RNA Technologies 12

Stefan Jurga
Jan Barciszewski *Editors*

Epitranscriptomics

Springer

RNA Technologies

Volume 12

RNA Technologies provides critical and comprehensive discussions of the most significant areas of RNA research, written by leading international authorities. Each volume in the series reviews both general topics, such as RNA interference, RNA biosynthesis and metabolism, microRNAs and diseases, transcriptome analysis and aptamers, as well as highly topical chapters on very recent findings in the field. Recurring themes of the series include the complexities associated with target recognition, RNA-protein interactions, analysis of genetic networks and the promise of small RNAs as cancer diagnostics and therapeutics in combating cancer.

RNA-based approaches have great potential to revolutionize molecular biology, cell biology, biomedical research and medicine. The nucleic acid-based molecules can be used to regulate the level of gene expression inside the target cells and their potential efficacy against several viruses, including SARS-CoV-2, and other chronic diseases represents a promising research area.

RNA Technologies publishes in both print and electronic format since 2010 and is intended for researchers from academia and industry, as well as graduates that look for a carefully selected collection of high quality review articles on their respective field of expertise.

More information about this series at http://www.springer.com/series/8619

Stefan Jurga • Jan Barciszewski
Editors

Epitranscriptomics

Springer

Editors
Stefan Jurga
Nanobiomedical Center
Adam Mickiewicz University
Poznań, Poland

Jan Barciszewski
Nanobiomedical Center of the Adam
Mickiewicz University, Poznań and
Institute of Bioorganic Chemistry of the Polish
Academy of Sciences,
Poznań, Poland

ISSN 2197-9731 ISSN 2197-9758 (electronic)
RNA Technologies
ISBN 978-3-030-71611-0 ISBN 978-3-030-71612-7 (eBook)
https://doi.org/10.1007/978-3-030-71612-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Introduction: Understanding Epitranscriptomics

The genetic alphabet consists of the four letters: C, A, G, and T in DNA and C, A, G, and U in RNA. 61 of 64 triplets of these four letters jointly encode 22 different amino acids which construct proteins. This system is universal and functions in all kingdoms of life.

Comparative transcriptomics between mammals has revealed that ∼66% of human genomic DNA is transcribed. RNA plays a critical role in regulating cellular functions. Interestingly, only ∼2% of the transcriptional production is protein-coding messenger RNA (mRNA), while ∼98% encompasses a wide variety of noncoding RNA (ncRNA) molecules. ncRNAs have been classified functionally as either housekeeping or regulatory. The housekeeping ncRNA genes include ribosomal RNA (rRNA), transfer RNA (tRNA), and small nuclear RNA (snRNA), while examples of regulatory ncRNAs are microRNA (miRNA) and long noncoding RNA (lncRNA).

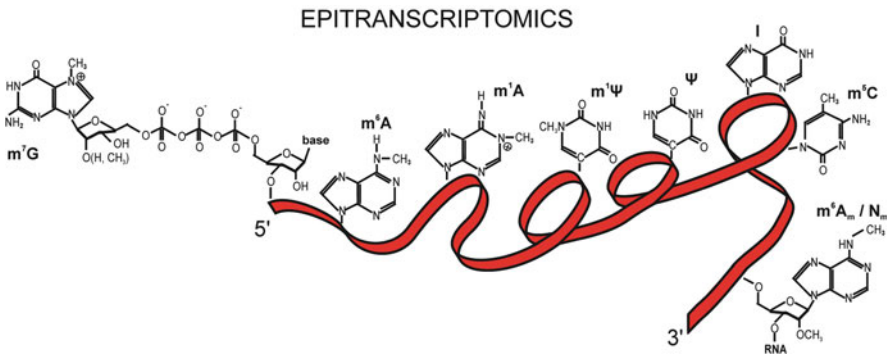
The year 2021 marks the diamond jubilee of the discovery of messenger RNA (mRNA) and the cracking of the genetic code. Sixty years ago, Sydney Brenner, François Jacob, and James Watson isolated mRNA; however, François Jacob and Jacques Monod put mRNA into a theoretical context, arguing for its role in gene expression and regulation. These observations stimulated a new way of thinking about gene function. At the same time, J. Heinrich Matthaei and Marshall Nirenberg showed that poly U encodes the synthesis of polyphenylalanine.

The complexity of RNA is further complicated by numerous post-transcriptional modifications. RNA harbors the potential of being dynamically and reversibly regulated by the addition and removal of distinct chemical moieties. Over 160 post-transcriptional modifications expand the RNA code. These modifications extend the RNA repertoire and alter its chemistry in various ways without changing the nucleotide sequence. The majority of RNA modifications consist of adding a methyl group to certain positions on the nucleobase. RNA methylation can be reversible and convey information *via* recognition of effector proteins. The different positions for methylation carry distinct implications. Although mRNA modification

has been known since the 1970s, its functional importance in mRNA metabolism and its effect on human biology have not been extensively studied in the past.

The process of mRNA maturation involving 5′-capping, splicing, and polyadenylation is well studied. The post-transcriptional modifications found in mRNA mark regions that potentially contribute to the regulation of cellular processes, including gene expression, protein translation, or RNA stability. These modifications can alter the structure and metabolism of mRNA. Only recently methodological and conceptual advances allowed systematic mapping and functional analysis to unfold the role they play in mRNA biology.

Mammalian messenger RNA contains tens of thousands of modified nucleotides, an essential addition to the standard genetic code of four nucleotides in animals, plants, and their viruses. Technological advances that allowed mapping of selected RNA modifications on a transcriptome-wide scale revealed the widespread distribution of N6-methyladenosine (m^6A), pseudouridine (Ψ), ribose 2′-O-methylation (Nm), N1-methyladenine (m^1A), and 5-methylcytidine (m^5C) on mRNA (Figure).



m^6A destabilizes pairing with uracil by altering the energetics of an AU pair through steric hindrance, but the pattern of hydrogen-bonding donors and acceptors remains the same. m^1A methylation endows a positive charge, which may result in strong electrostatic interactions. Also, in m^1A, the methyl group protrudes from the Watson–Crick hydrogen-bonding face of adenine, resulting in the nucleotide remaining unpaired. 2'-O-methylation augments hydrophobicity, protects against nucleolytic attack, and stabilizes RNA helices. Some modifications are incorporated enzymatically by various methyltransferases (writers) and removable by demethylases (erasers).

In comparison to U, Ψ contains an extra imino group (>C═NH), which serves as an additional hydrogen bond donor, while the carbon–carbon (C─C) glycosidic bond linking the sugar to the base is more stable than the carbon-nitrogen (C─N) in U. These two chemical changes confer rigidity to the sugar-phosphate backbone and enhances local base stacking. An additional class of nucleotide modifications, termed RNA editing, creates an irreversible change in the nucleotide sequence. These modifications include insertions, deletions, and base substitutions and occur

in all classes of RNA. When they occur in mRNA, the amino acid sequence of the protein will be altered relative to the sequence encoded by genomic DNA. RNA editing by deamination results in adenosine (A) to inosine (I) and cytosine (C) to uridine (U). A-to-I editing is an abundant class of RNA modifications found throughout metazoans. The conversion of A-to-I residues by base deamination results in the synthesis of distinct proteins, which creates functional diversity and enhances response to rapid environmental changes. RNA editing by deamination is mediated by two major classes of enzymes; the first class is a group of tissue-specific and context-dependent adenosine deaminases called ADARs. The ADAR enzyme class (adenosine deaminases acting on RNA) catalyzes hydrolytic deamination of A-to-I in double-stranded regions of RNA secondary structure. The second class of enzymes, the vertebrate-specific apolipoprotein B mRNA editing catalytic polypeptide-like (APOBEC) family, promotes C-to-U editing by cytosine deamination. APOBEC1, the first-discovered member of the APOBEC family, was characterized as the zinc-dependent cytidine deaminase which catalyzed a C-to-U modification, resulting in an in-frame stop codon in APOBEC mRNA.

Epitranscriptomics, or RNA epigenetics, is a branch of epigenetics and refers to RNA editing and noncoding RNA regulations. It is a young and fast developing field. mRNA modifications represent another layer of epigenetic regulation to gene expression. Epitranscriptomics is crucial in cell proliferation, migration, invasion, and epithelial–mesenchymal transition. Epitranscriptomics plays essential roles in alternative splicing, nuclear export, transcript stability, and translation of RNA targets to pioneer new ways of cancer treatment.

The rapid advancement of next-generation sequencing and mass spectrometry technologies has recently allowed for the identification and functional characterization of nucleotide modifications in protein-coding and noncoding RNA on a global transcriptome scale. In this book, there will be a summary of transcriptome-wide RNA modification mapping techniques. It will highlight studies exploring the functions of RNA modifications and their association with disease and finally offer insights into the future progression of epitranscriptomics.

Accurate regulation of the transcriptome is critical for gene expression and its subsequent control of cellular functions, including metabolism, proliferation, differentiation, and development. Thus, alterations in transcriptome regulation can disrupt cellular functions and lead to disease. Accumulating evidence has identified and functionally characterized several distinct types of chemical modifications of RNA nucleotides in both protein-coding RNAs and ncRNAs, further advancing the burgeoning field of epitranscriptomics.

The mechanisms and functions of different modifications in mRNA with an emphasis on its effect on human health and disease will be discussed. The low abundance of nucleotide modifications and technical limitations, however, have hampered systematic analysis of their occurrence and functions. Selective chemical and immunological identification of modified nucleotides has revealed global candidate topology maps for many modifications in mRNA.

There are many chapters in this book each dedicated to a particular topic. The role of RNA modifications which represent a novel layer of regulation of gene expression

and offer new possibilities to rapidly alter gene expression upon specific environmental changes will be analyzed. These chapters highlight the importance of epitranscriptomics in different diseases and drug resistance. They also provide new insights on potential therapeutic targets to reverse drug resistance. The field of epitranscriptomics is still in its infancy. We look forward to many exciting discoveries in the coming years.

Poznań, Poland

Stefan Jurga
Jan Barciszewski

Contents

The Emerging Neuroepitranscriptome

Andrew M. Shafik, Emily G. Allen, and Peng Jin

Contents

Abstract Chemical modifications on ribonucleic acid are prominent in the brain, are emerging as regulators of proper brain development, and have been implicated in neurological disorders. Similar to epigenetic modifications that regulate gene expression patterns, epitranscriptomic marks and their associated cellular machinery are involved in the expression of RNA that is required for developmental processes and maintenance of the nervous system. This work provides a detailed view on the current literature surrounding N6-methyladeonosine, N1-methyladeonsine, 5-methylcytidine, pseudouridine, and other RNA modifications in brain development and neurological disorders.

A. M. Shafik · E. G. Allen · P. Jin (✉)
Department of Human Genetics, School of Medicine, Emory University, Atlanta, GA, USA
e-mail: peng.jin@emory.edu

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_1

Keywords Epitranscriptomics · N6-methyladenosine · N1-methyladenosine · 5-methylcytidine · Pseudouridine · Neurodevelopment · Neurogenesis · Neurodegenerative disease

1 Introduction

Epigenetic modifications such as DNA methylation and histone modifications can regulate gene expression and have been implicated in many biological processes, particularly neurodevelopment. The mechanisms by which these modifications act can be dynamically regulated through "writer" and "eraser" enzymes that install or remove the modifications, respectively. Dysregulation of these mechanisms has been implicated in neurodegenerative, neurodevelopmental, and neurological disorders. For example, DNA methylation and histone modification-mediated gene regulation are crucial for neural cell differentiation (Feng et al. 2007; Hirabayashi and Gotoh 2010; Sun et al. 2011; Yao et al. 2016). However, the effect of RNA modifications in the regulation of biological processes has only just gathered momentum in the past few years. Furthermore, similar to epigenetic modifications, the reversibility and dynamic nature of some post-transcriptional RNA modifications allow them to quickly and tightly regulate various biological processes. Indeed, RNA modifications in the brain are emerging as a critical regulator of many different neuronal pathways, while deregulation of these RNA modifications may be associated with neurological disorders and developmental diseases, including tumorigenesis.

RNA modifications were first identified in abundant RNA species, such as ribosomal RNA and tRNA. However, the occurrence of RNA modifications has recently been documented in less abundant RNAs such as mRNA and long non-coding RNA. These modifications are known to influence many aspects of RNA processing, including splicing (Tang et al. 2018; Zhou et al. 2019), mRNA export (Yang et al. 2017), mRNA stability (Wang et al. 2014; Yang et al. 2019), mRNA translation (Liu et al. 2016a; Mao et al. 2019; Schumann et al. 2020; Wang et al. 2015), and miRNA processing (Alarcon et al. 2015; Song et al. 2020) (Fig. 1). Currently, over 170 chemical modifications have been documented on all four canonical bases and the ribose sugar across all RNA types (Boccaletto et al. 2018). These modifications comprise the "epitranscriptome." Furthermore, the identification of proteins that install ("writers"), remove ("eraser"), and specifically recognize ("readers") RNA modifications has extended our knowledge regarding the role and mechanism by which these modifications affect cellular, developmental, and disease processes (Jonkhout et al. 2017; Kadumuri and Janga 2018). However, research into many RNA modifications is in its infancy, but the advent of high-throughput approaches has paved the way to understand the functions and mechanisms of RNA modifications. In particular, N6-methyladenosine is an abundant modification that has been extensively studied and has been widely implicated as a critical regulator in higher brain processes (see, for example, Shafik et al. 2020). A picture

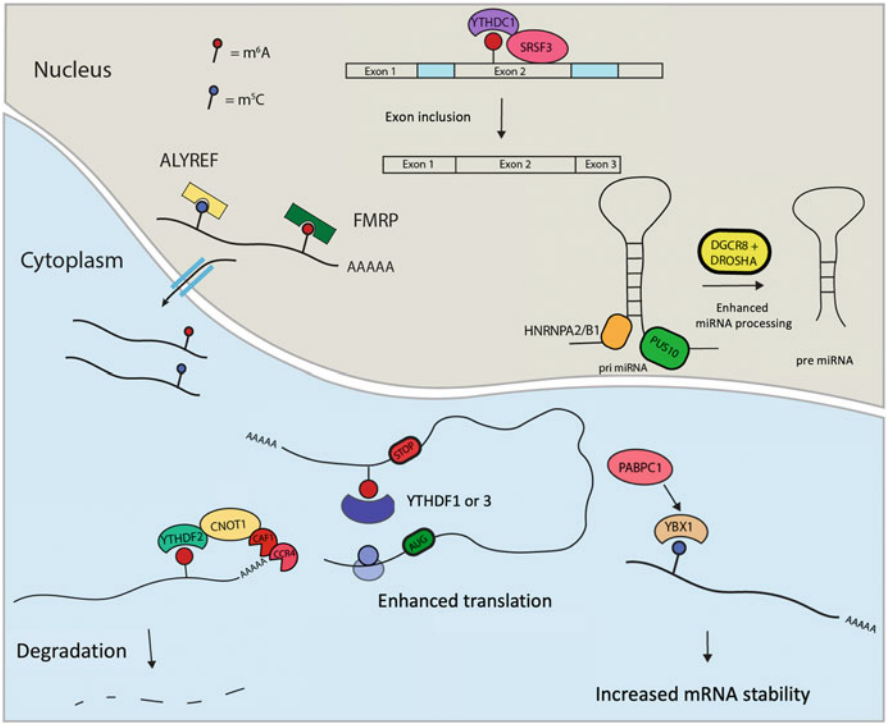


Fig. 1 RNA modifications affect many aspects of RNA metabolism. ALYREF and FMRP facilitate nuclear export of m^5C and m^6A marked transcripts, respectively. YTHDF1 and YTHDF3 specifically recognize m^6A marked transcripts, resulting in enhanced translation through a closed-loop model. YTHDC1 is an m^6A reader which directs the splicing factor SRSF3 to its target, resulting in exon inclusion. HNRNPA2/B1 or PUS10, in complex with DGCR8, recognize primary miRNAs modified by m^6A or pseudouridine respectively, delivering DGCR8 to its target for enhanced miRNA processing. YBX1 recognizes m^5C marked transcripts and functions in stabilizing the mRNA by recruiting PABPC1. m^5C = 5-methylcytidine, m^6A = N6-methyladenosine

of the neuroepitranscriptome is beginning to emerge, which solidifies the importance of RNA modifications in brain development and disease (Fig. 2).

2 N6-methyladenosine

N6-methyladenosine is one of the most abundant internal mRNA modifications and has been extensively studied over the last few years. m^6A has been shown to be a dynamic modification that is involved in the regulation of many biological processes, including mRNA stability, translation, splicing, export, and miRNA processing.

