ESSAY

GENOMICS AND PROTEOMICS

Roadmap to the epitranscriptome

N^6-methyl-adenosine signals the way and YTH proteins respond

By **Dan Dominissini**

The "RNA world" is not at all hypothetical but rather the biological world we live in (*1*). For RNA to function within a modern cellular milieu of proteins and DNA, numerous chemical modifications coevolved that help sculpt its interactions (*2*). To date, well over a hundred nucleotide modifications have been identified in diverse types of RNA molecules (*3*). Every position of pyrimidine and purine rings can be posttranscriptionally modified, with methylation predominating. Although the importance of some modifications—especially those on structured, stable, and catalytically active RNAs (like ribosomal and transfer RNA)—is appreciated, their mode of action is still largely unknown.

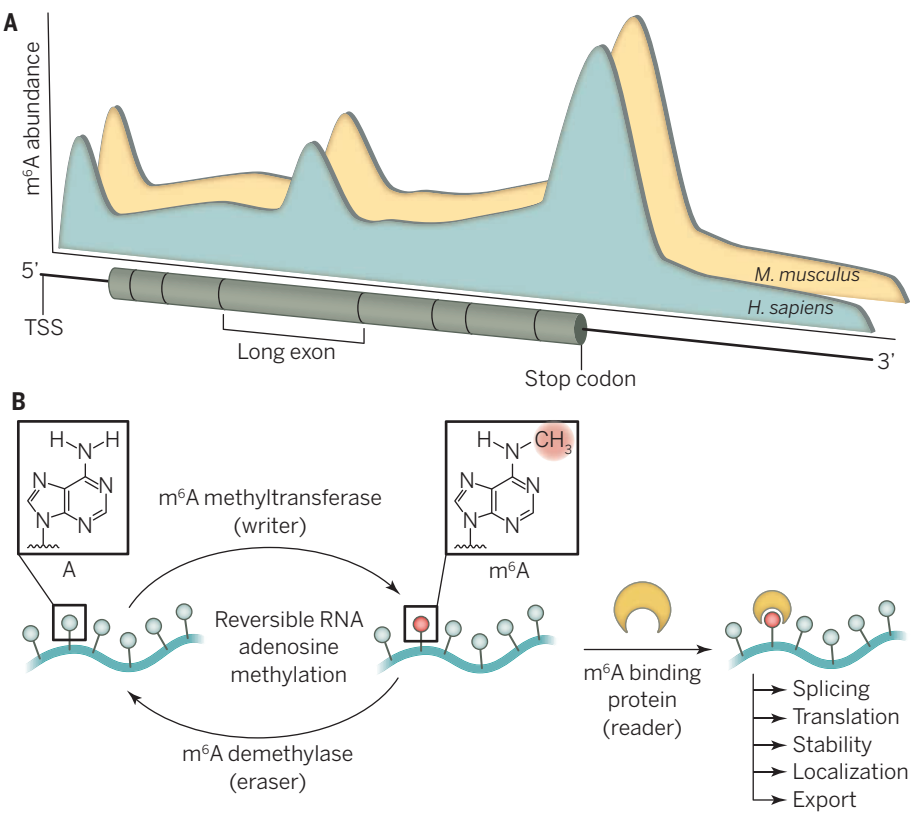
Recognizing the complex lives that mRNAs lead, involving highly interdependent processing events, one wonders whether messages convey additional information to that specified in their sequence. Whereas dynamic and reversible chemical modifications decorating DNA and proteins are an integral part of our understanding, the players worth mentioning when thinking of mRNA have primarily been proteins.

Methylation of adenosine on the N^6 position (m^6A) was first observed 40 years ago (*4*) and was recognized early as the most prevalent internal modification in eukaryotic mRNA. However, its function remained enigmatic, leaving us in the dark, partly because of a lack of experimental approaches. Fitful efforts over the years established that cellular life and organismal development depend on adenosine methylation (*5*–*7*) but have identified only a handful of methylated positions (*8*), and, consequently, further functional investigations were limited.

As a Ph.D. student studying adenosine-to-inosine RNA editing in the laboratory of Gideon Rechavi, I became intrigued by the possibility of adenosine N^6-methylation as a mechanism to regulate RNA editing, either directly through inhibiting deamination itself or indirectly through destabilizing the double helical RNA structure required for editing enzymes to bind.

