the other hand, a transcriptome-wide m^6^A profiling method was developed to decipher the m^6^A RNA landscape (Dominissini et al. 2012; Meyer et al. 2012). Intriguingly, m^6^A sites in mammalian polyadenylated RNA are dominated by the conserved Pu[G > A]m^6^AC[A/C/U] motif that localizes near stop codons, in 3′ untranslated regions (UTRs), within long internal exons, and at 5′ UTRs (Dominissini et al. 2012; Meyer et al. 2012; Schwartz et al. 2013; Li et al. 2014; Luo et al. 2014), immediately raising the question of how this specificity is achieved. The m^6^A RNA landscape is initially sculptured by a methyltransferase complex, but for a long time, METTL3 (methyltransferase-like 3) was the only known SAM (S-adenosyl methionine)-binding subunit associated with mRNA methylation (Bokar et al. 1997). In 2014, a new mammalian methyltransferase, METTL14, was discovered to catalyze m^6^A methylation. Together with METTL3, these two proteins form a stable heterodimer complex that mediates cellular m^6^A deposition on mammalian mRNAs (Liu et al. 2014; Wang et al. 2014b). Recently, the mammalian splicing factor WTAP (Wilms’ tumor 1-associating protein) was identified as the third auxiliary factor of the core methyltransferase complex that affects cellular m^6^A methylation (Liu et al. 2014; Ping et al. 2014). The identification and characterization of the complete mammalian m^6^A methylation machinery are the first steps toward deciphering the selectivity and biological functions of m^6^A deposition in eukaryotic mRNAs.

In this review, we mainly summarize recent progress in the study of m^6^A methylation in mRNA across different eukaryotes and discuss their newly discovered roles in post-transcriptional gene expression regulation. We first describe the features of m^6^A on a global scale and briefly introduce the mammalian m^6^A writers, erasers, and readers that specifically install, remove, or bind to m^6^A at defined sequence motifs [Fig. 1]. We then discuss the evolutionary conservation of the m^6^A methylation machinery across eukaryotic species that range from yeast, plants, and flies to mammals, highlighting the broad roles of methyltransferases and m^6^A in regulating cell status and embryonic development. Finally, we discuss the emerging functions of m^6^A in several mechanisms of post-transcriptional gene expression regulation with a special focus on the effects of m^6^A on differentiation and reprogramming of stem cells.

**Features of m^6^A on a global scale**

Studies in the 1970s revealed that m^6^A modification in mRNA mainly occurs at Pu[G > A]m^6^AC[U > A > C] (Pu represents purine) and is estimated to be present at an average level of approximately three m^6^A residues per mRNA (Rottman et al. 1974; Narayan and Rottman 1988; Csepany et al. 1990; Narayan et al. 1994). Transcriptome-wide mapping of m^6^A is hindered by the following two facts: (1) m^6^A, akin to A, reverse-transcribes to a thymine (T), and (2) m^6^A is not susceptible to chemical modifications that might promote its detection. In 2012, two groups independently developed an antibody-based high-throughput sequencing method (Dominissini et al. 2012; Meyer et al. 2012) and for the first time profiled the transcriptome-wide m^6^A distribution. In each method, mammalian mRNA is properly fragmented and immunoprecipitated by an m^6^A-specific antibody. Libraries are prepared from immunoprecipitated and input control fragments, respectively, and subjected to high-throughput sequencing. In general, ∼12,000 m^6^A sites in the transcripts

![Figure 1. Illustration of the cellular pathways of m^6^A in nuclear RNAs. The m^6^A methyltransferases and demethylases dynamically control the m^6^A methylation landscape within the nucleus. The m^6^A reader proteins preferentially bind to the methylated RNA and mediate specific functions. In the nucleus, m^6^A may affect alternative splicing of pre-mRNA and mature mRNA storage and export. After mature RNAs are exported to the cytoplasm, cytoplasmic m^6^A reader YTHDF2 can bind to the m^6^A-containing mRNAs to mediate mRNA decay. Other cytoplasmic readers could modulate mRNA translation and storage.](https://genesdev.cshlp.org/article/Fig1.jpg)
of ~7000 coding genes and ~300 noncoding ones are identified in human cells. The resolution of the m^6A peak site is ~100 nucleotides [nt], which was further improved by later optimization (Schwartz et al. 2013, 2014b; Chen et al. 2015a). However, transcriptome-wide m^6A detection at single-base resolution remains a challenge.