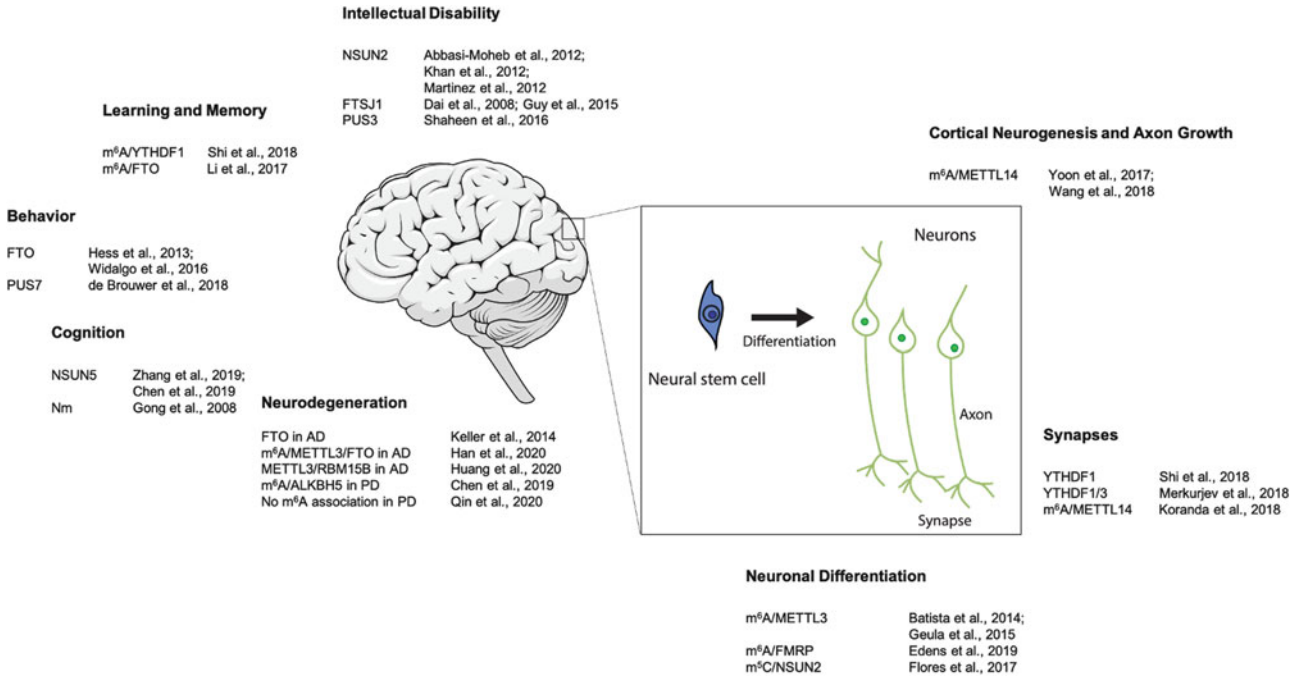


Fig. 2 A summary of the literature implicating RNA modifications and/or the machinery associated with those modifications in many neuronal processes. m^5C, m^6A, pseudouridine, and Nm modifications are broadly required during brain development and are involved in learning and memory, behavior, cognition, intellectual disability, neurodegeneration. Furthermore, m^5C and m^6A have a role in neuronal differentiation, while m^6A is also involved in cortical neurogenesis, axon growth, and synapses. m^5C = 5-methylcytidine, m^6A = N6-methyladenosine, Nm = 2′-*O*-methylation

m^6A is deposited on RNA targets by a complex that consists of the "writer" protein (methyltransferase-like 3) METTL3, which is supported structurally by (methyltransferase-like 14) METTL14 (Bokar et al. 1997; Liu et al. 2014). Further components of the complex include WT1-associated protein (WTAP), which functions in regulating the recruitment of the complex to its mRNA targets (Agarwala et al. 2012), while RNA-binding motif protein 15 directs the complex to appropriate m^6A sites (Patil et al. 2016). Lastly, Vir like m^6A methyltransferase associated (VIRMA) protein guides the catalytic core components (METTL3, METTL14, and WTAP) to selective methylation in the 3′UTR and near the stop codon (Yue et al. 2018). Methyltransferase-like 5 (METTL5) protein has also been shown recently to catalyze m^6A methylation at position 1832 in mouse 18S rRNA (Ignatova et al. 2020). Significantly, the authors demonstrate that *Mettl5* depletion in mouse embryonic stem cells resulted in a global reduction in mRNA translation and loss of pluripotency. Another recently discovered m^6A methyltransferase, methyltransferase-like 16 (METTL16), was shown to methylate a single mRNA (MAT2A) in its 3′ UTR, and this contributes to splicing induction (Pendleton et al. 2017). "Eraser" enzymes (FTO, ALKBH5) reverse the m^6A modification (Jia et al. 2011; Zhao et al. 2014; Zheng et al. 2013), while m^6A function is mediated by proteins that "read" it (e.g., YTHDF1-3, YTHDC1, and YTHDC2) (Mao et al. 2019; Shi et al. 2017; Wang et al. 2014, 2015). Following knockdown of FTO and ALKBH5, m^6A levels increase, validating these proteins as m^6A demethylases (Jia et al. 2011; Zhao et al. 2014; Zheng et al. 2013), while the affinity of "readers" is higher for m^6A methylated mRNA compared to unmethylated mRNA (Theler et al. 2014).

2.1 Detection of m^6A

Initially, m^6A was detected transcriptome-wide using a specific anti-m^6A antibody coupled to next generation sequencing. m^6A peaks are identified compared to a background RNA-seq input of the same sample in an approach termed methylated RNA immunoprecipitation sequencing (MeRIP-seq) (Dominissini et al. 2012; Meyer et al. 2012). Since then, another approach has been developed which allows for the single-nucleotide detection of m^6A (Linder et al. 2015). The authors cross-linked RNA to m^6A antibodies. Reverse transcription of the RNA resulted in a specific pattern of mutations or truncations, which allowed for the identification of m^6A at single-nucleotide resolution. More recently, antibody-free approaches have been developed. DART-seq or deamination adjacent to RNA modification targets coupled with next generation sequencing, employs a fusion protein consisting of cytidine deaminase, APOBEC1, and the m^6A-binding YTH domain (Meyer 2019). APOBEC1-YTH induces C-to-U deamination next to m^6A sites, which can then be identified using RNA-seq. Another antibody-free technique has been developed (Zhang et al. 2019b). They took advantage of the m^6A-sensitive endoribonuclease

(m^6A-sensitve RNA cleavage enzyme) to identify specific m^6A sites at single-nucleotide resolution (Fig. 3).

2.2 *m^6A in Neurodevelopment and Neurogenesis*

Specifically, m^6A is highly enriched in the mammalian brain, and many studies have implicated this RNA modification in neurodevelopment, neurogenesis, and neuropsychiatric disease. A decrease in protein expression of METTL3, METTL14, and WTAP during mouse cerebellar development from postnatal day 7 to postnatal day 60 has been detected (Ma et al. 2018). Furthermore, these proteins were detected in the external granule cell layer, Purkinje cell layer, and inner granule cell layer. With increasing age, the authors noted that expression levels decreased in the internal granular layer but increased in the Purkinje cell layer. Knockdown of *Mettl3* resulted in a dramatic alteration in Purkinje cell numbers, laminal structure, and stunted dendrites, suggesting that m^6A is crucial for proper cerebellar development. Loss of METTL3 has also been shown to result in the impaired embryonic stem cell exit from self-renewal toward differentiation into several lineages (Batista et al. 2014). On this note, Mettl14-induced m^6A depletion resulted in prolonged cell cycle of adult radial glial, while neuron production still occurred into postnatal stages, suggesting m^6A to be a major regulator of cortical neurogenesis (Yoon et al. 2017). Furthermore, Mettl14 conditional knockout results in a loss of late neurons, further suggesting a role for METTL14 in cortical neurogenesis (Wang et al. 2018). It was also demonstrated that METTL14 has a role in the striatum by (Koranda et al. 2018). The authors profiled m^6A in Mettl14-deleted striatum, where they showed a correlation between loss of m^6A and downregulation of mRNAs encoding neuron and synapse-specific proteins. More specifically, they found m^6A methylation in mRNAs known to function in synaptic plasticity, such as Homer1 and Cdk5r1, suggesting a role for m^6A in synaptic signaling. Also implicated WTAP has also been implicated in proper brain development (Ping et al. 2014). In *Drosophila,* they found that embryos injected with a *Wtap* morpholino exhibited smaller head, eyes, brain ventricle, and a curved notochord at 24 h post-fertilization compared to uninjected embryos. Interestingly, they found loss of *Mettl3* to only have a mild effect.

Ythdf1 KO mice have been shown to exhibit difficulties with learning and memory using hidden platform training Morris water maze tests (Shi et al. 2018). Furthermore, loss of Ythdf1 resulted in impaired basal synaptic transmission and long-term potentiation in hippocampal neurons, which further contributes to the defects in learning and memory exhibited by these mice. Re-expressing YTHDF1 in the hippocampus of the KO mice rescued defects in behavior and synapsis. Mechanistically, the authors found that in response to neuronal stimuli, YTHDF1, in the adult mouse hippocampus, recognizes a subset of m^6A marked transcripts, resulting in their enhanced translation and thus facilitating learning and memory. Another m^6A reader, YTHDF2, is highly expressed during the early stages of neural

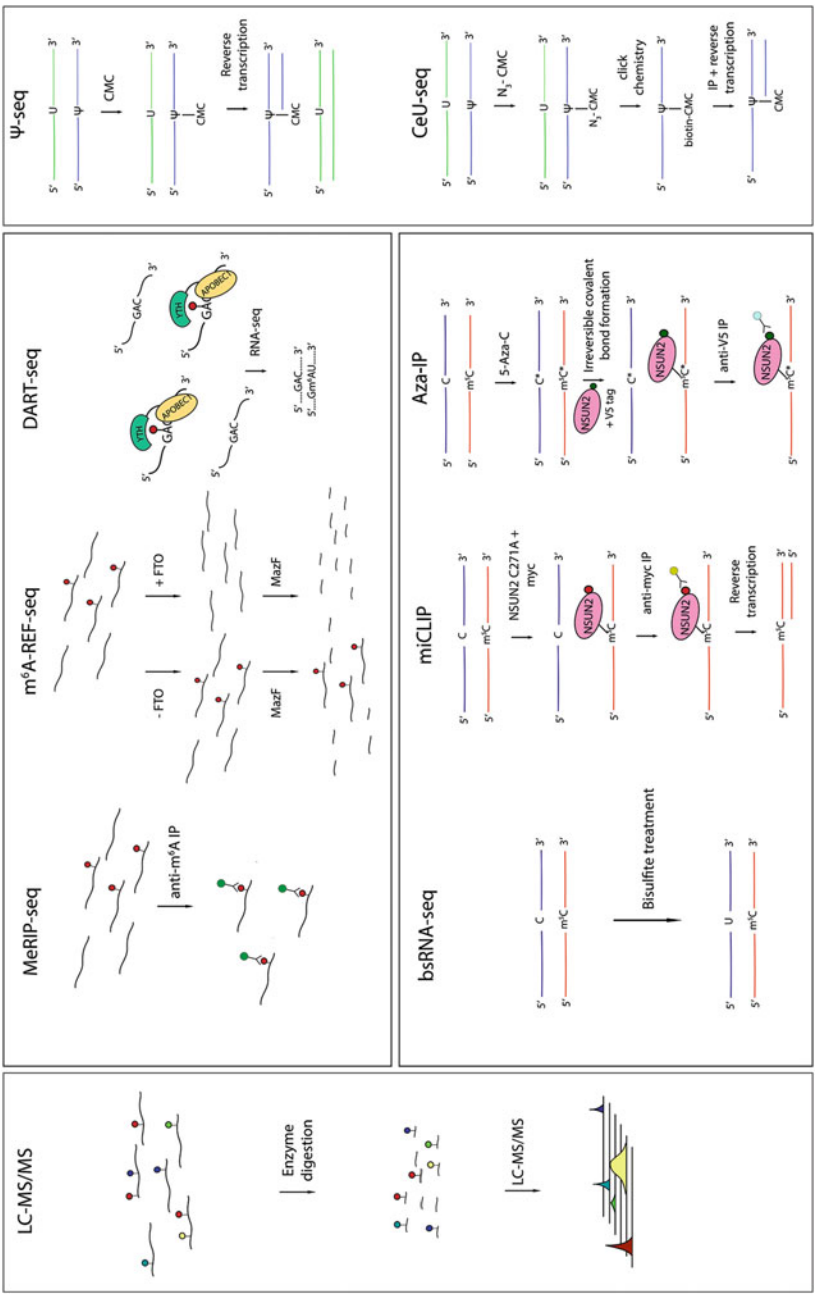


Fig. 3 Approaches to detect RNA modifications. LC-MS/MS allows for the identification of a plethora of RNA modifications in a sample but does not provide the user with information regarding the position of the modifications. The colored circles represent different types of RNA modifications. Each RNA modification produces a different signal. m^6A is detected using m^6A-REF-seq, MeRIP-seq, or DART-seq. m^6A-REF-seq and DART-seq are antibody-free methods. DART-seq uses an APOBEC1-YTH fusion protein to identify m^6A sites, while m^6A-REF-seq employs an m^6A-sensitive RNA cleavage enzyme

development and Ythdf2 −/− mouse embryos at E12.5 and E14.5 were alive but displayed reduced overall cortical thickness as described for Mettl14 cKO mice (Li et al. 2018b). They further demonstrate that loss of *Ythdf2* has a strong negative impact on neural stem/progenitor cell (NSPC) self-renewal and neuron generation in the embryonic neocortex. Also, YTHDF1, YTHDF2, and YTHDF3 are enriched in hippocampal dendrites, and the loss of YTHDF1 or YTHDF3 there resulted in altered spine morphology, dampened excitatory synaptic transmission, and altered cell-surface protein content (Merkurjev et al. 2018). Another protein, fragile X mental retardation protein (FMRP), was shown to bind m^6A marked mRNA (Edens et al. 2019; Zhang et al. 2018). It was demonstrated that upon loss of FMRP, 2035 genes were differentially expressed in the mouse cortex (Zhang et al. 2018). Interestingly, the majority of downregulated FMRP targets harbor an m^6A site. The authors show that FMRP binds the m^6A sites and interacts with YTHDF2, regulating the stability of its m^6A-marked targets. Furthermore, during neural differentiation, FMRP was shown to facilitate the nuclear export of m^6A marked transcripts (Edens et al. 2019). *Fmr1* KO mice were also observed to exhibit an extended neuronal progenitor cell cycle, with neural progenitors still proliferating into postnatal stages, similar to what was described in the Mettl14 cKO mice.

The correct balance of m^6A is essential for the proper development of the brain. Indeed, FTO-catalyzed m^6A demethylation has a significant role in neurogenesis, learning, and memory (Li et al. 2017a). They find FTO in both adult neural stem cells and neurons, observing an increased expression of FTO 1 day to 8 weeks post-birth. Furthermore, they find that many players in the brain-derived neurotrophic factor (BDNF) pathway are regulated by FTO, and this may, in turn, affect postnatal neurogenesis. Furthermore, it has been shown that the levels of FTO and ALKBH5 decreased in the cerebellum from P7 to P60 (Ma et al. 2018). Interestingly, they show that the cerebellum of Alkbh5-knockout mice is not significantly different in weight or morphology to wild-type mice. This may suggest that other demethylases, for example, FTO, may compensate for the lack of ALKBH5. However, since ALKBH5 may have a role in the regulation of cellular processes to hypoxia, the authors postulated that ALKBH5 may protect the brain against hypoxia.

Fig. 3 (continued) (MazF) to identify m^6A sites following treatment with or without the m^6A demethylase FTO. MeRIP-seq uses an antibody to specifically pull out m^6A-marked transcripts. All of these methods are coupled to next generation sequencing. m^5C is detected using either miCLIP, bsRNA-seq, or Aza-IP approaches. Both miCLIP and Aza-IP traps the RNA methyltransferase (NSUN2) to its target. Antibodies against a tagged version of NSUN2 are then used to immunoprecipitate the protein along with the covalently bound RNA, followed by next generation sequencing. bsRNA-seq is an approach where the RNA is treated with bisulfite, which converts unmethylated cytosines to uracil, while methylated cytosines are resistant to the treatment. Pseudouridine is detected through two similar methods. RNA is treated with CMCT, resulting in Ψ-CMC. As the CMC-Ψ is a bulky residue, it can block the reverse transcriptase, resulting in a stop or pause event one nucleotide before the pseudouridine site. The reverse transcription step is followed by next generation sequencing. m^5C = 5-methylcytidine, m^6A = N6-methyladenosine, Ψ = pseudouridine, CMC = carbodiimide

Indeed, they observed that ALKBH5 KO mice have significantly smaller whole brains and cerebellum compared to WT mice after being exposed to hypobaric hypoxia for 48 h. Another recent study found that ALKBH5 is generally expressed in the mouse brain, being more prominent in the cerebellum and olfactory bulb of the adult mouse brain (Du et al. 2020). Also, they detected that ALKBH5 co-localized with the neuronal marker NeuN, suggesting ALKBH5 to be mainly expressed in neurons. Finally, they also observed that ALKBH5 protein expression significantly decreased during brain development. Altogether, these observations suggest further work is required to fully understand the effect of these demethylates in different brain regions where redundancy exists and whether ALKBH5 and FTO are involved in different aspects of brain development (dependent on the region) and/or do they have different targets. Furthermore, m^6A demethylation by FTO plays an important role in dopaminergic transmission, which is involved in the control of complex behaviors (Hess et al. 2013). The authors found that levels of m^6A increased a subset of mRNAs associated with neuronal signaling, including the dopaminergic signaling pathway in the midbrain and striatum of Fto knockout mice. This resulted in several proteins with altered expression and impaired dopaminergic transmission. Another study found knocking down FTO in the mouse medial prefrontal cortex resulted in enhanced cued fear memory, while m^6A levels increased in the same tissue after behavioral training (Widagdo et al. 2016).