Addressing this hypothesis, I was in need of a map. The localization-function relation, analogous to structure-function, lies at the heart of genomics, and fundamental mechanistic discoveries in epigenetics were jump-started by global maps. As m^6A has little effect on Watson-Crick base pairing, it cannot be detected by reverse transcription. The advent of massively parallel sequencing enabled me to harness its powers to those of more traditional immunocapturing to develop a method for transcriptome-wide mapping of m^6A in high resolution, termed m^6A-seq (*9*, *10*). In essence, a highly m^6A-specific antibody is used to immunoprecipitate methylated RNA fragments out of a randomly fragmented transcriptome; fragments are then deep sequenced to identify positions of signal enrichment.

As it turned out, m^6A did not have much to do with editing, but as it happened, the landscape that unfolded proved even more dramatic. Applied to the human transcriptome, m^6A-seq revealed more than 12,000 methylated sites in transcripts originating from approximately 7000 coding genes and 250 noncoding ones. After years in the dark, we were instantly facing a wide vista. "Is there a pattern?" was the first question that came



Posttranscriptional regulation of RNA by reversible adenosine methylation. Human and mouse mRNA transcripts are punctuated by m^6A at specific, highly conserved, discrete locations: around stop codons, within long internal exons, and at transcription start sites (**A**). Methylation, a dynamic modification, is installed by a nuclear methyltransferase complex (writers)—composed of METTL3, METTL14, and WTAP—and removed by at least 2 demethylases (erasers), ALKBH5 and FTO. Methyl-specific binding proteins (readers), primarily of the YTH-domain family, bind to modified transcripts and mediate the effect (**B**).

Department of Chemistry and Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL 60637, USA. E-mail: dandominissini@gmail.com

Category Winner

Genomics and Proteomics: Dan Dominissini for his essay, "Roadmap to the epitranscriptome." Dr. Dominissini received his Bachelor of Medical Science degree from Tel-Aviv University, Israel, in 2007. He went on to study RNA posttranscriptional modifications for his Ph.D., focusing on adenosine deamination and methylation, with Gideon Rechavi at Tel-Aviv University. He is currently a Human Frontier Science Program postdoctoral fellow in the laboratory of Chuan He at The University of Chicago, where he develops novel chemistries for the study of nucleic acid modifications.

For the full text of all winning essays and further information, see http://scim.ag/SciLifeLab.

to mind. And indeed, methylation was nonrandomly distributed, which created a telltale pattern: It was strongly enriched around stop codons, within unusually long internal exons and at transcription start sites (TSSs) [see the figure (A)]. Although sites largely conformed to a strong sequence consensus, it is not sufficient to explain selective m^{6}A distribution. In the mouse transcriptome, m^{6}A-seq revealed a remarkable degree of conservation with humans in terms of topology (the way methylated sites are arranged relative to key positions in a typical RNA molecule), sequence consensus, and orthologous modified transcripts [see figure (A)]. The evolutionary conservation of this unique pattern provided the most telling evidence supporting function.

What can we make of the identity of methylated transcripts? Most expressed genes were methylated, so gene ontology terms were not very useful. A seeming lack of clear "functional enrichment" can be perplexing. However, if m^{6}A were as fundamental and far-reaching as we hypothesized, this is not surprising. In these cases, more can be learned from exceptions to the rule. We therefore asked which genes were least methylated. The most highly expressed ones was the answer. This observation was the first to draw an inverse link between methylation and transcript abundance and suggested that highly abundant messages avoid methylation to maximize stability.

The 2011 discovery that the fat mass and obesity–associated protein (FTO) was an m^{6}A demethylase (*11*) prompted us to look into dynamics. Dynamics is the combined result of active demethylation and shifting substoichiometric modification rates. Overall, the topology of m^{6}A across surveyed conditions was static. However, despite blind spots in our method, we detected a subset of sites that changed in response to physiological conditions and thus uncovered the dynamic nature of this modification.

How does m^{6}A exert its influence? Studies of methyl groups on amino acids in histones and cytosines in DNA (5mC) have shown that they confer considerable affinity and specificity to protein binding (*12*). Using synthetic RNA baits, I identified several specific m^{6}A-binding proteins, the top ones belonging to the YTH domain family (*13*). Finding the first m^{6}A-binding proteins, like the 5mC-binding proteins at the time, represents a breakthrough in our understanding and ability to investigate the function of m^{6}A. Subsequent studies provided a structural basis for selective m^{6}A binding (*14*) and showed that the destabilizing effect of m^{6}A is mediated by binding to a family member, YTHDF2 (*15*). Additional functions are expected to emerge as more binding proteins are uncovered and characterized.