To date, m^6A RNA methylomes across many eukaryotes, including human (Dominissini et al. 2012; Meyer et al. 2012; Batista et al. 2014; Schwartz et al. 2014b), mouse (Dominissini et al. 2012; Meyer et al. 2012; Batista et al. 2014; Schwartz et al. 2014b; Wang et al. 2014b; Geula et al. 2015), yeast (Schwartz et al. 2013), and plant (Li et al. 2014; Luo et al. 2014), have been profiled. In general, global mapping reveals a conserved, widespread, and dynamic mRNA methylation in eukaryotes. Three salient features of the m^6A methylome are evident. First, m^6A sites are mainly confined to the consensus motif Pu[G > A] m^6AC[U > A > C], which is consistent with earlier studies. Second, m^6A marks are not equally distributed across the transcriptome; they are preferentially enriched in a subset of consensus sequences near stop codons, in 3′ UTRs, and within long internal exons [Fig. 2]. In particular, this topology is preserved upon endodermal differentiation of stem cells (Batista et al. 2014; Geula et al. 2015). Last, m^6A-modified genes are well conserved between human and mouse embryonic stem cells (ESCs) and somatic cells (Batista et al. 2014). For instance, ~70% of human ESC genes are also m^6A-modified in the orthologous mouse gene, with ~46% of the m^6A peak sites in common. As expected, higher m^6A peak intensities were detected in conserved sites compared with those that are not conserved. On the other hand, distinct m^6A patterns can also be detected among different species or cells residing in different developmental stages [Meyer et al. 2012; Schwartz et al. 2013; Batista et al. 2014; Geula et al. 2015]. Certain m^6A modifications are tissue-specific and dynamically alter in response to different stimuli, indicating the potential role of m^6A in regulating diverse cellular processes.

**m^6A writers in mammals**

The m^6A modification is installed by a multicomponent methyltransferase complex (Fig. 1), which has not been fully characterized. In a pioneer work reported in 1997 [Bokar et al. 1997], a 200-kDa methyltransferase complex was isolated from the HeLa nuclear extract, which exhibits methyltransferase activity. Only a 70-kDa protein, termed MT-A70 or METTL3, was identified as one SAM-binding unit within this 200-kDa methyltransferase complex. The knockdown of METTL3 led to apoptosis of human HeLa cells and a concomitant reduction in cellular m^6A level [Bokar 2005]. METTL3 and m^6A appear to be strongly associated with development and gametogenesis, since the depletion of the METTL3 homologs in yeast (Agarwala et al. 2012), flies (Hongay and Orr-Weaver 2011), and plants (Zhong et al. 2008; Bodi et al. 2012) readily lead to developmental arrest or defects in gametogenesis.

A phylogenetic analysis of the MT-A70 [METTL3] family methyltransferase has suggested METTL14, which shares 43% identity with METTL3 but belongs to a different lineage, as a homolog of METTL3 [Fig. 3; Bujnicki et al. 2002]. The highly conserved nature of METTL14 in mammals together with the fact that the METTL14 protein can be pulled down by METTL3 has prompted researchers to consider METTL14 as a putative candidate for m^6A deposition on mRNA [Liu et al. 2014]. Intriguingly, knockdown of METTL14 results in a more pronounced decrease of m^6A in polyadenylated RNA compared with knockdown of METTL3 in both HeLa and human 293 FT cell lines [Liu et al. 2014]. The recombinant METTL3 and METTL14 proteins can form a stable METTL3–METTL14 complex in the gel filtration experiment, and subsequent two-dimensional native/SDS-PAGE analysis has further demonstrated the formation of a heterodimer between these two proteins, with a stoichiometry of 1/1 [Liu et al. 2014]. While the METTL14 protein itself exhibits higher methylation activity compared with METTL3 in vitro, the combination of both methyltransferases substantially enhances methylation efficiency, demonstrating a synergistic effect that is further confirmed by in vivo studies [Liu et al. 2014; Wang et al. 2014b]. The METTL3–METTL14 heterodimer preferentially methylates RNA substrates containing the previously identified consensus sequence GGACU and exhibits a modest preference for the less structured RNA substrate in vitro. Furthermore, the methyltransferase complex was isolated from the native HeLa cell nuclear extract. The nuclear extract fraction that exhibits the highest methylation activity was found to be mostly enriched with METTL3 and METTL14 [Liu et al. 2014], thus clearly indicating that the heterodimer of METTL3–METTL14 forms the catalytic core of the mammalian m^6A methyltransferase complex.

![Figure 2](https://genesdev.cshlp.org/figure2.png)

**Figure 2.** The normalized distribution (density) of m^6A peaks along the mRNA transcripts in HeLa cells (top panel) and Arabidopsis thaliana (bottom panel), where each mRNA transcript is divided into the 5′ UTR, coding sequences (CDS), and the 3′ UTR.
WTAP has been identified as the third crucial component of the mammalian m^6^A methyltransferase complex (Liu et al. 2014, Ping et al. 2014). WTAP was initially shown to act as a splicing factor that binds to Wilms’ tumor 1 protein (Little et al. 2000) and plays a regulatory role in cell cycle progression and early embryo development (Horiuchi et al. 2006, 2013). The first evidence of WTAP as a third component of the methyltransferase complex came from the coimmunoprecipitation result, which showed that WTAP readily binds to the METTL3–METTL14 heterodimer inside cells, although the interactions between WTAP and the two methyltransferases are weaker compared with that between METTL3 and METTL14 (Liu et al. 2014). WTAP itself does not possess methylation activity, consistent with its lack of a conserved catalytic methylation domain, but interacts with the METTL3–METTL14 heterodimer to substantially affect cellular m^6^A deposition (Liu et al. 2014; Schwartz et al. 2014b). A subsequent study suggests that WTAP helps to coordinate the localization of the METTL3–METTL14 heterodimer into nuclear speckles, thereby facilitating m^6^A deposition (Ping et al. 2014).

