Implications of Widespread Covalent Modification of mRNA

Thomas A. Cooper

Topology of the Human and Mouse m6A RNA Methylomes Revealed by m6A-seq
Dominissini et al

Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3' UTRs and Near Stop Codons
Meyer et al

Whole transcriptome analysis reveals that one third of human and mouse genes express mRNAs containing methylated adenosines. Fat mass and obesity-associated (FTO) gene, a risk gene for obesity and metabolic abnormalities, demethylates the modified residues, suggesting a link between RNA methylation and disease.

Vertebrate ribosomal RNA (rRNA), transfer RNA (tRNA), messenger RNAs (mRNAs) as well as viral RNAs can be methylated at the N6 position of adenosines to produce N6-methyladenosines (m6A). Although the presence of m6A modifications in mRNAs has been known for decades and a handful of vertebrate m6A-modified mRNAs have been identified, the extent of m6A modification in mRNAs was unknown. Two recent articles using massively parallel sequencing found that mRNAs from more than one third of human and mouse genes contain m6A.1,2 N6-methyladenosines were found to be enriched near the translational stop codon, as well as in mRNA segments derived from large exons. Furthermore, the results showed that the genes modified and the distribution of modifications within mRNAs is conserved between mouse and humans. Particularly intriguing is the recent discovery that FTO gene, a risk gene for obesity and metabolic abnormalities, is an m6A demethylase. These results suggest a potential regulatory role for covalent modification of RNA and a link with metabolic disorders.

Over the past 2 decades, RNA has been revealed as one of the most functionally diverse and information-rich molecules in the cell.3 RNA can function as a viral genome, message, decoding machine, guide to target proteins, enzyme, template and primer for DNA synthesis, and structural component. RNA carries information as primary nucleotide sequence and by using a huge structural space to provide a fit for diverse molecular interactions. Pre-mRNAs transcribed by RNA polymerase II are extensively modified and processed into mature mRNA by addition of the 5' cap, splicing, and endonucleolytic cleave to form the mRNA 3' end, which is further modified by addition of ≈250 nontemplated adenosines to produce the poly(A) tail. In addition, diverse species of RNA (tRNA, mRNA, small nuclear RNA, small nucleolar RNA, long intergenic non-coding RNAs etc) contain >100 chemical modifications.4 As for proteins, posttranscriptional modifications create the potential to increase and modulate functionality either through altering RNA–RNA or protein–RNA interactions.

The availability of antibodies that specifically recognize the m6A modification combined with the transcriptome-wide analysis provided by RNA-Seq in an approach analogous to chromatin immunoprecipitation-Seq (called MeRIP-Seq; Figure) has revealed several unexpected results. First, mRNA from up to one third of mouse protein–coding genes and >300 noncoding RNAs contain m6A modifications. Almost half of these mRNAs contain only a single m6A modification, whereas one third contain 2, one tenth contain 3, and 1 in 20 contain ≥4. Quantitative analysis demonstrated that the majority of mRNAs from individual genes are modified, supporting a physiological role for the modification.1 A second surprise was the topology of m6A modifications within mRNAs. Both studies found striking enrichment of the modification near the translational stop codon. The methylated adenosine sits within a preferred consensus sequence that is relatively short. The fact that only one tenth of the putative sequence motifs are methylated demonstrates the high level of selectivity, implicating tight regulation by a yet-to-be defined mechanism. Also highly suggestive of functionality is the finding that the m6A modifications are conserved between mouse and humans both in terms of the mRNAs modified and the topology of modifications. All in all, the results provide strong indications of functional consequences.

Now the question is what are the functions of the m6A modifications? There are 2 levels of inquiry relevant to function that arise from these data. First, what are the molecular functions of the m6A modification? The m6A residue does not affect coding potential and is not expected to affect base-pairing interactions. However, the modification is likely to have either positive or negative effects on interactions between RNA and protein. How does this potential modulate the functions of individual pre-mRNAs and mRNAs? The second level of inquiry is what is the larger physiological relevance of methylated mRNAs? That is, what is the role for m6A modification in cell, tissue, and organismal physiology?