2.3 *m^6A and Neurological Disorders*

Currently, there is little known regarding the specific role of m^6A in neurodegenerative and neuropsychiatric diseases. However, there is strong evidence to suggest a fundamental role for m^6A in neurodegenerative and neuropsychiatric disease. For example, genetic variants of FTO have been implicated in Alzheimer's disease (AD), e.g., see Ho et al. (2010), Li et al. (2018a), and Reitz et al. (2012). For example, the interaction between FTO and APOE was found, and individuals carrying both FTO-AA genotype and APOE $\epsilon 4$ were at a higher risk for dementia (Keller et al. 2011). Along those lines, it was shown that variation in introns 1 and 2 of the FTO gene may be associated with AD (Reitz et al. 2012). More recently, a possible association between m^6A and AD was revealed (Han et al. 2020). The authors quantified total RNA m^6A levels in an Alzheimer's mouse model (APP/PS1 transgenic mice), compared to C57BL/6 wild-type mice, revealing increased levels of total RNA m^6A in the cortex and the hippocampus of APP/PS1 transgenic mice. Furthermore, the authors found that the expression of METTL3 and FTO increased and decreased in AD mice, respectively, compared to wild type. This study suggests a possible role for m^6A in regulating Alzheimer's disease, but a more in-depth study is required to actually elucidate the role of m^6A. Another recent study used a public human Alzheimer's brain RNA-seq dataset to determine the expression profiles of m^6A players (Huang et al. 2020). They found that METTL3 and RBM15B were upregulated and downregulated in the hippocampus, respectively. Those observations were independently validated at the protein level in another set of postmortem

human brains. They also suggest that the accumulation of METTL3 positively correlates (albeit weakly) with Tau aggregates. The effect of m^6A has also been looked at in the context of Parkinson's disease (PD) (Chen et al. 2019b). They model Parkinson's disease in rats and PC12 cells by employing 6-OHDA to selectively destroy dopaminergic neurons. Upon this treatment, they found that the m^6A modification was reduced in the PC12 cells in the striatum of the PD rat but not the cortex, hippocampus, or midbrain. This observation seems to correlate with the fact that a significant increase in ALKBH5 is present only in the striatum of PD mice but not in any other region. Interestingly, an extensive study was performed to determine variants of the m^6A players in 1647 sporadic Parkinson patients and 1372 controls of Han Chinese origin (Qin et al. 2020). The authors identified 214 rare variants in METTL3, METTL14, WTAP, FTO, ALKBH5, YTHDF1, YTHDF2, YTHDF3, HNRNPC, and ELAVL1. However, gene-wise association studies did not achieve significance, suggesting that there is no biological association between these m^6A associated proteins and sporadic Parkinson's disease.

m^6A is present in synapses, and aberrant translation at synapses has been associated with autism, fragile X syndrome, and other intellectual disorders, suggesting there may be a role for m^6A in these diseases. For example, mutations in YTHDC2 have been implicated as a risk factor for autism spectrum disorder (Liu et al. 2016b). Also, depressed patients have altered levels of m^6A/m after glucocorticoid stimulation, suggesting that these modifications may also have a role in stress-related psychiatric disorders (Engel et al. 2018). Lastly, FTO polymorphisms have also been found in psychiatric diseases, including major depressive disorder (MDD). Individuals carrying the FTO rs9939609 allele were demonstrated to have an inverse association between obesity risk and depression (Milaneschi et al. 2014). Another study also found an association between MDD and allelic variants of ALKBH5 (Du et al. 2015).

3 N1-methyladenosine

Several enzymes have been reported to install m^1A on rRNA and tRNA, including TRM6, TRM10, and TRM61 (Bujnicki 2001), and the modification can be removed by ALKB, ALBH1, and ALKBH3 (Li et al. 2016; Liu et al. 2016a). No readers have been identified as yet. Furthermore, m^1A, unlike m^6A and m^6Am, perturbs Watson–Crick base pairing and potentially disrupts protein–RNA interactions and RNA secondary structures through electrostatic effects (Zhou et al. 2016).

3.1 Detection of m^1A

Currently, the literature surrounding m^1A is conflicting and not much is known about the modification. Using an antibody approach similar to the m^6A MeRIP-Seq

method, m^1A was recently mapped transcriptome-wide and proposed as a prominent modification in the 5′UTR of mRNAs (Dominissini et al. 2016; Li et al. 2017c; Safra et al. 2017). However, the further analysis presented in a recent study revealed that the m^1A antibody used in these studies had an affinity for the m^7G-cap and resulted in false-positives (Grozhik et al. 2019). Furthermore, the authors used another m^1A antibody that did not cross-react with the m^7G cap, and m^1A was not detected in the 5′ UTR. Indeed, they demonstrate that m^1A sites are rare, if not absent, from mRNAs. Furthermore, m^1A is known to cause truncation during reverse transcription, however, there is much uncertainty associated with this approach. To overcome this, under alkaline conditions, m^1A can be converted to m^6A in a process known as the Dimroth rearrangement. As m^6A does not have a signature during reverse transcription, a disappearance of a reverse transcription signal will be observed after the treatment, signaling the presence of m^1A (Dominissini et al. 2016).

3.2 *m^1A in the Brain*

m^1A abundance has been measured in various tissues using LC-MS/MS, and m^1A was shown to be the highest in the kidney and the brain (Dominissini et al. 2016). Furthermore, they detected four times as much m^1A in the brains of lean (wt/wt) mice compared to obese (ob/ob) mice. That study provides some evidence that m^1A levels vary according to tissue type, is an abundant modification in the brain, and also that the modification can possibly be a dynamic modification in response to physiological signaling. Also, another study determined m^1A levels across 39 different tissues and found that m^1A was highest in whole blood, brain, muscle, and nerve tissues (Ali et al. 2020). Furthermore, m^1A levels increased in the brain of FTO −/− mice compared to wild type, suggesting that in addition to being an m^6A and m^6Am demethylase, FTO mediates demethylation of m^1A (Wei et al. 2018) show that. As loss of m^1A tRNA methylation may regulate translation (Liu et al. 2016a), some of the FTO mediated phenotypes observed in brain tissues (see above) may be in part because of tuning translation through tRNA m^1A demethylation.

4 5-Methylcytidine

RNA m^5C methyltransferases belong to the superfamily of Rossman fold-containing enzymes that use *S*-adenosyl-*L*-methionine (SAM), which functions as a cofactor to donate the methyl group. Currently, known m^5C methyltransferases belong to either the DNMT2 or the NOL1/NOP2/SUN (NSUN) subgroups (Bujnicki et al. 2004; Motorin et al. 2010). In mammals, the NSUN family includes seven enzymes (NSUN1-NSUN7). NSUN2 is the main m^5C methyltransferase and has been shown to target tRNAs, non-coding RNAs, and mRNAs (Amort et al. 2017; Hussain et al. 2013; Khoddami and Cairns 2013; Squires et al. 2012). Substrates for NSUN1

and NSUN5 are 28S rRNA (Heissenberger et al. 2019; Sharma et al. 2013), NSUN4 targets 12S mitochondrial rRNA (Metodiev et al. 2014), NSUN3 and NSUN6 recognize tRNAs (Haag et al. 2015, 2016), and NSUN7 has been shown to methylate enhancer RNAs (Aguilo et al. 2016). Recently, the first m^5C "reader," ALYREF protein was discovered (Yang et al. 2017). The authors showed that ALYREF promotes mRNA export of m^5C-marked mRNAs. m^5C has also been associated with the YBX1 protein that was shown to stabilize a subset of maternal mRNAs by recruiting poly(A)-binding protein cytoplasmic 1 during the maternal to zygotic transition in zebrafish (Yang et al. 2019). The modification has also been implicated as potentially having a role in regulating mRNA translation. For example, specific m^5C methylation in the 3′UTR of the cell cycle regulators CDK1 and p21 resulted in enhanced translation in vitro and using a gene reporter in vivo (Li et al. 2017b; Xing et al. 2015). Also, m^5C may reduce translation when introduced at any position within a codon and altered codon specificity when in the second position of the codon (Hoernes et al. 2016). However, this observation was achieved using a bacterial system and still remains to be validated for eukaryotic translation.

4.1 Detection of m^5C

Currently, several high-throughput m^5C detection methods exist. m^5C was first mapped transcriptome-wide using bisulfite RNA sequencing (Squires et al. 2012). In the presence of sodium bisulfite, cytosine is converted into uracil by deamination while methylated cytosines are resistant to the treatment. A standard RNA-seq then allows for the identification of m^5C. Other currently used detection techniques are m^5C RNA immunoprecipitation (m^5C-RIP), 5-azacytidine-mediated RNA immunoprecipitation (Aza-IP) (Khoddami and Cairns 2013), or methylation-individual nucleotide resolution cross-linking immunoprecipitation (miCLIP) (Hussain et al. 2013). m^5C-RIP uses m^5C specific antibodies to selectively enrich for m^5C-containing RNA. Aza-IP randomly incorporates the cytidine analog, azacytidine, into nascent RNA transcripts in place of cytosine. This results in an RNA-methylase covalent adduct that can be immunoprecipitated with enzyme-specific or tag antibodies. Lastly, release of RNA from the complex results in hydrolytic opening of the azacytidine ring, which is read as a guanine following reverse transcription. The miCLIP approach employs a mutant form of the m^5C writer NSUN2, which allows for capturing NSUN2 bound RNA complexes. This is coupled with an individual-nucleotide resolution cross-linking immunoprecipitation (iCLIP) approach and next generation sequencing (Fig. 3).

4.2 m^5C in Neurodevelopment

There is some evidence to suggest that m^5C may be a significant regulator of brain development processes. For example, recently, an ultrahigh performance liquid

chromatography–multiple reaction monitoring tandem mass spectrometry (UHPLC–MRM-MS/MS) analysis identified m^5C in mRNAs from six mouse tissues (small intestine, heart, muscle, brain, kidney, and liver) (Yang et al. 2017). This revealed the highest concentration of m^5C in the brain compared to the other tested tissues. A total of 4371 m^5C sites in 1655 mRNAs were detected. Of those 4371 sites, 1918 are specific to the brain and correlate with the highest mRNA expression in the brain compared to the other tissues. These transcripts are involved in trans-synaptic signaling, nervous system development, cell-cell signaling, and neurogenesis processes, pointing to a role for m^5C in neurodevelopment. m^5C in the mouse brain was also detected by Amort et al. (2017). They found ~7500 m^5C sites (>20% methylation) mapping to 1650 mRNAs were detected in embryonic stem cells (ESCs) and 2075 m^5C sites mapping to 486 mRNAs in the brain, with m^5C being its highest in the 3′ UTR of transcripts methylated in the brain. Interestingly, mRNAs that are specifically methylated in ESCs were expressed but not methylated in the brain. The uniquely methylated transcripts in the brain are not expressed in ESCs and are involved in ion transport or synapse function biological processes. The authors also found an overlap between the 3′ UTR m^5C sites and RNA-binding protein sites, including UPF1 and splicing factors SRSF3 and SRSF4. Despite this association being statistically significant, very low numbers of overlapping sites were actually observed, suggesting that any gene regulation through such a mechanism may be specific to a particular transcript.

hm^5C is generated through the active demethylation of m^5C by the Tet enzyme. hm^5C in RNA was shown to be at the highest levels in the brainstem, hippocampus, and cerebellum using a dot blot approach (Miao et al. 2016). Interestingly, they detected less hm^5C in an MPTP-induced Parkinson's disease mouse model compared to wild type. Another study found that hydroxymethylated RNA and the Tet enzyme are prominent in the Drosophila brain (Delatte et al. 2016). Depletion of Tet results in decreased RNA hydroxymethylation and consequently impaired Drosophila brain development. Clearly, both m^5C and hm^5C have significant roles in the brain, and further work is required to fully elucidate the impact and mechanisms of m^5C on neurodevelopment.

Expression of NSUN2, the m^5C writer, is highest in the cortex, hippocampus, and striatum of the developing mouse brain (Blanco et al. 2014). Furthermore, loss of NSUN2 resulted in the increased cleavage of NSUN2 tRNA targets by the endonuclease angiogenin, which in turn resulted in the accumulation of 5′ tRNA fragments (Blanco et al. 2014) and consequently a decrease in global protein translation (Gebetsberger et al. 2012; Ivanov et al. 2011). This manifested itself as increased cellular stress and reduction in size in those tissues where NSUN2 levels are highest. Importantly, both cellular stress and microcephaly can be rescued through the inhibition of angiogenin (Blanco et al. 2014). In humans, NSUN2 is expressed in early neuroepithelial progenitors, and levels progressively decrease during differentiation of neuroepithelial stem (NES) cells in vitro, and the process is impaired by both absence of NSUN2 and the presence of angiogenin (Flores et al. 2017). The authors also observed that loss of *Nsun2* caused an accumulation of intermediate progenitors and a decrease in differentiated upper-layer neurons in the cortex. Also,

NSUN5 expression was found to be highest in oligodendrocyte precursor cells compared to neurons or astrocytes (Zhang et al. 2019a). On this note, the authors found that the loss of *Nsun5* inhibits NMDA receptor activity in neuronal cells, possibly as a result of impaired development and function of oligodendrocyte precursor cells. Another study focused on the role of NSUN5 in the mouse cerebral cortex (Chen et al. 2019a). At P10, the authors found a significant reduction in cortical thickness and abnormal laminar organization in Nsun5-KO mice compared to wild-type littermates. At embryonic day (E) 12.5 to E16.5, the authors find that NSUN5 is expressed in radial glial cells of the cerebral cortex, where it is required to properly maintain the radial glial scaffolds in order to regulate neocortical neuron migration. Clearly, proper connections within the brain and cortex organization are required to ensure the correct development of complex social-cognitive functions, and these findings suggest that NSUN5 plays a particularly important role in cortex development.

4.3 m^5C in Neurological Disorders

Loss-of-function mutations in the m^5C writer, NSUN2, in both mouse and human have been linked to intellectual disability phenotypes, including growth retardation, microcephaly, impaired cognition, and motor function (Abbasi-Moheb et al. 2012; Martinez et al. 2012). A missense mutation at a conserved residue in NSUN2 causes a failure to localize to the nucleolus of Purkinje cells in the cerebellum, contributing to the intellectual disability phenotype (Khan et al. 2012).

Williams–Beuren syndrome is a neurodevelopmental disorder characterized by social-cognitive disorder. The NSUN5 gene, which encodes an m^5C methyltransferase, has been shown to be deleted in the syndrome. Adult *Nsun5*-KO mice that have been generated show spatial cognitive defects (Chen et al. 2019a; Zhang et al. 2019a).

5 Pseudouridine

Pseudouridylation can occur through two different mechanisms. In the first, enzymes belonging to the pseudouridine synthase (PUS) family recognize sequence and/or secondary structural elements of the RNA target and catalyze the isomerization of uridine to pseudouridine. There are ten pseudouridine synthases that are compartmentalized into six families (TruA, TruB, TruD, RsuA, RluA, and Pus10) (Spenkuch et al. 2014); however, all PUS enzymes share a conserved catalytic domain and it is likely that they all function through a conserved catalytic mechanism. In the second mechanism, the pseudouridylation event is guided by snoRNPs (H/ACA snoRNA assembled with core proteins). The core proteins include Cbf5/NAP57/Dyskerin, Nhp2/L7Ae, Nop10, and Gar1 (Massenet et al. 2017). The

snoRNA functions as a guide that base pairs with the substrate RNA, directing the core protein to the site for pseudouridylation. Interestingly, all recent transcriptome-wide studies mapping pseudouridine showed that the majority of mRNA pseudouridylation is catalyzed by the independent PUS enzymes. Lastly, pseudouridylation has been shown to result in stop codon read through or nonsense suppression in mRNAs (Karijolich and Yu 2011), while another study found pseudouridylated mRNAs to be highly inducible in response to serum starvation in humans (Carlile et al. 2014). Recently, Pus10 has been implicated in regulating miRNA processing by directly binding to primary miRNAs interacting with the microprocessor to promote miRNA biogenesis (Song et al. 2020).

5.1 Detection of Pseudouridine

There are two main methods to detect pseudouridine in RNA. The first method is based on *N*-Cyclohexyl-*N*′-(2-morpholinoethyl) carbodiimide methyl-*p*-toluenesulfonate (CMCT) modification of pseudouridine (Li et al. 2015; Lovejoy et al. 2014). Uridine and guanine can also be modified with CMCT resulting in Ψ-CMC, U-CMC, and G-CMC adducts. Fortunately, mild alkaline conditions can reverse the CMCT modification of uridine and guanine but not pseudouridine. Finally, as the CMC-Ψ is a bulky residue, it can block the reverse transcriptase, resulting in a stop or pause event one nucleotide before the pseudouridine site. There are presently two very similar approaches that use this idea termed Ψ-seq and CeU-seq (Fig. 3). The second method exploits site-specific cleavage and labeling of RNA (Zhao and Yu 2004). In this approach, RNA is cleaved at the target uridine/pseudouridine sites, the cleaved site is then radiolabeled, followed by nuclease digestion into mononucleotides, and, finally, analysis by thin-layer chromatography.

Pseudouridylation was shown to be as prevalent as m^6A in mRNAs using quantitative mass spectrometry (Li et al. 2015). Furthermore, using the chemical pulldown-based approach, the authors found high levels of pseudouridine in mRNAs from various mouse tissues, with particular enrichment in the brain and lung. They detected 1741 sites in mouse brain, and the pseudouridine-marked genes were enriched in nervous system development and signal transduction. These findings, as well as the fact that pseudouridylation can be dynamically regulated (Schwartz et al. 2014), raise the possibility that pseudouridine modification is involved in the regulation of neurodevelopmental processes.

5.2 Pseudouridine in Neurological Disorders

Novel homozygous protein-truncating variants in PUS3 and PUS7, respectively, which encode RNA-independent pseudouridylate synthases (de Brouwer et al. 2018; Shaheen et al. 2016). Patients with these mutations exhibited an intellectual

disability phenotype, including speech delay, a smaller physique, microcephaly, and aggressiveness. Furthermore, a significant decrease in pseudouridine modification at positions 38 and 39 in tRNA of patient cells was observed, consistent with a loss of PUS3 in these patients (Shaheen et al. 2016), while these protein-truncating mutations resulted in the loss of PUS7, and consequently loss of pseudouridine on tRNA and mRNA targets (de Brouwer et al. 2018). They validated their findings by studying the effects of PUS7 knockout in Drosophila. Deletion of *Pus7* resulted in a number of behavioral defects, including increased activity, disorientation, and aggressiveness. These findings suggest that PUS7-mediated RNA pseudouridylation is required to maintain proper neurodevelopment and function. These observations were also independently validated by a related study (Shaheen et al. 2019). Those authors also found one missense and one frameshift mutation in PUS7 that resulted in the abolishment of the enzyme, with consequent loss of pseudouridine at position 13 in PUS7 tRNA targets. This manifested itself in a strong intellectual disability phenotype and progressive microcephaly.