Global analysis indicates that METTL3, METTL14, and WTAP share a large portion of common binding sites (∼36%) on their RNA substrates and exhibit a binding consensus motif similar, if not identical, to that of m^6^A (Liu et al. 2014). A PAR-CLIP (photoactivatable ribonucleoside-enhanced cross-linking and immunoprecipitation) assay revealed that a large fraction of the binding sites fall into intergenic regions (∼46%) and introns (∼31%). This observation suggests that the core methyltransferase complex might work on precursor mRNAs (pre-mRNAs); however, whether and how m^6^A is installed is not yet known (Fig. 1). The m^6^A mark may play a regulatory role in alternative splicing pathways because alternative splicing can be directly affected by the presence of the m^6^A modification in the spliced region (Fig. 1; Geula et al. 2015). In addition, silencing of the methyltransferase complex leads to enhanced abundance of their m^6^A target transcripts, supporting the role of m^6^A as a negative regulator of gene expression (Batista et al. 2014; Liu et al. 2014; Schwartz et al. 2014b; Wang et al. 2014a,b; Geula et al. 2015).

The discovery of the core mammalian m^6^A methyltransferase complex comprised of METTL3, METTL14, and WTAP reveals several new insights. It is surprising and interesting that the core complex of the mRNA m^6^A methyltransferase contains two parallel active methyltransferases. Each is active and seems to impact different sets of transcripts. One potential explanation points to the selective regulation of different pathways and functions inside cells. Each component may be subjected to different post-translational modifications or binding of partner proteins for the tuning of specific functions.
m^6^A erasers in mammals

FTO is the first identified demethylase that oxidatively reverses m^6^A to adenosine in mRNA (Jia et al. 2011). FTO is a member of the ALKB subfamily of Fe^III/^alpha-ketoglutarate-dependent dioxygenases, which has eight other family members in humans (ALKBH1–ALKBH8) and catalyzes the oxidation of diverse biological substrates (Kuworski et al. 2003; Gerken et al. 2007; Fu et al. 2010; Zheng et al. 2014). FTO was initially thought to work on 3-methylthymidine (3mT) in ssDNA (Gerken et al. 2007) and 3-methyluracil (3mU) in ssRNA (Jia et al. 2008). In 2011, FTO was discovered to efficiently demethylate m^6^A in nuclear RNA (Jia et al. 2011). A subsequent study showed that FTO can oxidize m^6^A to two previously unknown intermediates—N^6^-hydroxymethyladenosine (hm^6^A) and N^6^-formyladenosine (f^6^A)—in a stepwise manner (Fu et al. 2013). Intriguingly, this process is similar to the oxidation of 5-methylcytosine (5mC) in genomic DNA to 5-hydroxymethylcytosine (5hmC) and then 5-formylcytosine (5fC) by the TET (ten eleven translocation) family proteins (Tahiliani et al. 2009; Ito et al. 2010, 2011), which also belong to the general family of Fe^III/^alpha-ketoglutarate-dependent dioxygenases. TET proteins can further oxidize 5fC to 5-carboxycytosine (5caC) (He et al. 2011; Ito et al. 2011; Zhang et al. 2012). While 5hmC, 5fC, and 5caC are stable cytosine derivatives, hm^6^A and f^6^A are short-lived intermediates with half-lives of ∼3 h in aqueous solution under physiological conditions (Fu et al. 2013). The continuous oxidation of 5hmC by the TET family proteins is a critical step in the active DNA demethylation pathway in mammals (He et al. 2011; Pastor et al. 2013; Shen et al. 2014). It is not yet clear whether hm^6^A and f^6^A have specific biological functions.

Immunostaining revealed that the FTO protein mainly resides in the nucleus and partially colocalizes with nuclear speckles (Jia et al. 2011), suggesting a dynamic model of m^6^A demethylation on mRNA coupled with m^6^A deposition and RNA processing. A recent study found that FTO can modulate alternative splicing of the important adipogenic factor RUNX1T1 by removing the m^6^A residues around the splice sites (Zhao et al. 2014). It is proposed that loss of m^6^A on RUNX1T1 transcripts prevents the binding of the splicing factor SRSF2 protein and promotes the production of a shorter isoform, which in turn acts to induce preadipocyte differentiation. FTO is also found in the cytoplasm in several cell types, suggesting a possible role of FTO in modulating cytosolic mRNA processing (Gulati et al. 2013; Vujovic et al. 2013).