Can the conserved topology of m^{6}A illuminate its function? Admittedly, we are still unsure. Being invariably unstructured (*16*), the stop codon vicinity can be a good "landing pad" for proteins, and translational control comes naturally to mind. Methylation at internal exons is suggestive of splicing. Indeed, knockdown of the METTL3 subunit of the m^{6}A methyltransferase resulted in widespread alterations in splicing that correlated well with methylation, consistent with the location of methylases and demethylases in nuclear speckles along with known pre-mRNA splicing factors. The siting of m^{6}A at the TSS belongs to a different story and represents installation of m^{6}A on the first transcribed nucleotide (part of the 5′ cap structure) by a different enzyme. Thus, the unbiased power of m^{6}A-seq to uncover novel phenomena, regardless of location or sequence consensus, is underscored.

When we take a wider view, evidence gathered in recent years is consolidating our understanding that posttranscriptional regulation contributes as much and probably more than the better-characterized transcriptional regulation to determine gene expression (*17*). On the basis of its ubiquity, conservation, dynamics, and phenotypes, m^{6}A appears to make an essential contribution to posttranscriptional processes orchestrated by specific binding proteins and perhaps, as initially hypothesized, directly through RNA restructuring. Clearly, many questions remain.

The roadmap that my collaborators and I created has spawned a spate of studies, some refining the physiological processes linked to m^{6}A (including stem cell differentiation and circadian periods) and others delving into its molecular underpinnings. The discoveries of components in the methyltransferase holocomplex (METTL14 and WTAP) (*18*), of demethylating enzymes (FTO and ALKBH5) (*11*, *19*), and of m^{6}A-binding proteins can be combined with our detailed m^{6}A maps to make head-way in delineating pathways for posttranscriptional regulation of mRNA by reversible methylation, analogous to the reversible processes of DNA methylation and protein phosphorylation [see figure (B)]. ■

REFERENCES AND NOTES

1. T. R. Cech, *Cold Spring Harb. Perspect. Biol.* **4**, a006742 (2012).
2. H. Grosjean, Ed., *Fine-Tuning of RNA Functions by Modification and Editing*, vol. 12 of *Topics on Current Genetics* (Springer, Berlin, 2005).
3. W. A. Cantara *et al.*, *Nucleic Acids Res.* **39** (Database), D195–D201 (2011).
4. R. Desrosiers, K. Friderici, F. Rottman, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 3971–3975 (1974).
5. S. Zhong *et al.*, *Plant Cell* **20**, 1278–1288 (2008).
6. M. J. Clancy, M. E. Shambaugh, C. S. Timpte, J. A. Bokar, *Nucleic Acids Res.* **30**, 4509–4518 (2002).
7. C. F. Hongay, T. L. Orr-Weaver, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 14855–14860 (2011).
8. S. Horowitz, A. Horowitz, T. W. Nilsen, T. W. Munns, F. M. Rottman, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5667–5671 (1984).
9. D. Dominissini *et al.*, *Nature* **485**, 201–206 (2012).
10. D. Dominissini, S. Moshitch-Moshkovitz, M. Salmon-Divon, N. Amariglio, G. Rechavi, *Nat. Protoc.* **8**, 176–189 (2013).
11. G. Jia *et al.*, *Nat. Chem. Biol.* **7**, 885–887 (2011).
12. H. Li *et al.*, *Mol. Cell* **28**, 677–691 (2007).
13. Z. Zhang *et al.*, *J. Biol. Chem.* **285**, 14701–14710 (2010).
14. C. Xu *et al.*, *Nat. Chem. Biol.* **10**, 927–929 (2014).
15. X. Wang *et al.*, *Nature* **505**, 117–120 (2014).
16. Y. Wan *et al.*, *Nature* **505**, 706–709 (2014).
17. C. Vogel, E. M. Marcotte, *Nat. Rev. Genet.* **13**, 227–232 (2012).
18. J. Liu *et al.*, *Nat. Chem. Biol.* **10**, 93–95 (2014).
19. G. Zheng *et al.*, *Mol. Cell* **49**, 18–29 (2013).

ACKNOWLEDGMENTS

D.D. is supported by a Human Frontier Science Program long-term fellowship.

10.1126/science.aaa1807

PHOTO: DAN DOMINISSINI

Roadmap to the epitranscriptome

Dan Dominissini

Science **346** (6214), 1192.
DOI: 10.1126/science.aaa1807