6 Other RNA Modifications

Since RNA modifications on small non-coding RNAs (sncRNAs) have been implicated as important regulators of physiological and pathological processes (e.g., Abe et al. 2014), it was investigated whether RNA modifications on small RNAs have an impact on Alzheimer's disease (AD) (Zhang et al. 2020). To this end, they analyzed modification profiles on both small RNAs (15–25 nt) and tRNA fractions extracted from the human cortex using LC-MS/MS (Fig. 3). In the 15–25-nt RNA fraction, the authors found increases in 2′-*O*-methylcytidine (Cm), 7-methylguanosine (m^7G), 2′-*O*-methylguanosine (Gm) and decreased levels of *N2*,*N2*,*N7*-trimethylguanosine ($m^{2,2,7}G$) and *N2*,*N2*-dimethylguanosine ($m^{2,2}G$) modifications in AD brains compared with controls. Interestingly, in the 30–40-nt fraction, a significant reduction in rRNA-derived small RNA (rsRNA)-5S, tRNA-derived small RNA (tsRNA)-Tyr, and tsRNA-Arg fragments were observed in AD patients compared with controls; however, the authors do not show any association between these changing RNA levels and RNA modifications. Ultimately, more work is required to understand whether these RNA modifications play a part in AD etiology and how this is achieved mechanistically.

Also, various tRNA modifications have been indirectly linked with neurological disorders. For example, mutations in the FtsJ methyltransferase homolog 1 (FTSJ1) gene have been linked with non-syndromic X-linked mental retardation and intellectual disability (Dai et al. 2008), and genetic variants of FTSJ1 have been associated with general cognitive ability, verbal comprehension, and perceptual organization (Gong et al. 2008). This gene encodes for a methyltransferase that is responsible for 2′-*O*-methylation modifications in $tRNA^{Leu}$, $tRNA^{Trp}$, and $tRNA^{Phe}$ at positions 32 and 34 (Gong et al. 2008; Guy et al. 2015). Furthermore, two genetically independent lymphoblastoid cell lines from non-syndromic X-linked

mental retardation and intellectual disability patients had loss-of-function *FTSJ1* mutations and a near complete loss of 2′-*O*-methylation at position 32 and 32 of $tRNA^{Phe}$ (Guy et al. 2015). This observation suggests 2′-*O*-methylation is linked to neurodevelopmental disorder. However, whether loss of the modification is a driving force for onset or is just a by-product of *FTSJ1* mutation remains to be determined. In any case, aberrant 2′-*O*-methylation is clearly associated with neurological disorder.

Furthermore, a homozygous frameshift mutation in human tRNA methyltransferase 1 has been associated with recessive cognitive disorders (Najmabadi et al. 2011). tRNA methyltransferase 1 catalyzes the dimethylation of guanosines ($m^{22}G$) at position 26 of tRNAs. More recently, studies have associated ELP2, ELP3, and ELP4, three subunits of the multisubunit elongator protein complex, with neurodevelopmental disability, amyotrophic lateral sclerosis, and atypical Rolandic epilepsy, respectively (Kojic and Wainwright 2016). Moreover, the elongator complex has an essential role in tRNA uridine modifications. Altogether these observations suggest that these tRNA modifications may be linked to neurological disorders. However, whether the modification has a direct or indirect role in the onset of these disorders and the possible mechanisms remains to be fully elucidated.

7 Future Directions and Perspectives

RNA modifications clearly afford another layer of regulation in the maintenance of the mammalian brain. However, further work is required to understand the role of the mechanisms by which the RNA modifications act in brain development and disease. Mapping RNA modifications, especially at the single-nucleotide level, will go a long way to further understanding the role of these modifications. For example, what cues (cellular or environmental) control whether a transcript is modified or unmodified, and how do the different modifications work in concert? Single-cell sequencing studies have started to emerge outlining the view of the transcriptome in each cell type in the brain, and this is being extending to look at the epitranscriptome. Currently, the advent of single-molecule direct RNA sequencing may shed light on the different RNA modifications present on a given transcript. This will allow us to begin to fully determine the impact of modifications on gene expression. Such technologies, as those that are being developed by Nanopore Sequencing, will be important to determine, using a small input of RNA, the combinations of modifications that are present on individual transcripts, in a specific cell type, at a specific time. This is critical to gain an understanding of how different modifications work together or against each other to contribute to development and disease. On this note, it is important to know whether it is the modification itself or the cellular machinery that "writes," "reads," or "erases" the modification that actually contributes to disease etiology. For example, it is the loss of methylation or the loss of the writer that plays a role in the onset of disease? Clearly, more work is required to fully

understand the mechanism by which these RNA modifications exert their function in neurodevelopment and disease.

References

Abbasi-Moheb L, Mertel S, Gonsior M et al (2012) Mutations in NSUN2 cause autosomal-recessive intellectual disability. Am J Hum Genet 90(5):847–855

Abe M, Naqvi A, Hendriks GJ et al (2014) Impact of age-associated increase in 2′-O-methylation of miRNAs on aging and neurodegeneration in Drosophila. Genes Dev 28(1):44–57

Agarwala SD, Blitzblau HG, Hochwagen A et al (2012) RNA methylation by the MIS complex regulates a cell fate decision in yeast. PLoS Genet 8(6):e1002732

Aguilo F, Li S, Balasubramaniyan N et al (2016) Deposition of 5-methylcytosine on enhancer RNAs enables the coactivator function of PGC-1alpha. Cell Rep 14(3):479–492

Alarcon CR, Lee H, Goodarzi H et al (2015) N6-methyladenosine marks primary microRNAs for processing. Nature 519(7544):482–485

Ali AT, Idaghdour Y, Hodgkinson A (2020) Analysis of mitochondrial m1A/G RNA modification reveals links to nuclear genetic variants and associated disease processes. Commun Biol 3 (1):147

Amort T, Rieder D, Wille A et al (2017) Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain. Genome Biol 18(1):1

Batista PJ, Molinie B, Wang J et al (2014) m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell 15(6):707–719

Blanco S, Dietmann S, Flores JV et al (2014) Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. EMBO J 33(18):2020–2039

Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46(D1):D303–D307

Bokar JA, Shambaugh ME, Polayes D et al (1997) Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA 3 (11):1233–1247

Bujnicki JM (2001) In silico analysis of the tRNA:m1A58 methyltransferase family: homology-based fold prediction and identification of new members from eubacteria and archaea. FEBS Lett 507(2):123–127

Bujnicki JM, Feder M, Ayres CL et al (2004) Sequence-structure-function studies of tRNA:m5C methyltransferase Trm4p and its relationship to DNA:m5C and RNA:m5U methyltransferases. Nucleic Acids Res 32(8):2453–2463

Carlile TM, Rojas-Duran MF, Zinshteyn B et al (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515(7525):143–146

Chen P, Zhang T, Yuan Z et al (2019a) Expression of the RNA methyltransferase Nsun5 is essential for developing cerebral cortex. Mol Brain 12(1):74

Chen X, Yu C, Guo M et al (2019b) Down-regulation of m6A mRNA methylation is involved in dopaminergic neuronal death. ACS Chem Neurosci 10(5):2355–2363

Dai L, Xing L, Gong P et al (2008) Positive association of the FTSJ1 gene polymorphisms with nonsyndromic X-linked mental retardation in young Chinese male subjects. J Hum Genet 53 (7):592–597

de Brouwer APM, Abou Jamra R, Kortel N et al (2018) Variants in PUS7 cause intellectual disability with speech delay, microcephaly, short stature, and aggressive behavior. Am J Hum Genet 103(6):1045–1052

Delatte B, Wang F, Ngoc LV et al (2016) RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. Science 351(6270):282–285

Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485(7397):201–206
Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S et al (2016) The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. Nature 530(7591):441–446
Du T, Rao S, Wu L et al (2015) An association study of the m6A genes with major depressive disorder in Chinese Han population. J Affect Disord 183:279–286
Du T, Li G, Yang J et al (2020) RNA demethylase Alkbh5 is widely expressed in neurons and decreased during brain development. Brain Res Bull 163:150–159
Edens BM, Vissers C, Su J et al (2019) FMRP modulates neural differentiation through m(6)A-dependent mRNA nuclear export. Cell Rep 28(4):845–854. e845
Engel M, Eggert C, Kaplick PM et al (2018) The role of m(6)A/m-RNA methylation in stress response regulation. Neuron 99(2):389–403. e389
Feng J, Fouse S, Fan G (2007) Epigenetic regulation of neural gene expression and neuronal function. Pediatr Res 61(5 Pt 2):58R–63R
Flores JV, Cordero-Espinoza L, Oeztuerk-Winder F et al (2017) Cytosine-5 RNA methylation regulates neural stem cell differentiation and motility. Stem Cell Rep 8(1):112–124
Gebetsberger J, Zywicki M, Kunzi A et al (2012) tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in Haloferax volcanii. Archaea 2012:260909
Gong P, Li J, Dai L et al (2008) Genetic variations in FTSJ1 influence cognitive ability in young males in the Chinese Han population. J Neurogenet 22(4):277–287
Grozhik AV, Olarerin-George AO, Sindelar M et al (2019) Antibody cross-reactivity accounts for widespread appearance of m(1)A in 5′UTRs. Nat Commun 10(1):5126
Guy MP, Shaw M, Weiner CL et al (2015) Defects in tRNA anticodon Loop 2′-O-methylation are implicated in nonsyndromic X-Linked intellectual disability due to mutations in FTSJ1. Hum Mutat 36(12):1176–1187
Haag S, Warda AS, Kretschmer J et al (2015) NSUN6 is a human RNA methyltransferase that catalyzes formation of m5C72 in specific tRNAs. RNA 21(9):1532–1543
Haag S, Sloan KE, Ranjan N et al (2016) NSUN3 and ABH1 modify the wobble position of mt-tRNAMet to expand codon recognition in mitochondrial translation. EMBO J 35 (19):2104–2119
Han M, Liu Z, Xu Y et al (2020) Abnormality of m6A mRNA methylation is involved in Alzheimer's disease. Front Neurosci 14(98)
Heissenberger C, Liendl L, Nagelreiter F et al (2019) Loss of the ribosomal RNA methyltransferase NSUN5 impairs global protein synthesis and normal growth. Nucleic Acids Res 47 (22):11807–11825
Hess ME, Hess S, Meyer KD et al (2013) The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. Nat Neurosci 16(8):1042–1048
Hirabayashi Y, Gotoh Y (2010) Epigenetic control of neural precursor cell fate during development. Nat Rev Neurosci 11(6):377–388
Ho AJ, Stein JL, Hua X et al (2010) A commonly carried allele of the obesity-related FTO gene is associated with reduced brain volume in the healthy elderly. Proc Natl Acad Sci USA 107 (18):8404–8409
Hoernes TP, Clementi N, Faserl K et al (2016) Nucleotide modifications within bacterial messenger RNAs regulate their translation and are able to rewire the genetic code. Nucleic Acids Res 44 (2):852–862
Huang H, Camats-Perna J, Medeiros R et al (2020) Altered expression of the m6A methyltransferase METTL3 in Alzheimer's disease. eNeuro 7(5)
Hussain S, Sajini AA, Blanco S et al (2013) NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep 4(2):255–261
Ignatova VV, Stolz P, Kaiser S et al (2020) The rRNA m(6)A methyltransferase METTL5 is involved in pluripotency and developmental programs. Genes Dev 34(9-10):715–729
Ivanov P, Emara MM, Villen J et al (2011) Angiogenin-induced tRNA fragments inhibit translation initiation. Mol Cell 43(4):613–623

Jia G, Fu Y, Zhao X et al (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7(12):885–887
Jonkhout N, Tran J, Smith MA et al (2017) The RNA modification landscape in human disease. RNA 23(12):1754–1769
Kadumuri RV, Janga SC (2018) Epitranscriptomic code and its alterations in human disease. Trends Mol Med 24(10):886–903
Karijolich J, Yu YT (2011) Converting nonsense codons into sense codons by targeted pseudouridylation. Nature 474(7351):395–398
Keller L, Xu W, Wang HX et al (2011) The obesity related gene, FTO, interacts with APOE, and is associated with Alzheimer's disease risk: a prospective cohort study. J Alzheimers Dis 23 (3):461–469
Khan MA, Rafiq MA, Noor A et al (2012) Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. Am J Hum Genet 90 (5):856–863
Khoddami V, Cairns BR (2013) Identification of direct targets and modified bases of RNA cytosine methyltransferases. Nat Biotechnol 31(5):458–464
Kojic M, Wainwright B (2016) The many faces of elongator in neurodevelopment and disease. Front Mol Neurosci 9:115
Koranda JL, Dore L, Shi H et al (2018) Mettl14 is essential for epitranscriptomic regulation of striatal function and learning. Neuron 99(2):283–292. e285
Li X, Zhu P, Ma S et al (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol 11(8):592–597
Li X, Xiong X, Wang K et al (2016) Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. Nat Chem Biol 12(5):311–316
Li L, Zang L, Zhang F et al (2017a) Fat mass and obesity-associated (FTO) protein regulates adult neurogenesis. Hum Mol Genet 26(13):2398–2411
Li Q, Li X, Tang H et al (2017b) NSUN2-mediated m5C methylation and METTL3/METTL14-mediated m6A methylation cooperatively enhance p21 translation. J Cell Biochem 118 (9):2587–2598
Li X, Xiong X, Zhang M et al (2017c) Base-resolution mapping reveals distinct m(1)A methylome in nuclear- and mitochondrial-encoded transcripts. Mol Cell 68(5):993–1005. e1009
Li H, Ren Y, Mao K et al (2018a) FTO is involved in Alzheimer's disease by targeting TSC1-mTOR-Tau signaling. Biochem Biophys Res Commun 498(1):234–239
Li M, Zhao X, Wang W et al (2018b) Ythdf2-mediated m(6)A mRNA clearance modulates neural development in mice. Genome Biol 19(1):69
Linder B, Grozhik AV, Olarerin-George AO et al (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods 12(8):767–772
Liu J, Yue Y, Han D et al (2014) A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol 10(2):93–95
Liu F, Clark W, Luo G et al (2016a) ALKBH1-mediated tRNA demethylation regulates translation. Cell 167(3):816–828. e816
Liu X, Shimada T, Otowa T et al (2016b) Genome-wide association study of autism spectrum disorder in the East Asian populations. Autism Res 9(3):340–349
Lovejoy AF, Riordan DP, Brown PO (2014) Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*. PLoS One 9(10):e110799
Ma C, Chang M, Lv H et al (2018) RNA m(6)A methylation participates in regulation of postnatal development of the mouse cerebellum. Genome Biol 19(1):68
Mao Y, Dong L, Liu XM et al (2019) m(6)A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. Nat Commun 10(1):5332
Martinez FJ, Lee JH, Lee JE et al (2012) Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. J Med Genet 49(6):380–385
Massenet S, Bertrand E, Verheggen C (2017) Assembly and trafficking of box C/D and H/ACA snoRNPs. RNA Biol 14(6):680–692

Merkurjev D, Hong WT, Iida K et al (2018) Synaptic N(6)-methyladenosine (m(6)A) epitranscriptome reveals functional partitioning of localized transcripts. Nat Neurosci 21 (7):1004–1014
Metodiev MD, Spahr H, Loguercio Polosa P et al (2014) NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly. PLoS Genet 10(2):e1004110
Meyer KD (2019) DART-seq: an antibody-free method for global m(6)A detection. Nat Methods 16(12):1275–1280
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149(7):1635–1646
Miao Z, Xin N, Wei B et al (2016) 5-hydroxymethylcytosine is detected in RNA from mouse brain tissues. Brain Res 1642:546–552
Milaneschi Y, Lamers F, Mbarek H et al (2014) The effect of FTO rs9939609 on major depression differs across MDD subtypes. Mol Psychiatry 19(9):960–962
Motorin Y, Lyko F, Helm M (2010) 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. Nucleic Acids Res 38(5):1415–1430
Najmabadi H, Hu H, Garshasbi M et al (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature 478(7367):57–63
Patil DP, Chen CK, Pickering BF et al (2016) m(6)A RNA methylation promotes XIST-mediated transcriptional repression. Nature 537(7620):369–373
Pendleton KE, Chen B, Liu K et al (2017) The U6 snRNA m(6)A methyltransferase METTL16 regulates SAM synthetase intron retention. Cell 169(5):824–835. e814
Ping XL, Sun BF, Wang L et al (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res 24(2):177–189
Qin L, Min S, Shu L et al (2020) Genetic analysis of N6-methyladenosine modification genes in Parkinson's disease. Neurobiol Aging 93:143 e149–143 e113
Reitz C, Tosto G, Mayeux R et al (2012) Genetic variants in the fat and obesity associated (FTO) gene and risk of Alzheimer's disease. PLoS One 7(12):e50354
Safra M, Sas-Chen A, Nir R et al (2017) The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature 551(7679):251–255
Schumann U, Zhang HN, Sibbritt T et al (2020) Multiple links between 5-methylcytosine content of mRNA and translation. BMC Biol 18(1):40
Schwartz S, Bernstein DA, Mumbach MR et al (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell 159(1):148–162
Shafik AM, Allen EG, Jin P (2020) Dynamic N6-methyladenosine RNA methylation in brain and diseases. Epigenomics 12(4):371–380
Shaheen R, Han L, Faqeih E et al (2016) A homozygous truncating mutation in PUS3 expands the role of tRNA modification in normal cognition. Hum Genet 135(7):707–713
Shaheen R, Tasak M, Maddirevula S et al (2019) PUS7 mutations impair pseudouridylation in humans and cause intellectual disability and microcephaly. Hum Genet 138(3):231–239
Sharma S, Yang J, Watzinger P et al (2013) Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. Nucleic Acids Res 41(19):9062–9076
Shi H, Wang X, Lu Z et al (2017) YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. Cell Res 27(3):315–328
Shi H, Zhang X, Weng YL et al (2018) m(6)A facilitates hippocampus-dependent learning and memory through YTHDF1. Nature 563(7730):249–253
Song J, Zhuang Y, Zhu C et al (2020) Differential roles of human PUS10 in miRNA processing and tRNA pseudouridylation. Nat Chem Biol 16(2):160–169
Spenkuch F, Motorin Y, Helm M (2014) Pseudouridine: still mysterious, but never a fake (uridine)! RNA Biol 11(12):1540–1554
Squires JE, Patel HR, Nousch M et al (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40(11):5023–5033