Shortly after the discovery of FTO, ALKBH5 was identified and characterized as a second mammalian m^6^A demethylase that displays distinct biological functions (Zheng et al. 2013). Like FTO, ALKBH5 preferentially binds ssRNAs due to the presence of a unique loop in ALKBH5 that confers single-stranded substrate selectivity (Aik et al. 2014; Xu et al. 2014a). Distinct from FTO, though, ALKBH5 directly reverses m^6^A to adenosine with no detected intermediates. ALKBH5 is primarily colocalized with nuclear speckles and affects mRNA export and RNA metabolism in a demethylation-dependent manner (Zheng et al. 2013). ALKBH5 knockout mice exhibit impaired male fertility, consistent with the highest expression level of ALKBH5 being in the testis (Zheng et al. 2013). In contrast, FTO is most highly expressed in mouse brains, and FTO-deficient mice mainly suffer from early mortality and reduced body mass (Gerken et al. 2007; Fisher et al. 2009). Taken together, the diverse functions regulated by these two demethylases suggest broad physiological roles of m^6^A.

Further research is needed to delineate the mechanisms by which demethylases act on specific mRNAs and IncRNAs. Advanced sequencing methods coupled with global analysis approaches will help to define the demethylomes of FTO and ALKBH5.

m^6^A readers in mammals

While the transcriptome-wide RNA m^6^A landscape is sculpted by methyltransferases and demethylases in a dynamic and reversible manner, proteins that preferentially recognize m^6^A (termed m^6^A readers) bind to methylated RNA and confer downstream functions. Studies using methylated RNA probes to pull down binding proteins followed by mass spectrometry identification have identified several m^6^A reader candidates in mammalian cells (Dominissini et al. 2012). Among them, the YTH domain-containing family proteins (YTHDF1–3) were validated as m^6^A readers in cytoplasm, with binding affinities to methylated RNA ranging from ∼180 nM to ∼520 nM (Wang et al. 2014a). Subsequently, YTHDC1, another member of the YTH domain family, was identified as a mammalian m^6^A reader in the nucleus (Xu et al. 2014b). Mrg1 (methylated RNA-binding 1), a yeast protein with an YTH domain, was also shown to be an m^6^A reader (Schwartz et al. 2013). Crystal structure characterizations of the YTH domain containing a bound m^6^A further reveal a conserved hydrophobic pocket used for the binding of the methyl group of m^6^A as well as the preferential
Conservation of \( m^6A \) RNA methylation machinery and its related biological functions across eukaryotes

The identification and characterization of the \( m^6A \) methylation machineries are the first steps toward elucidating the biological roles of \( m^6A \) in mRNAs. Phylogenetic analysis revealed that the MT-A70 [METTL3] superfamily consists of four lineages of proteins with varied degrees of interrelatedness (Bujnicki et al. 2002). The simplification and updated version is shown in Figure 3. Lineages A, B, and C are unique to eukaryotes, while lineage D corresponds to a small group of bacterial DNA \( m^6A \) methyltransferases associated with restriction/modification systems. Among eukaryotes, humans, mice, pufferfish, *Drosophila melanogaster*, and *Arabidopsis thaliana* each contain representatives of the A, B, and C lineages. For instance, humans have representative proteins METTL3, METTL14, and METTL4 that belong to the A–C subfamily, respectively. The budding yeast *Saccharomyces cerevisiae* specifies IME4 (inducer of meiosis 4) and KAR4 (karyogamy protein) in the A and B lineages, respectively, while the fission yeast *Schizosaccharomyces pombe* seems to have only one member in lineage C. Conservation of the methylation signature motifs such as DPPW and EPPL (Fig. 3) in the MT-A70 superfamily members suggests a common ancestry. Genetic studies of methyltransferases in different organisms have been performed in order to understand functional roles of \( m^6A \) methylation on mRNA (Table 1). Below we focus on reviewing methyltransferases in different organisms and their associated biological functions.