With regard to molecular functions, the evidence from Dominissini et al1 associates m6A modification with pre-mRNA splicing in 2 ways: first is a potential role in alternative splicing. A subunit of the methyltransferase that is required for the m6A modification, METTL3 was knocked down in
the HepG2 cell line. Interestingly, METTL3 knockdown was associated with a change in the ratios of mRNA splice variants that correlated with exons and introns that contain m6A modifications. Furthermore, m6A modifications were enriched among exons and introns that are alternatively spliced. A second link between m6A and splicing is the finding that modified residues are highly enriched within unusually long internal exons (>400 nt). The average size of an internal exon is ≈140 nt, and internal exons larger than ≈300 nt are spliced less efficiently as a result of constraints within the splicing machinery, which remain to be defined. In addition to the splice sites at the exon–intron junctions that are essential for splicing, vertebrate pre-mRNAs contain a splicing code of short (5–8 nt) motifs within introns and exons that prevent recognition of cryptic splice sites and enhance recognition of bona fide splice sites by the splicing machinery. Large exons are thought to require additional information for their efficient splicing. The results of Dominissini et al. suggest that m6A could contribute to the splicing code to ensure efficient splicing of large exons. The preferential modification of mRNAs near the translation termination signal suggests a role for methylation either in translation regulation or in an aspect of RNA metabolism associated with the 3′ untranslated region. For example, Meyer et al. found that m6A modifications are enriched within mRNAs that contain microRNA (miRNA)-binding sites which are typically found within the 3′ untranslated region. Although modified mRNAs are more likely to contain miRNA-binding sites, there was an inverse correlation between the locations of miRNA-binding sites and m6A modifications such that the modification was located 5′ of the miRNA-binding sites. It is possible that mRNAs influence methylation sites or that methylation affects miRNA binding or function. It is also possible that this correlation reflects an indirect relationship reflective of a yet-to-be-discovered phenomena.

Beyond defining the specific molecular functions of m6A modification of individual mRNAs, results from both articles address the broader biological significance of m6A modifications. One question addressed is the dynamic nature of the modification: are changes in methylated adenosines used to regulate gene expression programs? The extent of m6A modification increases dramatically during postnatal brain development, suggesting functionality during this critical transition. Comparison between adult mouse brain and the human HEK293 cell line showed extensive similarities by MeRIP-Seq illustrating striking conservation, but also suggesting the lack of cell-specific regulation on a global scale. Similarly, MeRIP-Seq on HepG2 cells subjected to a variety of stresses showed only minimal differences in m6A sites. Although the results do not explain why a large fraction of mRNAs are modified, they provide a limited number of genes subject to altered methylation, with potential relevance to a physiological consequence; something that is often not available after a genome-wide assay.

A particularly intriguing recent finding is that the FTO gene encodes an m6A demethylase, suggesting an important link between m6A modifications and metabolism, obesity, and heart disease. Several genome-wide association studies have linked FTO with obesity and metabolic abnormalities. The next task is to connect the molecular function of FTO in m6A demethylation to a physiological role in metabolism. One can imagine a role for FTO in setting a critical baseline modification of the transcriptome or in the dynamic modification of a few specific mRNAs. The FTO demethylase is localized in the same nuclear structures as the methyltransferase (using the METTL3 subunit as a proxy). One can also imagine that, in the end, the mA6 signature results from coordinated methylation and trimming by the FTO demethylase.

Although m6A is thought of as a posttranscriptional modification, the actual timing of modification with regard to transcription is not clear. mRNA processing (capping, splicing, cleavage/polyadenylation) is primarily cotranscriptional and because m6A modifications were found within introns as well as exons, it is likely that a large fraction of the target pre-mRNAs is cotranscriptionally methylated. Consistent with this is the finding that METTL3 and FTO are located in nuclear speckles that also contain components of the splicing machinery. Although speckles are thought to be storage areas rather than sites of active splicing, colocalization suggests a potential splicing-associated function. It is also likely that the fraction of modified sites within introns is an underestimate, because the procedures used for MeRIP-Seq were biased toward the more abundant spliced mRNAs. An analysis of nuclear RNA for m6A modifications would produce an interesting map of m6A enriched within pre-mRNA. It would also be interesting to use MeRIP-Seq in a time course after acute gene activation and in RNA from different cellular components as recently described to define the dynamics of modifications within different RNA populations.

The results from these studies overwhelmingly support an important role for the m6A modification in one or more aspects of gene regulation. The results also illustrate the bursts of insight...
that result from systematic analysis using genome-wide approaches. Ultimately however, it will be an analysis of individual genes that will connect the effects of m6A modifications within specific mRNAs to a broader understanding of physiological significance. These results have turned our collective heads to look in a new direction. Now, it will require the tools and creativity of molecular, cellular, and computational biologists to zero in on the specific targets of m6A modification that are relevant to each of the areas of potential physiological impact.

Disclosures

None.

References


Implications of Widespread Covalent Modification of mRNA
Thomas A. Cooper

Circ Res. 2012;111:1491-1493
doi: 10.1161/CIRCRESAHA.112.281071

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/111/12/1491