Sun J, Sun J, Ming GL et al (2011) Epigenetic regulation of neurogenesis in the adult mammalian brain. Eur J Neurosci 33(6):1087–1093

Tang C, Klukovich R, Peng H et al (2018) ALKBH5-dependent m6A demethylation controls splicing and stability of long 3′-UTR mRNAs in male germ cells. Proc Natl Acad Sci USA 115 (2):E325–E333

Theler D, Dominguez C, Blatter M et al (2014) Solution structure of the YTH domain in complex with N6-methyladenosine RNA: a reader of methylated RNA. Nucleic Acids Res 42 (22):13911–13919

Wang X, Lu Z, Gomez A et al (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505(7481):117–120

Wang X, Zhao BS, Roundtree IA et al (2015) N(6)-methyladenosine modulates messenger RNA translation efficiency. Cell 161(6):1388–1399

Wang Y, Li Y, Yue M et al (2018) N(6)-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications. Nat Neurosci 21(2):195–206

Wei J, Liu F, Lu Z et al (2018) Differential m(6)A, m(6)Am, and m(1)A demethylation mediated by FTO in the cell nucleus and cytoplasm. Mol Cell 71(6):973–985. e975

Widagdo J, Zhao QY, Kempen MJ et al (2016) Experience-dependent accumulation of N6-methyladenosine in the prefrontal cortex is associated with memory processes in mice. J Neurosci 36(25):6771–6777

Xing J, Yi J, Cai X et al (2015) NSun2 promotes cell growth via elevating cyclin-dependent kinase 1 translation. Mol Cell Biol 35(23):4043–4052

Yang X, Yang Y, Sun BF et al (2017) 5-methylcytosine promotes mRNA export – NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res 27(5):606–625

Yang Y, Wang L, Han X et al (2019) RNA 5-methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA decay. Mol Cell 75(6):1188–1202. e1111

Yao B, Christian KM, He C et al (2016) Epigenetic mechanisms in neurogenesis. Nat Rev Neurosci 17(9):537–549

Yoon KJ, Ringeling FR, Vissers C et al (2017) Temporal control of mammalian cortical neurogenesis by m(6)A methylation. Cell 171(4):877–889. e817

Yue Y, Liu J, Cui X et al (2018) VIRMA mediates preferential m(6)A mRNA methylation in 3′UTR and near stop codon and associates with alternative polyadenylation. Cell Discov 4:10

Zhang F, Kang Y, Wang M et al (2018) Fragile X mental retardation protein modulates the stability of its m6A-marked messenger RNA targets. Hum Mol Genet 27(22):3936–3950

Zhang T, Chen P, Li W et al (2019a) Cognitive deficits in mice lacking Nsun5, a cytosine-5 RNA methyltransferase, with impairment of oligodendrocyte precursor cells. Glia 67(4):688–702

Zhang Z, Chen LQ, Zhao YL et al (2019b) Single-base mapping of m(6)A by an antibody-independent method. Sci Adv 5(7):eaax0250

Zhang X, Trebak F, Souza LAC et al (2020) Small RNA modifications in Alzheimer's disease. Neurobiol Dis 145:105058

Zhao X, Yu YT (2004) Pseudouridines in and near the branch site recognition region of U2 snRNA are required for snRNP biogenesis and pre-mRNA splicing in Xenopus oocytes. RNA 10 (4):681–690

Zhao X, Yang Y, Sun BF et al (2014) FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res 24(12):1403–1419

Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49(1):18–29

Zhou H, Kimsey IJ, Nikolova EN et al (2016) m(1)A and m(1)G disrupt A-RNA structure through the intrinsic instability of Hoogsteen base pairs. Nat Struct Mol Biol 23(9):803–810

Zhou KI, Shi H, Lyu R et al (2019) Regulation of co-transcriptional Pre-mRNA splicing by m(6)A through the low-complexity protein hnRNPG. Mol Cell 76(1):70–81. e79

RNA Modifications in Neurodegenerations

Eleonora Calzoni, Chiara Argentati, Alessio Cesaretti, Nicolò Montegiove, Ilaria Tortorella, Martina Bazzucchi, Francesco Morena, Sabata Martino, and Carla Emiliani

Contents

Abstract Considerable evidence is accumulating about the regulatory function of coding- and non-coding RNAs through chemical modifications on their own nitrogen bases. The mechanisms by which the amount and types of these modifications are built to each type of RNA are yet poorly clarified, although some classes of proteins have been identified as actors able to introduce (*writers*), specifically recognize (*readers*) or delate (*erasers*) such modifications. In this context, advances have been made thanks to bioinformatic tools, which have allowed a comprehensive database, where RNA processing and post-transcriptional modification are integrated and elucidated, to be generated. Now, it is accepted that the abovementioned *epitranscriptome* interplay is involved in orchestrating development, health, and disease, also including the development, function, and dysfunction of the nervous system. In this chapter, we aim at reviewing and describing the potential

Eleonora Calzoni, Chiara Argentati and Alessio Cesaretti contributed equally with all other contributors.

E. Calzoni · C. Argentati · A. Cesaretti · N. Montegiove · I. Tortorella · M. Bazzucchi · F. Morena · S. Martino · C. Emiliani (✉)
Department of Chemistry, Biology and Biotechnology, University of Perugia, Perugia, Italy
e-mail: sabata.martino@unipg.it; carla.emiliani@unipg.it

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_2

mechanisms associated with the dysregulation of the *epitranscriptome* and neuropathologies, focusing specifically on brain cancer and neurodegenerative disorders.

Keywords Epitranscriptome · Brain cancer · Neurodegenerative diseases · Non-coding RNAs · RNA modifications

1 Introduction

According to the central dogma of molecular biology, the genetic information passes from DNA to RNA to proteins, and this allows a specific phenotype for a cell or an organism to be obtained (Garcia et al. 2020).

The technical development of the last 10 years has, however, exposed the incredible complexity of RNA. In the past, RNAs were generally known to be intermediates between DNA and proteins, except for infrastructural RNAs, that is, ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs). In recent decades, the rapid evolution of high-throughput technologies for the sequencing of the eukaryotic genome has revealed the central role of RNA in gene regulation. This has led to re-examining the function of RNAs, besides protein translation, in the development and evolution of higher organisms (Morena et al. 2018; Li and Liu 2019).

In fact, only a small proportion (less than 3%) of genetic transcripts encodes for proteins, and this fact raises the question of whether the remaining non-protein-coding transcripts are transcriptional "noise" or contain genetic information. As a matter of fact, a small part of the mammalian genome is transcribed into messenger RNAs (mRNAs), while the great majority of the genome is transcribed to what has been referred to as "dark matter:" the so-called non-coding RNAs (ncRNAs), which encode no information about proteins. Among ncRNAs, the long ncRNAs (lncRNAs) represent the most prevalent and functionally diverse class and are defined as transcripts longer than 200 nucleotides with low/no protein-coding potential (Hon et al. 2017; Li and Liu 2019).

Recent bioinformatic data analysts have indeed shown how a portion of these transcripts contains Open Reading Frames (ORFs) and joins with ribosomes. This finding indicates the possibility for the lncRNAs, in rare cases, to encode small peptides of less than 100 amino acids (AA) in length, called micropeptides, which are involved in different biological processes. It thus seems rational to suppose that RNA demarcation is somehow undefined and partly complicated and not exclusively depending upon its coding or non-coding status. In fact, RNAs cannot be categorically classified either as mRNA or as ncRNA, but rather their roles converge and overlap (Yeasmin et al. 2018; Hartford and Lal 2020).

Compared to ENCODE (Encyclopedia of DNA Elements) consortium data, when these sequences of ncRNA were analyzed, several groups of these molecules were found to be produced by pathways similar to those of the protein-coding genes. These results indicate a variety of ncRNAs, generally divided into two main

categories: structural non-coding RNAs and regulatory non-coding RNAs. Non-coding structural RNAs include rRNAs and tRNAs. Regulatory ncRNAs, besides the lncRNA, also include classes of RNAs of less than 200 nucleotides in length, known as small and medium RNAs (Alvarez-Dominguez and Lodish 2017).

In detail, microRNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA), telomere specific small RNA (tel-sRNA), cis-regulatory RNA (cisRNA), are classified as small ncRNA (sncRNA) made up of 20–50 nucleotides; while small nucleolar RNA (SnoRNA), tRNA-derived stress-induced RNA (tiRNA), small nuclear RNA (SnRNA), etc. are medium ncRNA (mncRNA) with a length ranging from 50 to 200 nucleotides. However, the largest class of RNA with maximum regulatory potency remains the lncRNAs (e.g., intronic, antisense, lincRNA, cisRNA, ceRNA, etc.) (Fig. 1) (Dahariya et al. 2019).

In this panorama, it has now been established that over a hundred chemical marks on the various RNAs control the activities of the transcripts and consequently regulates the *epitranscriptome*, that is to say, the biochemical modifications the RNA can undergo (Martino et al. 2014; Helm and Motorin 2017; Romano et al. 2018). In detail, various reversible and diverse chemical modifications of the nitrogen bases, that may arise post-transcriptionally in both coding and non-coding RNAs, have been discovered thanks to sequence databases and *transcriptome-wide* mapping. These modifications can affect the structure, metabolism, function, and structural stability of RNA, as well as protein recognition, resulting in an extension of the transcriptome variety (Bicchi et al. 2013; Morena et al. 2018).

Furthermore, these epigenetic modifications are stable, heritable, and they can change the gene expression and cellular function without modifying the original DNA sequence. In particular, they can generate a signature of gene expression that is held constant over several divisions of cells and reprogrammed under the influence of particular signals that control the tissue specification and its development (Morena et al. 2018). RNA molecules thus follow a specific program and perform a crucial role in biological and pathological processes, inasmuch as they can act as an actor or a target depending on the type of modification. The first RNA modification was detected in yeasts nearly 60 years ago and, after that, more than 150 modifications in RNAs have been described, including the canonical and non-canonical mRNA modifications and the main ncRNA (tRNAs, rRNAs, and regulatory RNAs) modifications, such as N6-methyladenosine (m^6A), N1-methyladenosine (m^1A), 5-methylcytosine (m^5C), pseudouridine (Ψ) and A-to-I editing (Table 1) (Roundtree et al. 2017; Morena et al. 2018).

However, to date, the studies that investigate the function and the biological effects of post-transcriptional modifications in the ncRNA types are still limited, apart from those concerning tRNA and rRNA (Dykes and Emanueli 2017; Yang et al. 2020).

These modifications are collectively possible thanks to the proteins involved in RNA chemical modifications that are known as RNA-modifying proteins (RMPs). The RMPs are classified into three distinct categories: *writers*, the enzymes that deposit RNA chemical marks; *erasers*, the enzymes that can remove them; and

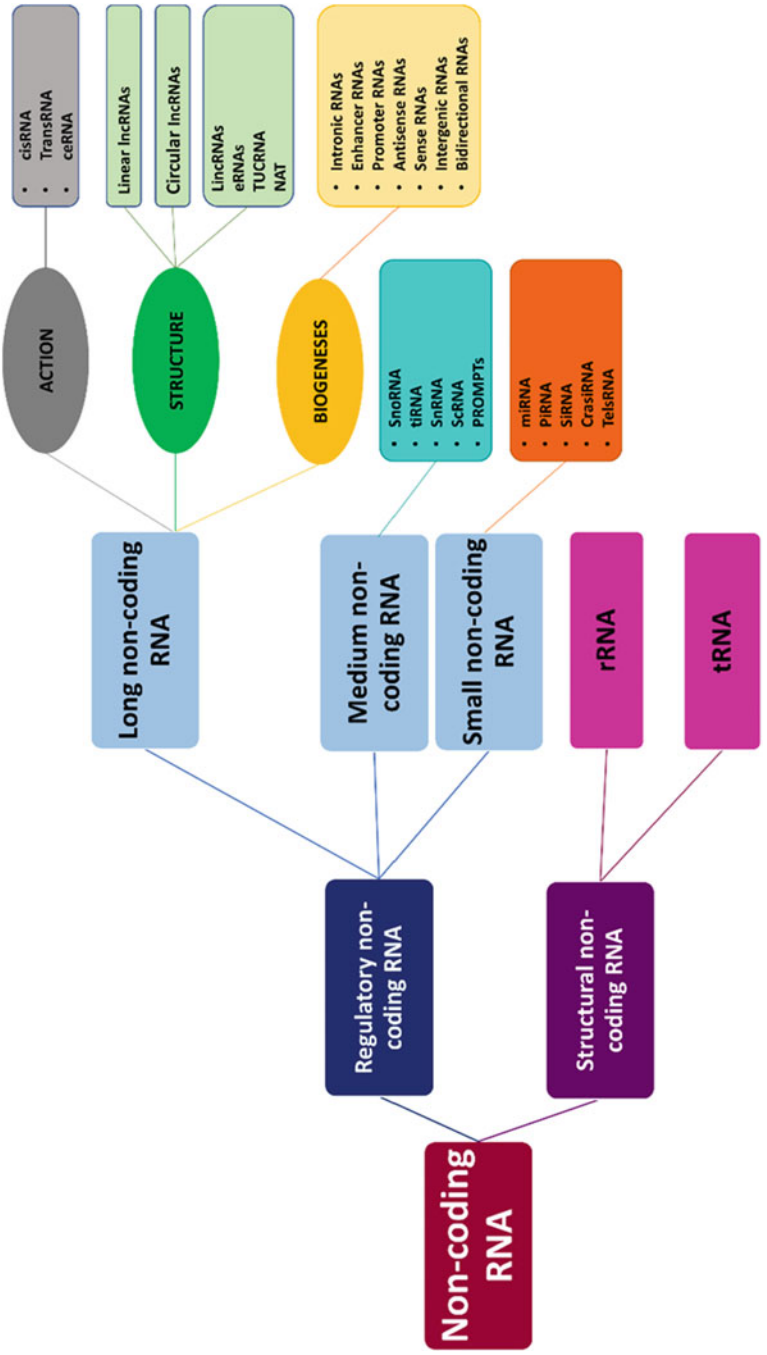


Fig. 1 Classification of non-coding RNAs

Table 1 Post-transcriptional modifications in coding and non-coding RNAs

RNA species	Functions of RNA species	Post-transcriptional modifications	Type of modifications and functions	References
mRNA	The mRNA transmits genetic information from DNA to the ribosome, where it is translated into proteins	Canonical: 5′-end	*cap-0* includes the addition of an N7-methylguanosine (m^7G) via a 5′–5′ triphosphate, which binds to the 5-′-terminal nucleotide of the pre-mRNA. The cap protects the mRNA from degradation by the activity of exonucleases. *cap-1 and cap-2* consist in a methylation ribose 2′-*O*-methylation on the first and second nucleosides to be transcribed. These modifications are fundamental for mRNA processing	Gagliardi and Dziembowski (2018), Morena et al. (2018), Song and Yi (2019)
		Canonical: 3′-end	*poly(A)-tail* contains at least 150–200 residues of adenylate and a conserved sequence of AAUAAA bounded by less conserved sequences. The mRNA stability is ensured during the phase of translation by this modification	Gagliardi and Dziembowski (2018), Morena et al. (2018), Song and Yi (2019)
		Canonical: splicing	It consists in the intron excision from pre-mRNAs by the splicing process obtained with the splicing machinery, which includes U1, U2, U4, U5, and U6 snRNPs. The splicing guarantees that pre-mRNA precursors enter into the post-transcriptional regulation of gene expression	Morena et al. (2018), Yan et al. (2019)

(continued)

Table 1 (continued)