**\( m^6A \) methylation machinery in yeast: the MIS [MUM2 (muddled meiosis 2)–IME4–SLZ1 (sporulation-specific leucine zipper 1)] complex mediates \( m^6A \) RNA deposition during yeast meiosis**

Unlike mammals, \( m^6A \) methylation in yeast *S. cerevisiae* is confined to meiosis; \( m^6A \) starts to accumulate on mRNA at the onset of meiosis, peaks in premeiotic S and G2/prophase, and decreases as strains enter into the meiotic divisions. In fact, the modification is hardly detected in yeast undergoing mitotic growth (Clancy et al. 2002; Bodi et al. 2010; Agarwala et al. 2012). High-resolution mapping of \( m^6A \) sites in meiotic yeast transcripts reveals that the methylation sites are primarily enriched in a consensus motif—RGAC (R = A/G), similar to the consensus motif in mammals—and are strongly biased toward the 3′ end of the transcripts (Schwartz et al. 2013). IME4 (yeast homolog of mammalian METTL3) is identified as an essential component for \( m^6A \) deposition on yeast mRNA and regulates meiotic progression via RNA methylation. Depletion of IME4 in yeast is not lethal but delays cellular entry into meiosis divisions and hinders sporulation (Shah and Clancy 1992; Hongay et al. 2006; Agarwala et al. 2012). A two-hybrid screen in yeast has identified a core \( m^6A \) RNA methyltransferase complex (termed MIS) composed of IME4, MUM2 (yeast homolog of mammalian WTAP), and a third crucial component, SLZ1 [not conserved in mammals] (Table 1; Agarwala et al. 2012). Intriguingly, each component of the MIS complex is expressed in a meiosis-specific manner, consistent with meiosis-confined methylation (Agarwala et al. 2012; Schwartz et al. 2013). Upon the induction of meiosis, SLZ1 shuttles IME4 and MUM2 from the cytoplasm into the nucleolus. Notably, nucleolar entry of the MIS complex is essential for \( m^6A \) deposition on yeast mRNA, and the global \( m^6A \) level subsequently reaches its maximum at meiotic prophase. After that, down-regulation of \( m^6A \) deposition is induced by activation of NDT80, a transcription factor required for exit from meiotic G2/prophase (Chu and Herskowitz 1998). As a result, the MIS complex exits from the nucleolus, and \( m^6A \) abundance returns to the basal level as cells enter into the meiotic divisions. Interestingly, researchers have found that IME4 also regulates IME1, which implies a putative positive feedback loop between \( m^6A \) deposition and IME1 expression (Schwartz et al. 2013).
m\(^6\)A methylation in D. melanogaster

*D. melanogaster* IME4 shows significant amino acid similarity to and a conserved catalytic domain with its eukaryotic homologs (Table 1; Fig. 3). Unlike in yeast, elimination of the full-length *D. melanogaster* IME4 in *Drosophila* is lethal (Hongay and Orr-Weaver 2011). Partial deletion of *D. melanogaster* IME4 is semilethal, with the rare viable adults showing significantly reduced fecundity. The catalytic domain of *D. melanogaster* IME4 is required for the rescue of this semilethality (Hongay and Orr-Weaver 2011), indicating a potential role for m\(^6\)A RNA methylation in metazoan development. Further studies showed that *D. melanogaster* IME4 was primarily expressed in the gonads of adult flies. In females, *D. melanogaster* ime4-deficient females exhibit compound egg chambers accompanied by significant defects in the Notch signaling pathway. The ancillary factor FL(2)D (female-lethal 2 D), the homolog of yeast MUM2 and mammalian WTAP, is conserved in *Drosophila*. This protein is required for the splicing regulation of *Sxl* (Sex lethal) and *tra* (transformer) pre-mRNAs, two critical gene transcripts associated with *Drosophila* sex determination and dosage compensation (Penalva et al. 2000; Ortega et al. 2003; Penn et al. 2008).

### Table 1. Evolutionary conservation of nuclear RNA m\(^6\)A methylation machinery

<table>
<thead>
<tr>
<th>Species</th>
<th>Methyltransferases</th>
<th>Auxiliary factors</th>
<th>Biological roles</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Saccharomyces cerevisiae</em></td>
<td>IME4</td>
<td>MUM2, SLZ1</td>
<td>Required for meiosis and sporulation (Clancy et al. 2002), SLZ1 localizes the complex to the nucleolus for m(^6)A methylation (Schwartz et al. 2013).</td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em></td>
<td>IME4</td>
<td>FL(2)D</td>
<td>IME4 is essential for viability (Hongay and Orr-Weaver 2011). IME4 is required for Notch signaling during oogenesis (Hongay and Orr-Weaver 2011). FL(2)D is required for splicing of <em>Sxl</em> and <em>tra</em> pre-mRNAs that are responsible for sexual determination (Penalva et al. 2000).</td>
</tr>
<tr>
<td><em>Danio rerio</em></td>
<td>METTL3, METTL14</td>
<td>WTAP</td>
<td>METTL3 and WTAP are required for normal embryogenesis (Ping et al. 2014).</td>
</tr>
</tbody>
</table>

m\(^6\)A methylation in plants

m\(^6\)A is a ubiquitous modification found in the mRNAs of various plants, including monocot plants maize (Nichols 1979), wheat (Kennedy and Lane 1979), oat (Haugland and Cline 1980), *A. thaliana* (Zhong et al. 2008, Luo et al. 2014), and rice (Li et al. 2014). MTA (encoded by At4g10760), a METTL3 ortholog in *Arabidopsis*, has been identified as an active component of the m\(^6\)A methyltransferase complex (Zhong et al. 2008). MTA interacts with FIP37 (encoded by At3g54170), an *Arabidopsis* homolog of mammalian WTAP and *Drosophila* FL(2)D, highlighting the highly conserved nature of the methyltransferase components across eukaryotes (Table 1). Intriguingly, MTA tends to be expressed in higher levels in dividing tissues, such as developing seeds, shoot meristems, and emerging lateral roots (Craigon et al. 2004; Zhong et al. 2008). Disruption of either MTA or FIP37 in *Arabidopsis* leads to developmental arrest of embryos at the globular stage (Vespa et al. 2004; Zhong et al. 2008), coupled to a loss of m\(^6\)A from the mRNA in arrested seeds (Vespa et al. 2004; Zhong et al. 2008). Later in development, perturbation of MTA causes multiple growth defects, including reduced apical dominance, organ abnormality, and increased trichome branching (Bodi et al. 2012). Collectively, these results demonstrate that the methyltransferase and hence m\(^6\)A methylation in mRNA play a crucial role in plant development. Very recently, transcriptome-wide m\(^6\)A profiling was performed in two accessions of *Arabidopsis* (Luo et al. 2014)—Can-0 and Hen-16—as well as in the rice callus and leaf (Li et al. 2014). It is worth noting that *Arabidopsis* and rice are unique in their enrichment of m\(^6\)A not only around the stop codon and within 3′ UTRs—as observed in yeast and mammals—but also around the start codon (Fig. 2). As genes possessing m\(^6\)A sites around the start codon are associated with photosynthesis and appear to be highly expressed in *Arabidopsis*, this suggests a potential direct role of m\(^6\)A at the 5′ UTR during translation (Luo et al. 2014). It will be interesting to determine whether this feature observed in plants is conserved in other organisms such as mammals.
m⁶A methylation machinery in vertebrates and mammals

We previously discussed the m⁶A methylation machinery of mammals in our description of writer proteins. The core m⁶A methyltransferase complex METTL3–METTL14–WTAP is highly conserved from zebrafish to mammals. In zebrafish, both METTL3 and WTAP proteins are ubiquitously expressed during embryogenesis and specifically enriched in the brain 36 h after fertilization (Ping et al. 2014). Embryos injected with either METTL3 or WTAP antisense morpholinos [MOs] suffer from various developmental defects, including smaller heads, eyes, and brain ventricles and curved notochord. In comparison with embryos injected with single-gene-targeted MOs, simultaneous knockdown of these two genes leads to a more pronounced phenotype in embryonic development as well as more severe decreases in the m⁶A level, indicating the in vivo synergistic effect of the methyltransferase complex. How METTL14 affects m⁶A deposition and zebrafish tissue differentiation remains to be studied.

Methyltransferases METTL3 and METTL14 are also shown to mediate the m⁶A formation in mouse ESCs (mESCs) (Batista et al. 2014, Wang et al. 2014b; Geula et al. 2015). Recent work has identified m⁶A as a crucial regulator in the differentiation and reprogramming of stem cells, which are discussed next.

Biological consequences of m⁶A methylation of mRNA and the underlying mechanisms

m⁶A RNA methylation determines stem cell fate by regulating pluripotency transition toward differentiation

ESCs are pluripotent stem cells derived from the inner cell mass (ICM) of a preimplantation embryo, exhibiting prolonged undifferentiated proliferation and stable developmental potential to form derivatives of all three embryonic germ layers [Thomson et al. 1998]. The ESCs reside in a so-called “ naïve” pluripotent state, while epiblast stem cells (EpiSC) that are derived from a post-implantation epiblast reside in a more differentiation-prepared, “primed” pluripotent state (Geula et al. 2015). The transition from naïve pluripotency to differentiation is tightly regulated by a plethora of pluripotency markers and developmental factors. Transcriptome-wide m⁶A profiling in mESCs and human ESCs showed that the majority of these core pluripotent genes (e.g., Nanog, Sox2, Klf4, Myc, Jarid2, and Smad3) and developmental regulators (e.g., Foxa2 and Sox17) have m⁶A modifications on their transcripts, with most of them being targets of Mettl3 (Batista et al. 2014; Wang et al. 2014b; Geula et al. 2015). Meanwhile, siRNA screening also identified Mettl3 as an epigenetic repressor that specifically destabilizes the primed EpiSCs (Geula et al. 2015). Importantly, both of the two methyltransferases, Mettl3 and Mettl14, are shown to catalyze m⁶A RNA deposition in mESCs (Batista et al. 2014; Wang et al. 2014b; Geula et al. 2015). Wang et al. (2014b) reported that the partial depletion of Mettl3 or Mettl14 by shRNAs leads to decreased m⁶A levels and reduced self-renewal of mESCs. However, in more recent studies (Batista et al. 2014; Geula et al. 2015) complete Mettl3 knockout mESCs and epiblasts were generated that actually displayed increased self-renewal but substantially impaired differentiation into mature cardiomyocytes and neurons [Batista et al. 2014]. When subcutaneously injected into immunodeficient mice, Mettl3 knockout mESCs readily generate larger but poorly differentiated teratomas in vivo, further indicating that depletion of m⁶A in mESCs enhances self-renewal but hampers differentiation (Batista et al. 2014).