RNA species	Functions of RNA species	Post-transcriptional modifications	Type of modifications and functions	References
		Non-Canonical:	It consists in the addition of a methyl group in the adenosine base at the nitrogen-6 position of the GAC sequence. The m^6A methylation of mRNA tends to be a process used by cells in response to different stress, such as nutrient deprivation, oxygen deficiency, exposure to high temperatures, or toxins and is involved in splicing, stability, and translation processes of mRNA	Morena et al. (2018), Liu et al. (2020b)
		Non-Canonical: N1-methyladenosine	It is characterized by a methyl group that is added to adenosine at the nitrogen-1 position. The role of m1A modification is not clear yet; first studies suggest that it is crucial in mRNA for blocking base pairing, which affects reverse transcription and protein translation	Xiong et al. (2018), Morena et al. (2018)
		Non-Canonical: 5-methylcytosine	It consists in a methylation at the position 5′ of cytosine that increases mRNA hydrophobicity and base stacking without affecting base pairing. Evidence suggests the interdependence between cytosine methylation levels and mRNA translation	Morena et al. (2018), Huang et al. (2019a)
		Non-Canonical: Pseudouridine	Ψ modification is characterized by the isomerization of uridine where the uracil is bound via a carbon–carbon bond instead of a nitrogen–carbon glycosidic bond. This type of modification improves the ability of mRNA to stack and makes the sugar-phosphate	Jacob et al. (2017), Morena et al. (2018)

			backbone more rigid. It could alter the efficiency of the translation initiation	
		Non-Canonical: Adenosine-to-Inosine editing	Adenosine-to-Inosine (A-to-I) editing, consisting of different base-pairing properties that are also observed in the combination Inosine-to-Cytosine. Adenosine is deaminated to produce inosine thanks to the activity of the ADAR enzyme (adenosine deaminase acting on RNA). This modification is critical for the editing and fate of mRNA	Fisher and Beal (2018), Morena et al. (2018)
		Non-Canonical: Cytidine-to-Uridine editing	Cytidine-to-Uridine (C-to-I) deamination is catalyzed by the APOBEC enzyme (apolipoprotein B mRNA editing catalytic polypeptide-like). This modification can alter the RNA base-pairing capacity and lead during the translation process to the recording of the message through the substitution of a different amino acid	Fisher and Beal (2018)
		Non-Canonical: 5-Hydroxymethylcytosine	The exact function remains unknown, but the m^5C might form hm^5C following oxidative processing. This modification is the most recurring one in exon sequences and increases the translation performance of mRNA	Lewis et al. (2017), Morena et al. (2018)
tRNA	The tRNA act as a translator for nucleotide language that will be converted into amino acid language	Anticodon loop	Inosine: modification found at three different positions in tRNAs (34, 37, and 57) and consists in a deamination reaction of adenines. Inosine found in tRNAs is essential for the correct translation of	Shafik et al. (2016), Asano et al. (2018), Morena et al. (2018)

(continued)

Table 1 (continued)

RNA species	Functions of RNA species	Post-transcriptional modifications	Type of modifications and functions	References
			the genetic code in wobble base pairs. Queuosine: is an analog of hypermodified guanosine, which comprises a 7-deaza-guanine core structure covalently linked to an amino-methyl side chain and a cyclo-pentanediol moiety. Queuosine improves the accuracy of translation compared to guanosine. 5-methylcytosine: methylation of carbon 5 in cytosine. m^5C could affect RNA stability in tRNA. 5-methoxycarbonylmethyl-2-thiouridine**:** uridine is transformed into the 5-methoxycarbonylmethyl-2-thiouridine by a three-step transformation. This nucleotide is critical for proper mRNA decoding and protein translation. Threonyl-carbamoyl-adenosine: modification located in the anticodon stem-loop at position 37 (t^6A37), adjacent to the anticodon of tRNAs that decodes ANN codons. t^6A37 contributes to translational efficiency and is essential for restricting initiation codons to AUG and limiting frameshifting at tandem of ANN codons. Wybutosine: it consists in a first modification at the guanosine residue and proceeds through 6 steps: N1 methylation of	

			guanosine, condensation of 1-methylguanosine with pyruvate, the addition of an aminocarboxypropyl group to obtain yW-86, methylation of yW-86 into yW-72, methylation of yW-72 to obtain yW-58, and finally methoxycarbonylation of yW-58 to obtain wybutosine. During codon recognition, Wybutosine and its derivatives are considered to play a crucial role. They support the codon-anticodon base pairing, thanks to the maintaining of the translational reading frame, and also prevent the formation of extended Watson–Crick base pairing. As a rule, all these anticodon loop modifications are required for the interaction of tRNA and mRNA within the ribosome	
		Adenosine-to-Inosine editing	It consists in the deamination mechanism of adenosine at positions 34, 37, and 57, and it is useful to improve the interactions of codon-anticodon in the ribosome	Rafels-Ybern et al. (2018), Morena et al. (2018)
		5-methyluridine	It consists in the uracil methylation by the enzyme hTRMT9 at position 5 (m^5U) and is identified and associated with an increase in decoding activity	Lentini et al. (2018), Morena et al. (2018)
		N1-methyladenosine	It consists in a methylation on the N1 atom of adenosine to form 1-methyladenosine in nucleotides 9, 14, 22, 57, and 58. It is associated with the correct folding and the increased of tertiary structure stability of tRNA	Oerum et al. (2017), Morena et al. (2018)

(continued)

Table 1 (continued)

RNA species	Functions of RNA species	Post-transcriptional modifications	Type of modifications and functions	References
		N1-methylguanosine	It is characterized by a methylation at the N-1 position of guanosine, and it is involved in preventing frameshifting and helping the tRNA with its accommodation in the ribosome during the translational elongation process	Kawarada et al. (2017), Pan (2018), Morena et al. (2018)
		N6-threonylcarbamoyl-adenosine	It consists in the formation of a threonylcarbamoyl group on adenosine, and it is found at position 37, 3′-adjacent to the anticodon in the tRNAs responsible for ANN codons. It may be involved in preventing frameshifting or aiding the tRNA while it accommodates in the ribosome during the translational elongation process	Kawarada et al. (2017), Pan (2018), Morena et al. (2018)
		N6-isopentenyladenosine	Modification that includes the isopentenylation of the adenosine at position 37, leading to the formation of isopentenyladenosine. This modification is involved in preventing framework changes or in assisting tRNA to accommodate in the ribosome during the translational elongation process	Kawarada et al. (2017), Pan (2018), Morena et al. (2018)
rRNA	rRNAs play a direct role in the creation of the peptic link between two amino acids in the ribosome	2′-*O*-methylation (2′-O-Me)	It consists in the addition of a methyl group at the 20 positions with the modification of one-two specific nucleotides. It is involved in ribosome assembly, in the protection of the rRNA from	Sergiev et al. (2018), Morena et al. (2018)

			hydrolysis, and in the modification of the rRNA strand flexibility	
		Uridine-to-Pseudouridine isomerization	It consists in the isomerization of uridines to pseudouridines (Ψs) in rRNA: this reaction is better known as pseudouridylation. Compared to uridine, this modification improves the rigidity of the phosphate-sugar backbone, the stacking of local bases, and the stability of pseudouridylation duplexes. Specifically, it is implicated in the folding, processing, and assembly of rRNA in ribosomes guaranteeing the maintenance of advanced structures	Vandivier and Gregory (2017), Morena et al. (2018), Meng et al. (2019)
lncRNA	lncRNA plays a crucial role in different biological functions, including post-transcriptional regulations and chromatin modifications	N6-methyladenosine	It consists in the addition of a methyl group to the adenosine base at the nitrogen-6 position of the GAC sequence. In this case, the modification can generate the segment 78 m^6A residues, playing a crucial role in regulating the lncRNA X-inactive specific transcript (XIST) activity and in silencing genes in the X chromosome. Moreover, several studies demonstrated that the m^6A is involved in splicing, processing, and biogenesis of lncRNA	Bohnsack and Sloan (2018), Dai et al. (2018), Morena et al. (2018)
		N1-methyladenosine	It consists in the addition of a methyl group to the N1 position of adenosine and in lncRNAs. Its role has not been extensively explored, even if it seems to maintain biological functions	Zhang and Jia (2018), Morena et al. (2018)

(continued)

Table 1 (continued)

RNA species	Functions of RNA species	Post-transcriptional modifications	Type of modifications and functions	References
		Pseudouridine	The Ψ modification is characterized by the isomerization of uridine with the uracil that is bound via a carbon–carbon bond instead of a nitrogen–carbon glycosidic bond. The role of Ψ in lncRNA is still unknown, but it seems to serve a function in the activation of the steroid receptor RNA activator (SRA) and to carry several implications in different diseases	Shafik et al. (2016), Zhao et al. (2018), Morena et al. (2018)
		5-methylcytosine	It consists in the methylation at the position 5′ of cytosine. This modification can alter the structure of RNA and affect the binding of lncRNA with proteins	Morena et al. (2018), Liu et al. (2020a)
miRNA	The miRNA operates at the post-transcriptional level to silencing and regulating of gene expression	Pseudouridine	The Ψ modification is characterized by the isomerization of uridine with the uracil that is bound via a carbon–carbon bond instead of a nitrogen–carbon glycosidic bond. Several studies demonstrated that the incorporation of this modification tends to decrease the regulation of miRNA switches	Ge and Yu (2013), Morena et al. (2018), Lockhart et al. (2019)
		5-methylcytosine	It consists in the methylation at the position 5′ of cytosine. This modification seems to be involved in the miRNA biogenesis	Morena et al. (2018), Trixl and Lusser (2019)

		N6-methyladenosine	It consists in the addition of a methyl group in the adenosine base at the nitrogen-6 position of the GAC sequence. This modification is involved in the life cycle of RNA metabolism and in particular in miRNA processing and biogenesis	Morena et al. (2018), Zhu et al. (2020)
		Adenosine-to-Inosine editing	Adenosine-to-Inosine (A-to-I) editing consists of different base-pairing properties that are also observed in the combination Inosine-to-Cytosine. This modification can affect miRNA biosynthesis and target recognition	Xu et al. (2018), Morena et al. (2018)
circRNA	The circRNA acts as a miRNA sponge, thus playing a significant role in gene expression regulation	N6-methyladenosine	It consists in the addition of a methyl group in the adenosine base at the nitrogen-6 position of the GAC sequence. This modification may happen during or after the formation of circRNAs, and it can work as a switch to control their functionality	Zhou et al. (2017), Morena et al. (2018)

readers, the proteins that can selectively recognize and bind to specific RNA chemical modifications (Esteve-Puig et al. 2020).

As of today, the available data highlight the role of RNA modifications in stem cell self-renewal, commitment, and differentiation processes; while ncRNA dysregulation, their modification, and mutations in RMPs have been related to human disorders including neurological diseases, cancer, obesity, and infertility (Bicchi et al. 2013; Mongelli et al. 2020; Barbieri and Kouzarides 2020).

2 Databases and Web Servers of ncRNA

For years, research on RNA modifications has mainly focused on tRNAs, because of their relative abundance and small size; and later, when technological advances in sequencing methodology were finally made, attention has turned to ribosomal rRNA as well (Torres et al. 2014; Lafontaine 2015; Kirchner and Ignatova 2015). However, only in the last couple of years, after the emergence of next-generation sequencing (NGS) technology, it has been feasible to shift the scope of research towards *transcriptome-wide* modification studies.

Today it has become evident that RNA modifications are more prevalent and chemically diverse than their DNA counterparts (Boccaletto et al. 2018). They are highly dynamic, and at least some are reversible, which makes them a critical component of the post-transcriptional gene regulatory landscape. Hence, it seems perfectly clear that RNA modifications and alterations of the RNA modification machinery can have detrimental effects on human diseases (Jonkhout et al. 2017).

Given the important functionality of ncRNAs and that their dysregulation is associated with a large number of pathologies, such as cancer, cardiovascular diseases, and neurodegenerative diseases (Esteller 2011), a database for ncRNA–disease association is an important tool for researchers (Orlacchio et al. 2008; Bicchi et al. 2013). In fact, ncRNA–disease association databases can prove useful to either biomedical scientists investigating the roles of these RNAs under pathological conditions or bioinformatics scientists eager to discover patterns of ncRNAs in such diseases and to develop novel ncRNA–disease association prediction algorithms.

Here, we provide a state-of-the-art summary of the principal databases and web servers of ncRNA (Table 2). We also include databases with genetic variants and principal chemical editing that can affect RNA modifications and functions.

3 Epitranscriptome in the Central Nervous System

It has long been known that epigenetic modifications can contribute to and influence changes in gene expression and are at the basis of synaptic plasticity and the development of diseases affecting the Central Nervous System (CNS). In recent

Table 2 Principal databases and web servers of ncRNA

Name	Type	Year of the first version	Objective	Link
HMDD v3.0	Database	2008	Experimentally supported human miRNA–disease associations	http://www.cuilab.cn/hmdd
miRwayDB	Database	2018	Information of experimentally validated microRNA-pathway associations in various pathophysiological conditions	http://www.mirway.iitkgp.ac.in
RMDisease	Database	2020	Genetic variants that affect RNA modifications, with implications for epitranscriptome pathogenesis	https://www.xjtlu.edu.cn/en/study/departments/academic-departments/biological-sciences/rmdisease
miRNASNP-v3	Database	2012	SNPs and disease-related variations in miRNAs and miRNA targets	http://bioinfo.life.hust.edu.cn/miRNASNP
LncRNADisease	Database	2012	lncRNA-associated diseases	http://cmbi.bjmu.edu.cn/lncrnadisease
LncRNADisease v2.0	Database	2019	lncRNA-associated diseases	http://www.rnanut.net/lncrnadisease/index.php/home
Lnc2Meth	Database	2017	Regulatory relationships between long non-coding RNAs and DNA methylation associated with human disease	http://www.bio-bigdata.com/Lnc2Meth/
CAMi-Finder	Database	2003	Human cancer miRNA associations	http://www.isical.ac.in/~bioinfo_miu/web_isi.html
dbDEMC	Database	2010	Differentially expressed miRNAs in cancers	http://www.oncomir.umn.edu/
mirdsnp	Web server	2012	Disease-associated SNPs and miRNA target sites on 3′UTRs of mRNAs	http://mirdsnp.ccr.buffalo.edu
miRo v 2.0	Web server	2009	Associations between miRNAs and diseases	http://microrna.osumc.edu/miro/
miRSel	Web server	2010	Associations between microRNAs and genes from the biomedical literature	https://services.bio.ifi.lmu.de:1047/mirsel/
miRWalk 2.0	Web server	2011	miRNA and disease associations	http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/
MMIA	Web server	2009	Linking miRNA functions to diseases	http://147.46.15.115/MMIA/index.html

(continued)

Table 2 (continued)

Name	Type	Year of the first version	Objective	Link
PhenomiR	Web server	2010	miRNA and disease relations	http://mips.helmholtz-muenchen.de/phenomir/
PolymiRTS 3	Web server	2014	Linking miRNA polymorphism and diseases	http://compbio.uthsc.edu/miRSNP/
SomamiR DB 2.0	Web server	2013	Somatic mutations and microRNA targeting in cancer	http://compbio.uthsc.edu/SomamiR/
miRcancer	Database	2013	MicroRNA cancer association	http://mircancer.ecu.edu/
Starbase v2.0 ENCORI	Web server	2011	Deciphering Pan-Cancer Networks of lncRNAs, miRNAs, ceRNAs, and RNA-binding proteins	http://starbase.sysu.edu.cn
YM500v2	Database	2013	Cancer miRNAome	http://ngs.ym.edu.tw/ym500v2/index.php
EpimiRBase	Database	2016	Information on up- and downregulated microRNAs in the brain and blood	http://www.epimirbase.eu/
ExcellmiRDB	Web server	2015	Extracellular miRNAs	http://www.excellmirdb.brfjaisalmer.com/
miRandola	Database	2012	Circulating non-coding RNAs in diseases	http://mirandola.iit.cnr.it/index.php
MirSNP	Database	2012	Polymorphisms altering miRNA target sites	http://bioinfo.bjmu.edu.cn/mirsnp/search/
Mirsnpscore	Database	2011	SNP effects on microRNA targeting	http://www.bigr.medisin.ntnu.no/mirsnpscore/
multiMiR	Database	2014	miRNA and disease relations	http://multimir.ucdenver.edu
Patrocles	Database	2009	Linking polymorphisms in miRNA target with human disease	http://www.patrocles.org/
SurvMicro	Database	2014	Expression of miRNAs in cancers	http://bioinformatica.mty.itesm.mx:8080/Biomatec/Survmicro.jsp
MODOMICS	Database	2013	A database for RNA modifications	http://modomics.genesilico.pl
RNAMDB	Database	2011	A database for RNA modifications	http://rna-mdb.cas.albany.edu/RNAmods/
RADAR	Database	2014	Collection of A-to-I RNA editing sites with annotation	http://RNAedit.com

(continued)

Table 2 (continued)

Name	Type	Year of the first version	Objective	Link
MeT-DB	Database	2015	A database for publicly available m^6A data sets	http://compgenomics.utsa.edu/methylation/
REDIportal	Database	2016	Collection of A-to-I RNA editing sites with annotation	http://srv00.recas.ba.infn.it/atlas/
HAMR	Web server	2013	Predict RNA modification site (location and methylation class)	http://wanglab.pcbi.upenn.edu/hamr
iRNA-Methyl	Web server	2015	To predict m^6A site	http://lin.uestc.edu.cn/server/iRNA-Methyl
iRNA-PseU	Web server	2016	To predict W site	http://lin.uestc.edu.cn/server/iRNA-PseU
iRNA-PseColl	Web server	2017	To predict occurrence sites of RNA modification	http://lin.uestc.edu.cn/server/iRNA-PseColl
RNAMethPre	Web server	2016	To predict m^6A site	http://bioinfo.tsinghua.edu.cn/RNAMethPre/index.html

years, however, it has been understood that the *epitranscriptome* also plays a pivotal regulatory role during the development and onset of diseases involving the CNS (Flamand and Meyer 2019). To date, more than 160 modifications have been discovered that can affect RNAs and which in turn can affect many aspects of nervous development, synaptic plasticity, and the outbreak of neurological diseases (Hsu et al. 2017; Angelova et al. 2018). Therefore, studying these mechanisms and understanding how the *epitranscriptome* operates on CNS is a very topical line of research. About 98% of transcribed RNAs are non-coding RNAs (ncRNAs) but capable of performing a particularly important function in gene expression programming (Xylaki et al. 2019; Yan et al. 2019). As mentioned before, these non-coding RNAs are usually classified depending on their length in sncRNAs (which among others include miRNAs and piRNAs), mncRNAs, and lncRNAs. Recently, circular long non-coding RNAs (circRNAs) have also joined this family (Neueder 2019).