Recently, Geula et al. (2015) demonstrated that the m⁶A modification plays a key role in facilitating transition of mESCs from the naïve state to the primed state upon differentiation (Fig. 4). To resolve the role of m⁶A in the naïve pluripotent state, genetic ablation of Mettl3 was performed in mESCs, and mating the Mettl3+/− heterozygote mice yielded the Mettl3−/− knockout blastocysts. Consistent with previous results of Batista et al. (2014), Mettl3-depleted mESCs showed an almost complete loss of m⁶A and preserved naïve pluripotency but failed to proceed into the primed EpiSC-like state. Like Mettl3−/− mESCs, Mettl14−/− knockout mESCs resisted progression out of the naïve state. Taken together, this evidence suggests that m⁶A ablation in naïve mESCs impairs the transition of naïve mESCs into the primed state and hence blocks the subsequent differentiation. In contrast, mouse EpiSCs [mEpiSCs] at a primed pluripotency state showed a distinct response to m⁶A depletion, namely, Mettl3 knockdown in mEpiSCs resulted in attenuated stability and an enhanced tendency to lineage priming, which finally led to fast differentiation and/or cell death.

The balance between naïve pluripotency and lineage priming is fine-tuned by the relative expression of naïve pluripotency markers and lineage commitment factors. Global analysis of methylomes of naïve ESCs and primed EBs showed that m⁶A modification was detected in 80% of the transcripts of naïve pluripotency genes (e.g., Nanog, Klf4, Sox2, and Esrrb) as well as multiple lineage commitment regulators (e.g., Foxa2 and Sox17). In general, m⁶A deposition in mESCs decreases the expression of methylated transcripts and directly reduces their stability. For both types of regulators, loss of m⁶A results in increased abundance of transcripts and longer mRNA lifetime (Batista et al. 2014; Wang et al. 2014b; Geula et al. 2015), reminiscent of the role of YTHDF2 in mediating the degradation of methylated mRNA [Wang et al. 2014a]. Thus, depletion of m⁶A acts to boost the expression of the dominant regulators [pluripotent-promoting or lineage commitment genes] at a given pluripotency state, thereby driving stem cell differentiation. In the ground naïve state, where pluripotency-promoting transcripts prevail, Mettl3 depletion further amplifies the already highly expressed naïve pluripotency genes but leads to only a marginal increase in lineage commitment transcripts, resulting in a so-called “hyper naïve” pluripotency phenotype (Fig. 4; Batista et al. 2014; Geula et al. 2015). In the primed state, where lineage commitment transcripts dominate, Mettl3 depletion primarily up-regulates lineage commitment
m6A significantly increases the frequency of two types of m6A modification in the spliced region. Depletion of Mettl3 significantly enhances the reprogramming efficiency (Warren et al. 2010). Inhibition of m6A formation by silencing METTL3 causes an mRNA processing delay and circadian period

Functions of m6A modification in eukaryotic mRNA

Figure 4. Methyltransferases set m6A marks on mRNAs to balance the expression levels of pluripotency genes and lineage commitment genes in naïve and primed states of the ESCs. In the naïve state, the expression level of the pluripotency genes is dominant over that of lineage commitment genes, while in the primed state, the trend exhibits the opposite. The m6A methyltransferase depletion in naïve pluripotent cells further up-regulates already highly abundant naïve pluripotency genes, while the lineage commitment genes remain at very low residual levels. As a result, cells stay in a “hypernaïve” pluripotent state and fail to progress into the primed state. If the methyltransferase depletion occurs in the primed state, the expression level of the differentiation priming markers is further boosted, which pushes cells above the critical threshold toward differentiation, leading to fast differentiation and/or cell death.

m6A RNA methylation controls the circadian clock

The mechanism of the mammalian circadian clock involves a negative transcription-translation feedback loop in which the transcription of the clock genes is suppressed by their own encoded proteins. The period of the circadian cycle is set according to this general principle. Around 10% of the transcriptome in livers is known to be rhythmic, but only about one-fifth is driven by de novo transcription, which indicates that mRNA processing could serve as a major circadian component. Recent work showed that many clock genes as well as clock output gene transcripts bear m6A modifications [Fustin et al. 2013]. Inhibition of m6A formation by silencing METTL3 causes an mRNA processing delay and circadian period.
It appears that m^6^A depletion prolongs nuclear retention of mature mRNAs of the clock genes Per2 and Arntl. This result reveals an important physiological function of m^6^A methylation in setting the pace of the circadian cycle and determining clock speed and stability.

**Perspectives**

The last few years have witnessed breakthrough discoveries on biological functions of m^6^A in mRNA, but the field is still in its infancy. Methylation specificity stands out as one of several challenging questions that remain to be addressed. In mammals, m^6^A occurs in only ~15% of all methylation consensus Pu[GU] > A[m^6^A]AC[A/C/U] motifs, and these methylated sites are primarily enriched near the stop codon, at the 3′ UTR, within long exons, and at the 5′ UTR. How the methylation machinery selectively targets a subset of consensus motifs in the transcriptome remains to be understood. This specificity likely has functional implications on the methylated RNAs. The METTL3–METTL14 heterodimer exhibits higher activity to the GGACU sequence located in a random structure region compared with that residing in the stem or loop [Liu et al. 2014]. In agreement with the biochemistry results, global analysis also shows that methylated sites are significantly less structured when compared with randomly selected counterparts from the same genes, possibly because these sites are more exposed and accessible to the methylation machinery [Schwartz et al. 2013]. However, more complicated pathways/mechanisms must be involved to achieve target selectivity. A recent study indicated that microRNAs [miRNAs] could partially regulate m^6^A modification via a sequence-pairing mechanism [Chen et al. 2015b], whereby miRNA expression may modulate the binding of METTL3 to mRNA substrates. Further biochemical and cellular validations are required to confirm this model. Interestingly, another recent study revealed that the m^6^A mark on primary miRNA (pri-miRNA) plays critical roles in miRNA maturation [Alarcon et al. 2015]. METTL3 methylates pri-miRNAs, which facilitates their recognition and processing by the RNA-binding protein DGCR8 in the initiation of miRNA biogenesis. Collectively, these studies suggest a potential regulatory network between the miRNA-based regulation and the m^6^A-dependent regulation as two main pathways that post-transcriptionally control gene expression [Alarcon et al. 2015; Berulava et al. 2015; Chen et al. 2015b].

The multicomponent mammalian methyltransferase complex still needs to be completely resolved because auxiliary components in the complex may play roles in recruiting the catalytic core to the particular locations of the cognate pre-mRNAs and/or tuning activities of the methyltransferases. Thus, careful characterizations of proteins that interact with METTL3/METTL14/WTAP within the nuclear speckles will shed further insights on the origination of the m^6^A specificity.

Transcriptome-wide mapping of m^6^A at single-base resolution will greatly facilitate our understanding of selective m^6^A installation by the methyltransferase complex. With a base-resolution m^6^A map, single and clustered m^6^A sites can be differentiated from each other; m^6^A fractions on particular transcripts and nearby cis elements can be derived. Additionally, one can study the knockout cell lines to determine whether METTL3 and METTL14 control individual groups of transcripts or share the same targets. Most m^6^A-seq studies to date have profiled the steady-state polyadenylated RNA inside cells, with the majority of them being mature mRNA rather than highly labile pre-mRNA. Therefore, it is necessary to carry out m^6^A sequencing on pre-mRNA in order to thoroughly examine the prevalence and distribution of m^6^A within the intronic regions and estimate the percentage of mRNAs that could be methylated either cotranscriptionally or, potentially, post-transcriptionally.

Emerging results suggest that m^6^A serves as a dynamic mark on a large number of mRNAs and IncRNAs, which help cells rapidly respond and/or adapt to external signaling and stimuli. By virtue of the reversible nature of the m^6^A modification, the stability, localization, and translatability of a large group of mRNA transcripts and IncRNAs can be regulated by m^6^A reader proteins and thereby participate in a timely manner in various biological pathways. The methyltransferases, demethylases, and reader proteins can all direct the methylation-based signaling process. Development of small molecule inhibitors or gene therapy tools for targeting these proteins could lead to new ways of controlling gene expression and potential new therapies for human diseases.

Last, m^6^A in eukaryotic mRNA exhibits substantial contributions to post-transcriptional gene expression regulation. This same modification, N^6^-methyladenine (m^6^A or m^6^dA), in DNA has been known to play important roles in bacterial genomes. Very recently, three independent studies reported the presence and characterizations of m^6^A/m^6^dA in three different eukaryotic genomes (green alga, worm, and fly) with proposed transcription regulation functions [Fu et al. 2015; Greer et al. 2015; Zhang et al. 2015]. Indeed, the adenine methylation appears to be a common mechanism to control gene expression.

**Acknowledgments**

This work was supported by the National Institutes of Health (GM071440 to C.H.). C.H. is an investigator of the Howard Hughes Medical Institute. S.F. Reichard contributed editing.

**References**


Functions of m^A modification in eukaryotic mRNA

GENES & DEVELOPMENT 1353
Yue et al.