It has been shown that lncRNAs, circRNAs, and miRNAs are particularly expressed in the CNS and correlated with its correct functionality, or else with the onset of different pathologies (Briggs et al. 2015; Xylaki et al. 2019). The presence of aberrant tRNAs, rRNAs, and mRNAs has also been correlated with CNS dysfunctions and with the outbreak of pathologies affecting the brain (Angelova et al. 2018). Through modern strategies based on the use of high-throughput sequencing, it has been possible to map the main modifications affecting the transcriptome and, in particular, mRNAs and lncRNAs (Jung and Goldman 2018). As an example, the methylation of some mRNAs, which code for key proteins of synaptic plasticity, has proved to be associated with neurodevelopmental disorders such as Autism

(Flamand and Meyer 2019). Understanding the mechanisms by which these RNAs act is extremely important in order to disclose their involvement in the development of pathological phenomena and to be able to use them as possible biomarkers of the disease and as targets to develop new treatments for pathologies affecting the CNS (Neueder 2019). In fact, in light of the numerous modifications that can affect the transcriptome, RNA is now seen as a potential pharmacological target for the treatment of some neurodegenerative diseases. In this respect, a viable strategy involves the use of small molecules interacting with RNA (SMIRNAs), whose capability of binding RNA can, for example, be detected by changes in their fluorescence emission (Botti et al. 2019, 2020; Ursu et al. 2019).

The main and best-known modifications to the transcriptome are those reported above in Table 1. As described in recent studies, these modifications are catalyzed by molecular enzymatic machinery generally known as *writers*, *erasers*, and *readers* (Jung and Goldman 2018; Morena et al. 2018; Livneh et al. 2020). In the case of Pseudouridine (Ψ), an isomer of uridine that is particularly abundant in tRNAs, mRNAs, and some snRNAs, it has been shown how mutations that inactivate the enzyme Pseudouridylase 3 (Pus3), which catalyzes the isomerization of uridine-to-Pseudouridine (Ψ), are associated with intellectual disorders (Jung and Goldman 2018). Queuosine (Q), a hypermodified 7-deazaguanosine nucleoside, is an adduct of many tRNAs, and it has been found to influence the production in mouse models of neurotransmitters such as dopamine and to contribute to the development of encephalomyelitis in Multiple Sclerosis (Varghese et al. 2017; Jung and Goldman 2018). More generally, all these types of mutations can be indiscriminately present in both mRNAs and ncRNAs, and specifically in lncRNAs.

The lncRNAs perform an extremely important action in the brain as they are associated with cell differentiation, synaptogenesis, and other vital functions (Quan et al. 2017; Xylaki et al. 2019). For this reason, mutations or dysregulation of these lncRNAs is closely connected with the onset of neurodegenerative diseases (Fig. 2), such as Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Lobe Degeneration (FTLD) (Quan et al. 2017; Wei et al. 2018; Argentati et al. 2020).

The modifications affecting RNA can also be caused by oxidative stress mostly induced by Reactive Oxygen Species (ROSs). As with DNA, even in the case of RNA, the base which is most affected by this type of modification is guanosine. 8-hydroxyguanosine (8-OHG) is in fact one of the main biomarkers of RNA oxidation induced by the hydroxyl radical (•OH). The brain appears to be the organ most susceptible to oxidative damage, as it features the highest oxygen consumption but the least presence of antioxidant enzymes (Liu et al. 2020c). Oxidative damage affecting the transcriptome is closely correlated with neuronal fragility, and it too can promote the onset of neurodegenerative diseases such as AD and PD (Essack et al. 2020; Kilchert et al. 2020; Liu et al. 2020c). Indeed, it has been demonstrated that an increase in oxidized RNA is present both in AD patients and in PD and ALS patients (Liu et al. 2020c).

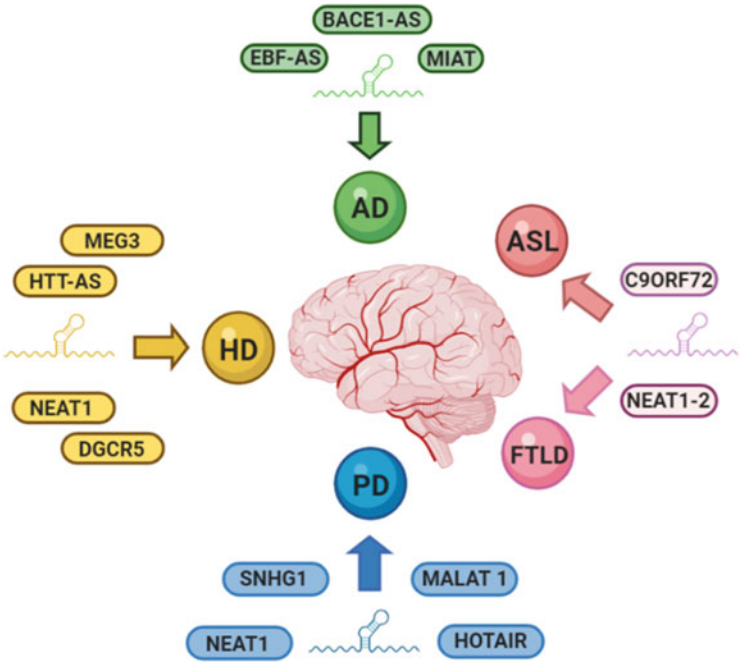


Fig. 2 Dysregulation of lncRNAs associated with neurodegenerative diseases. The main lncRNAs dysregulated in Alzheimer's disease (AD) are reported in green, those in Parkinson's disease (PD) in blue, those in Huntington's disease (HD) in yellow, and those in Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Lobe Degeneration (FTLD) in purple

In addition to neurodegenerative diseases, it has emerged that transcriptome modifications are also critical regulators in carcinogenesis, particularly in the case of cancers affecting the brain (Huang et al. 2020a).

3.1 Epitranscriptome and Brain Cancer

In recent years, the modifications involving RNA are attracting the attention of researchers inasmuch as they appear to serve a crucial role in tumorigenesis. In particular, some post-transcriptional modifications of RNA are particularly correlated with the onset of Glioblastoma (GBM), one of the most aggressive cancers affecting the brain, and due to which life expectancy, to date, does not exceed 15 months from diagnosis (Huang et al. 2020a; Dong and Cui 2020). However, the role of *epitranscriptome* in the onset of brain gliomas has only begun to be investigated in recent years thanks to the advent of Next-Generation Sequencing (NGS), which shed light on the possible mechanisms connected to the development and progression of brain tumors, as well as other neurological disorders (Delaidelli

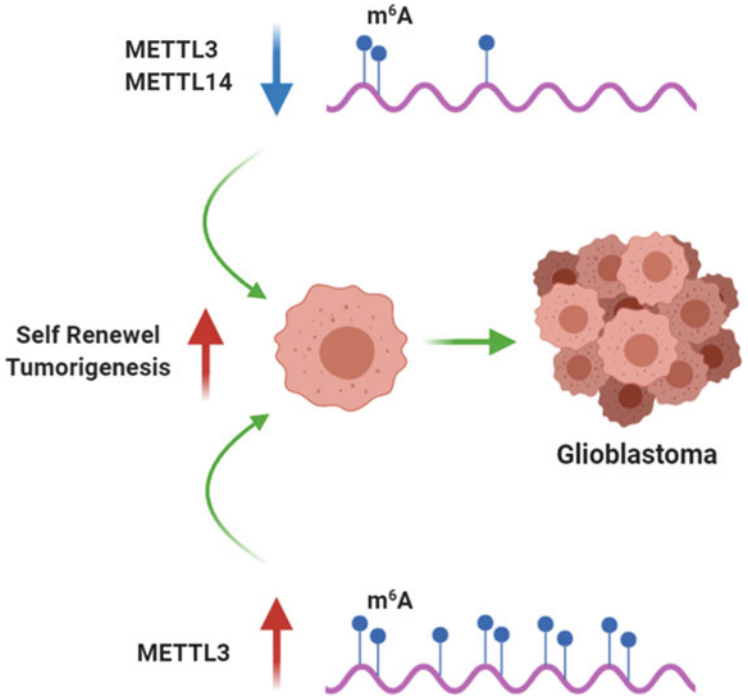


Fig. 3 Effect of the increase or decrease of m^6A in GBM. The knockdown of METTL3 and METTL14 reduces the m^6A levels inducing GBM stem cells' (GSCs) proliferation and their self-renewal capacity (Cui et al. 2017). The overexpression of METTL3 induces m^6A RNA methylation and the capacity of cancer cells' self-renewal and tumorigenesis (Visvanathan et al. 2018)

et al. 2019; Huang et al. 2020a). N6-methyladenosine (m^6A) is one of the main modifications affecting mRNA and whose deposition is catalyzed by a molecular complex called RNA methyltransferase complex (MTC), a heterodimer consisting of methyltransferase-like 3 (METTL3) and methyltransferase-like 14 (METTL14). This type of modification, correlated with the overexpression or depletion of METTL3 and METTL14, appears to be one of the most important alterations found in GBM (Galardi et al. 2020). To date, however, there are conflicting opinions about the role of N6-methyladenosine: while some researchers argue that a reduction of m^6A levels is protective, others claim the opposite (Fig. 3). Interestingly, the involvement of m^6A methylation of RNA and the related proteins in the development of GBM has recently been described (Cui et al. 2017; Galardi et al. 2020). The knockdown of METTL3 and METTL14 has been found to reduce the general methylation of m^6A, increase the proliferation of GBM stem cells (GSCs), and the capacity for self-renewal. According to these studies, there is therefore an inverse correlation between the levels of m^6A RNA and the tumorigenicity of GBM; data confirmed by in vivo studies in which the knockdown of either METTL3 or

METTL14, or both, has been shown to reduce the levels of m^6A RNA and increase the proliferation of GBM (Cui et al. 2017; Galardi et al. 2020). Conversely, other researchers argue that an increase in m^6A RNA methylation fulfills a major function in the development and progression of GBM (Visvanathan et al. 2018; Galardi et al. 2020). In particular, overexpression of METTL3 induces m^6A RNA methylation and the capacity of self-renewal of cancer cells and tumorigenesis (Visvanathan et al. 2018). The conflicting results obtained can partly be offset by the high heterogeneity that characterizes this type of brain tumor, which is also reflected in the different cellular models used in these studies (Cui et al. 2017; Galardi et al. 2020).

Another type of modification that has recently been found in cerebral glioma is 5-methylcytidine (m^5C), which is found not only in mRNAs but also in tRNAs and lncRNAs (Huang et al. 2020a; Galardi et al. 2020). This type of modification is introduced by the action of the RNA cytosine methyltransferases (RCMTs), in particular from the RCMT subgroup NOL1/NOP2/SUN (NSUN), and it has especially been detected in the 28S rRNA of glioma (Janin et al. 2019; Trixl and Lusser 2019; Galardi et al. 2020). By silencing the promoter of a member of the NSUN family, i.e., the NSUN5 gene, and thereby preventing m^5C methylation, tumor suppression has been demonstrated, both in vitro and in vivo (Janin et al. 2019; Trixl and Lusser 2019). There are also other post-transcriptional modifications of RNA that are involved in the pathogenesis of brain tumors and that are still under study, as in the case of alternative polyadenylation, alternative splicing, and the Adenosine-to-Inosine editing (Huang et al. 2020a). Many studies have also shown that the dysregulation of some ncRNAs, such as lncRNAs and miRNAs, is correlated with the development and progression of GBM. On the one hand, some miRNAs are upregulated in GBM and are functionally classified as oncomiR, inasmuch as they have an oncogenic and antiapoptotic effect; on the other hand, some oncosuppressive and pro-apoptotic miRNAs are downregulated (Banelli et al. 2017). Table 3 lists the main deregulated miRNAs with an oncogenic or oncosuppressive effect, together with other ncRNAs involved in the onset of this pathology. Figure 4 shows the entire network of dysregulated miRNAs involved in the onset and progression of GBM.

The RNA modifications and their regulation play therefore a fundamental function in carcinogenesis; however, the high heterogeneity of brain tumors has highlighted a great multiplicity and plasticity of the mechanisms involved. Further and deeper studies are thus needed to better understand the functioning of these molecular mechanisms, so that they can be used as a target for brain cancer treatment.

3.2 Epitranscriptome and Alzheimer's Disease

Alzheimer's disease (AD) is a complex progressive neurodegenerative disorder, which causes dementia, clinically characterized by memory and cognitive

Table 3 Principal lncRNAs and miRNAs dysregulated in GBM, type of their deregulation, and their mechanism in GBM

Name	Type of RNA	Deregulation in GBM	Mechanism in GBM	References
hsa-miR-21	miRNA	Upregulation	The upregulation of miR-21 promotes GBM tumorigenesis by downregulating the insulin-like growth factor-binding protein-3 (IGFBP3)	Yang et al. (2014)
hsa-miR-221/222	miRNA	Upregulation	The upregulation of miR-221/222 inhibits cell apoptosis by targeting the pro-apoptotic gene p53 upregulated modulator of apoptosis (PUMA) in human glioma cells. Moreover, miR-222 and -221 induce an increase in cell migration by targeting the protein phosphate PTP mu	Zhang et al. (2010), Quintavalle et al. (2012)
hsa-miR-335	miRNA	Upregulation	The upregulation of miR-335 promotes tumorigenic features, such as growth and invasion of GBM	Shu et al. (2011)
hsa-miR-218	miRNA	Downregulation	The downregulation of miR-218 promotes GBM invasion/migration, proliferation, apoptosis, and stemness by targeting different genes such as the E2F2 gene involved in the cell cycle	Zhang et al. (2015b)
hsa-miR-451	miRNA	Downregulation	The downregulation of miR-451 promotes GBM cell growth and invasive ability	Nan et al. (2010)
MALAT1	lncRNA	Downregulation	The downregulation of MALAT1 (Metastasis-Associated Lung Adenocarcinoma Transcript 1) promotes migration in GBM cells	Vassallo et al. (2016)
HOTAIR	lncRNA	Upregulation	The upregulation of HOTAIR promotes the cell cycle progression in glioma as a result of the binding of its 5′ domain to the PRC2 complex	Shen et al. (2018)
CDR1-AS	circRNA	Downregulation	The axis miR-671-5p/CDR1-AS/CDR1/VSNL1 is functionally altered in GBM cells and is involved in the modification of their biopathological profile	Barbagallo et al. (2016)
circSMARCA5	circRNA	Downregulation	The downregulation of CircSMARCA5 promotes migration of GBM	Barbagallo et al. (2018)

(continued)

Table 3 (continued)

Name	Type of RNA	Deregulation in GBM	Mechanism in GBM	References
TUG1	lncRNA	Upregulation	The upregulation of TUG1 enhances tumor-induced angiogenesis by inhibiting microRNA-299 in human GBM	Cai et al. (2017)
LINC01116	lncRNA	Upregulation	The upregulation of LINC01116 promotes tumor cells' growth and the stemness and radiation response of glioma stem cells	Brodie et al. (2017)
SNHG7	lncRNA	Upregulation	The upregulation of SNHG7 promotes the progression and growth of GBM via inhibition of miR-5095	Ren et al. (2018)
lncHERG	lncRNA	Upregulation	The upregulation of lncHERG promotes cell proliferation, migration, and invasion in GBM	Shi et al. (2017)
H19	lncRNA	Upregulation	The upregulation of H19 promotes invasion, angiogenesis, and stemness of GBM cells	Jiang et al. (2016b)

dysfunctions (Orlacchio et al. 2007; Urbanelli et al. 2008; Magini et al. 2015; Long and Holtzman 2019; Ceyzériat et al. 2020; Argentati et al. 2020).

Primary hallmarks of AD are well-known, and the disease has a specific neuropathological profile: extracellular neuritic plaques, mainly consisting of misfolded and aggregated amyloid-β (Aβ) peptides, and neurofibrillary tangles (NFTs), primarily containing highly phosphorylated Tau protein (microtubule-associated protein tau, MAPT), occur within neurons; whereas gliosis, neuroinflammation, and synaptic loss are also evident in the hippocampi and brain cortices of affected individuals (Alzheimer's Association 2020). Broadly speaking, AD can be either sporadic or familial and have either a late- or early-life onset. Sporadic AD (SAD) in most cases (90–95%) begins after age 65 (late-onset AD, LOAD), and it is thus typical of the elderly. Rarely, AD may onset at a young age (<65), being referred to as early-onset AD (EOAD) (Chen and Mobley 2019; Penke et al. 2020). In some EOAD cases, there is a family history of dementia, allowing it to be defined as Familial AD (FAD). FAD cases are typically caused by dominant mutations in three genes: the amyloid precursor protein (APP) gene, and the Presenilin (PSEN) genes, PSEN1 and PSEN2, whose products participate in processing APP (Dorszewska et al. 2016).

SAD has more complex pathogenesis than FAD, as it may have different potential causes not yet fully understood. During the past decade, clinical and experimental studies have identified many genetic and non-genetic risk factors for SAD (Dorszewska et al. 2016; Shao et al. 2017) via genome-wide association studies (GWAS), Whole Exome Sequencing (WES), and Whole Genome Sequencing

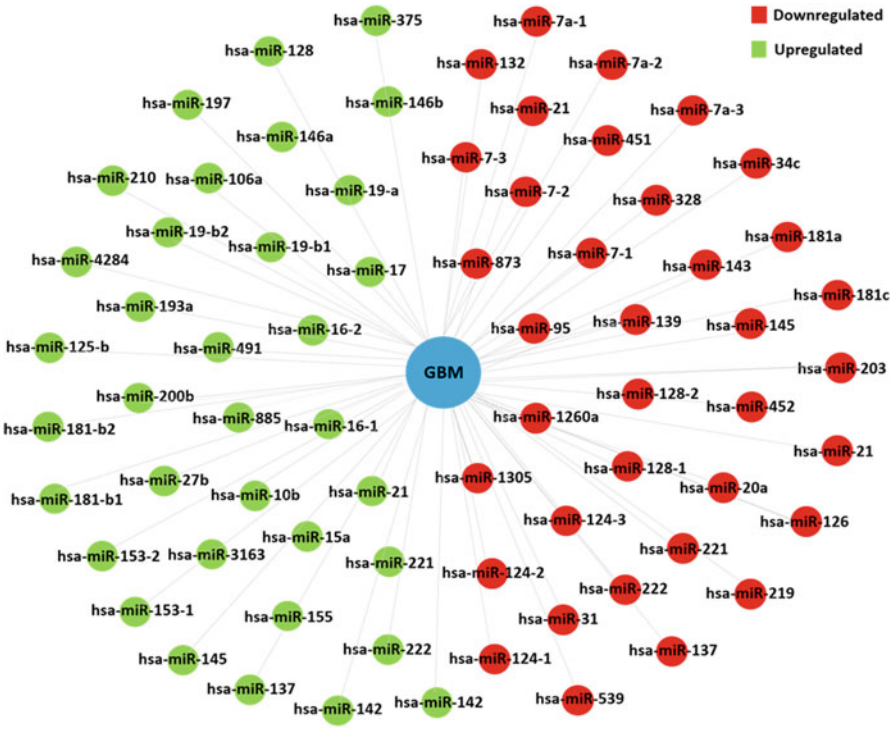


Fig. 4 Network of the miRNA dysregulation in GBM. Upregulated miRNAs are reported in green, and downregulated miRNAs are reported in red

(WGS). In particular, multiple variants have been discovered to pose an increased risk of AD (Chen and Mobley 2019). One of these factors is the ε4 allele of APOE, as harboring one or two ε4 alleles is also linked to an earlier age of AD onset. On the contrary, the ε2 allele of APOE is associated with decreased risk of AD as well as with a later age of onset (van der Lee et al. 2018; Armstrong 2019).

Together with a multiplicity of genetic factors (Bertram and Tanzi 2019), other factors linked to the etiology of AD include age, gender, vascular disorders, stroke, blood–brain barrier, diabetes, infection, inflammation, synaptic, mitochondrial and innate immune system dysfunction, dysfunction of the meningeal lymphatic system, sleep deprivation, chronic alterations in circadian rhythm, gut microbiome, hearing loss, protein misfolding and processing, as well as environmental factors that range from pollution to diet and from exercise to educational and socioeconomic status (Morena et al. 2018; Guo et al. 2020). Since AD is not ascribable to a single pathogenesis factor, the molecular mechanism underlying the disease still remains unclear.

Moreover, recent studies have shown alterations in the splicing programs in AD patients, contributing to the complexity of AD pathogenesis (Han et al. 2019; Braggin et al. 2019; Biamonti et al. 2019). In fact, although little is known about

AD at the molecular level, alternative splicing of the APP, TAU, APOE4, or PSEN1-2 genes may contribute to the pathology (Han et al. 2019). Moreover, a recent study showed that exposing primary neurons to nanomolar amounts of Aβ increases the translation of specific CYFIP2/FMRP-regulated mRNA. Furthermore, the reduction of endogenous CYFIP2 expression in the early model of AD leads to the development of key features of AD in mice with aging (Ghosh et al. 2020). It has been demonstrated that the overexpression of an activator of cdk5, namely p25, causes increased levels of BACE1 mRNA and protein, both in vitro and in vivo, which in turn increase the amyloidogenic processing of APP (Wen et al. 2008).

Transcriptomics analysis is essential in studying complex diseases, such as neurodegenerative disorders (Bagyinszky et al. 2020). A recent study from Wan et al. has developed an atlas of the human brain transcriptome in AD, based on 2114 postmortem samples, highlighting transcriptional networks altered by human brain pathophysiology and also identifying correspondences with mouse models for AD preclinical studies (Wan et al. 2020).

To date, many studies have been focusing on the potential roles of epigenetics in AD pathogenesis; in fact, several genetic and epigenetic factors and gene-environmental interactions could be involved in the disease onset (Bagyinszky et al. 2020). *Epitranscriptomics* allowed RNA modifications to be studied and N6-methyladenosine (m^6A), the most abundant modification of mRNA among these, was identified as a leading player in the reliability of brain functions and the nervous system (Flamand and Meyer 2019; Livneh et al. 2020).

Some studies have correlated the m^6A modification with neurodegenerative diseases (Li et al. 2018b) and AD in particular (Han et al. 2020; Huang et al. 2020b). A recent study has shown that the m^6A methylation was augmented in the cortex and hippocampus of APP/PS1 transgenic mice relative to control mice. Interestingly, the expression of the m^6A methyltransferase METTL3 was found by means of genome-wide maps of the m^6A mRNA to enhance in AD mice, while that of the m^6A demethylase FTO was found to decrease. These data demonstrate that the m^6A methylation of RNA promotes the development of AD (Han et al. 2020).

Further studies have also shown that both through genetic and epigenetic mechanisms, cerebral hypoxia increases Aβ deposition by altering the expression levels of the enzymes involved in the production/degradation of the protein (Jakubauskienė et al. 2020).

In addition, there are a series of non-coding RNAs, including miRNAs, lncRNAs, and circRNAs, that may be strongly implicated in AD pathogenesis, being involved in the formation and development of Aβ-amyloid plaques and NFTs, synaptic loss, and neuronal death (Millan 2017). For example, some ncRNAs regulate Aβ accumulation in AD models, inhibiting the clearance of secretase or modulating its cleavage (Wu and Kuo 2020). In another study carried out on the hippocampal tissue from Aβ1-42-induced AD model rats, a total of 555 circRNAs, 183 miRNAs, and 319 mRNAs, identified by microarray analysis technology, turned out to be significantly dysregulated (fold-change ≥ 2.0 and p-value < 0.05), suggesting the different ncRNA expression patterns as a cause of AD pathogenesis (Wang et al. 2018b).

As of today, the involvement of miRNAs in the pathogenesis of AD has been explored, and their expression may be either up-or downregulated in AD (Juźwik et al. 2019; Bagyinszky et al. 2020). In particular, miRNAs may be associated with AD through APP processing, amyloid formation, and Tau phosphorylation. For example, miR-346 can upregulate APP translation and Aβ production, thus causing AD pathogenesis (Millan 2017; Long et al. 2019). Moreover, miR-339-5p, miR-29c, miR-15b, miR-195, and miR-124, are involved in the Aβ metabolism by modulating the activity of β-secretases, such as BACE1 (Selkoe and Hardy 2016; Das et al. 2016). Not only can miRNA both affect the Tau protein synthesis and have a direct effect on MAPT (Santa-Maria et al. 2015), but it can also affect Tau phosphorylation via the regulation of specific enzymes. For example, miR-124-3p was shown to reduce the abnormal Tau phosphorylation (Zhou et al. 2019). Moreover, miR-219-5p downregulates GSK3 to inhibit Tau phosphorylation in AD (Li et al. 2019). miRNAs can also regulate synaptic plasticity, participate in neuronal growth, differentiation and apoptosis, and mediate immuno-inflammatory responses in AD (Wei et al. 2020). As an example, miR-142a-5p, miR-146a-5p, miR-155-5p, and miR-455-5p are upregulated in AD brains and are involved in neuronal functions, supposedly having a role in brain development and neurodegeneration (Arena et al. 2017; Sierksma et al. 2018). Interestingly enough, miRNAs can play a pivotal role in AD brains by either promoting apoptosis or inhibiting it. For example, miR-146a inhibits LRP2 translation, thus inducing cell apoptosis (Zhang et al. 2016); whereas, miR-98 reduces Aβ production, inhibits the Notch signaling pathway, and suppresses the apoptosis of hippocampal neurons, thereby promoting their survival (Chen et al. 2019a).

The miRNAs associated with human AD disease, according to experiment-supported evidence curated by the HMDD database (Human microRNA Disease Database), are reported in Table 4.

A recent study has shown that, among the 8098 miRNAs individually measured in blood cells, six of them, i.e., miR-107, miR-125b, miR-146a, miR-181c, miR-29b, and miR-342, were significantly downregulated in individuals with AD compared to controls (Fransquet and Ryan 2018). These results suggest that miRNAs could also be potential biomarkers for AD (Zhang et al. 2019; Swarbrick et al. 2019). In this regard, it has been confirmed that AD patients are characterized by alterations of miRNA in the brain and biological fluids, including serum, plasma, and cerebrospinal fluid (CSF) (Zendjabil 2018; Zetterberg and Burnham 2019). It has been hypothesized that miRNAs are transported within liposomes, HDLs, exosomes, and other proteins protecting the miRNA from degradation (van den Berg et al. 2020). A panel of 12 miRNAs able to discriminate AD from controls was first reported in 2013 (Leidinger et al. 2013). A successive work later described a novel 9-miRNA signature (hsa-miR-26a-5p, hsa-miR-181c-3p, hsa-miR-126-5p, hsa-miR-22-3p, hsa-miR-148b-5p, hsa-miR-106b-3p, hsa-miR-6119-5p, hsa-miR-1246, and hsa-miR-660-5p) to be used as a biomarker for detecting AD, helping in improving the AD diagnosis, especially at an early stage, and in classifying its clinical stages (Guo et al. 2017).

Table 4 Principal miRNAs dysregulated in AD, type of their deregulation, and their mechanism in AD

Name of miRNA	Deregulation in AD	Mechanism in AD	References
hsa-miR-101	Downregulation	miR-101 is involved in the control of APP translation and Aβ fibril accumulation miR-101b also directly targets the 3′ UTR of *AMPK* to post-transcriptionally regulate its protein levels. HDAC2/miR-101/AMPK pathway is thus a critical mediator of AD pathogenesis	Vilardo et al. (2010), Liu et al. (2017a)
hsa-miR-124	Downregulation	miR-124 plays neuroprotective roles in AD Drosophila by targeting Delta in the Notch signaling pathway. Moreover, it reduces abnormal Tau phosphorylation and regulates the expression of BACE1/-secretase	Fang et al. (2012), Kong et al. (2015), Zhou et al. (2019)
hsa-miR-125b	Upregulation	In vitro, miR-125b overexpression promotes the expression of amyloid precursor protein and β-secretase 1 and produces the Aβ-peptide, while decreasing the sphingosine kinase 1 (SphK1) protein expression. miR-125b induces apoptosis, inhibits cell proliferation, and induces oxidative stress and inflammation	Jin et al. (2018)
hsa-miR-128	Upregulation	miR-128 targets the peroxisome proliferator-activated receptor gamma (PPAR-γ). Inhibition of miR-128, causing PPAR-γ upregulation, decreases the Aβ-mediated cytotoxicity through the inactivation of NF-κB in AD models in vitro	Geng et al. (2018)
hsa-miR-132	Downregulation	miR-132 targets Tau mRNA and regulates its expression. Levels of miR-132/212 clusters are associated with insoluble Tau and cognitive impairment in humans	Smith et al. (2015)
hsa-miR-139	Upregulation	miR-139 targets the cannabinoid receptor type 2 (CB2) and regulates its expression. In particular, miR-139 inversely modulates the responses to pro-inflammatory stimuli	Tang et al. (2017)
hsa-miR-146a	Upregulation	Overexpression of miRNA-146a in vitro significantly decreases the expression of low-density lipoprotein receptor-related protein-2 (Lrp2), resulting in the reduction of Akt activation and induction of pro-apoptotic caspase-3, thereby increasing cell apoptosis.	Lukiw et al. (2008), Zhang et al. (2016)

(continued)

Table 4 (continued)

Name of miRNA	Deregulation in AD	Mechanism in AD	References
		miR-146a modulates the complement factor H (CFH) gene expression via NF-kB-sensitive miRNA-146a. This pathway may regulate an inflammatory response in AD brains and in cell models of AD	
hsa-miR-16	Downregulation	miR-16 targets the gene encoding amyloid precursor protein (APP). Downregulation of miR-16 decreases apoptosis, while its upregulation increases it	Zhang et al. (2015a)
hsa-miR-186	Downregulation	miR-186 may inhibit the development of AD through the downregulation of interleukin-2 (IL2) through suppression of the Janus kinase/signal transducers and activators of the transcription (JAK–STAT) signaling pathway	Wu et al. (2018)
hsa-miR-188-3p	Downregulation	In transgenic mice, the endocannabinoid 2-arachidonoylglycerol (2-AG) upregulates miR-188-3p by the PPARγ and NF-κB signaling pathway. This results in the inhibition of BACE1 expression and Aβ formation	Zhang et al. (2014)
hsa-miR-206	Upregulation	miR-206 upregulation increases lipopolysaccharide (LPS)-induced inflammation and Aβ release in microglia by directly targeting the 3′-untranslated region of (Insulin-like growth factor 1) IGF1	Xing et al. (2016)
hsa-miR-212	Downregulation	Cluster miR-132/212 deficiency leads to increased Tau expression, phosphorylation, and aggregation in mice	Smith et al. (2015)
hsa-miR-29a/b-1	Downregulation	miR-29a, -29b-1 can regulate BACE1 expression in vitro. Decreasing of specific miRNAs may contribute to increasing the BACE1 and Aβ levels in sporadic AD	Hébert et al. (2008)
hsa-miR-29c	Downregulation	miR-29c expression is correlated with BACE1 expression	Yang et al. (2015)
hsa-miR-330	Downregulation	miRNA-330 targets VAV1 via the MAPK signaling pathway; this causes protective effects on Aβ production, oxidative stress, and mitochondrial dysfunction in AD	Zhou et al. (2018)

(continued)

Table 4 (continued)

Name of miRNA	Deregulation in AD	Mechanism in AD	References
hsa-miR-339	Downregulation	miRNA-339-5p regulates BACE1 expression	Long et al. (2014)
hsa-miR-34a	Downregulation	p79 isoform (TAp73) regulates miR-34a during neuronal differentiation. When neurons are altered by the Aβ42 presence, MEK-ERK signaling is activated, causing the deregulation of TAp73/miR-34a and promoting cyclin D1 expression, which leads to apoptosis	Modi et al. (2016)
hsa-miR-34a	Upregulation	Altered expression of miR-34a could result in dysfunction of energy metabolism, synaptic plasticity, and resting-state network activity	Sarkar et al. (2016)
hsa-miR-34c	Upregulation	miR-34c inhibition upregulates the vesicle-associated membrane protein 2 (VAMP2) expression and rescues the memory deficit and synaptic failure caused by Aβ	Hu et al. (2015)
hsa-miR-922	Upregulation	miR-922 upregulates the level of phosphorylated Tau by regulating soluble ubiquitin carboxy-terminal hydrolase (L1UCHL1) levels	Zhao et al. (2014)
hsa-miR-98	Downregulation	miR-98-5p negatively regulates the sorting of nexin 6 (SNX6) expression by targeting the 3′-UTR of SNX6 mRNA. This pathway is involved in the regulation of Aβ production	Liu et al. (2015)

Moreover, microRNA-200b-5p was identified as a potential biomarker for AD by virtue of its high levels in AD tear fluid samples compared to controls (Kenny et al. 2019). These data suggest that tears may be a useful novel source of biomarkers for AD, and these may allow for the development of a non-invasive and cost-effective diagnostic test for AD.

In addition, the observation that miRNAs interfere with (or modulate) the expression of candidate genes in AD has pushed the researchers to develop miRNA-based therapeutic strategies. This is also because the therapies that aim at targeting Aβ have not been successful in treating or even slowing down the disease (Marsh and Alifragis 2018). In this regard, the possibility of modulating the mRNA expression of an AD candidate's gene has been recently explored by using antisense oligonucleotides (ASOs), which act similarly to miRNA by inhibiting mRNA transcription (DeVos et al. 2017). These strategies may be especially useful against Tau pathology, as murine models have already shown improved learning and memory, as a response to ASO treatment against GSK-3β, one of the primary kinases responsible

for the hyperphosphorylation and subsequent dysfunction of Tau (Orlacchio et al. 2007; Farr et al. 2014).

For the first time in 2019, piRNAs were found to be differentially expressed in human CSF exosomes of AD patients. In particular, Next-Generation Sequencing data, collected to study the small non-coding RNAome (sncRNAome) in exosomes derived from human cerebrospinal fluid (CSF), show that three piRNAs in CSF exosomes together with three miRNAs define a recognizable signature to detect AD. Furthermore, the piRNA signature could help predict the conversion of mild cognitive impairment (MCI) patients to AD (Jain et al. 2019). These data are in line with a recent report suggesting that piRNAs might play a crucial role in AD pathology, and more specifically that Tau pathology disrupts piRNAs functions leading to genome instability (Sun et al. 2018). However, when compared with miRNAs, the role of piRNAs is still less understood.

As for lncRNAs, they are implicated in alternative splicing, which is a central component of human brain complexity; in fact, many types of lncRNA are known to be normally expressed in the nervous system (Briggs et al. 2015). Moreover, lncRNAs have been demonstrated as transcriptional regulators. A number of studies have been performed on the abnormal expression of lncRNA-antisense transcripts in subjects with AD (Amlie-Wolf et al. 2019; Cao et al. 2019). In particular, most of these studies detected the involvement of lncRNA on Aβ metabolism, synaptic impairment, neurotrophin depletion, inflammation, mitochondrial dysfunction, and stress responses, which are all involved in AD pathogenesis (Millan 2017; Li et al. 2020a). However, their functions and mechanisms have yet to be investigated (Riva et al. 2016; Wan et al. 2016).

LncRNAs are differentially expressed in both AD patients and animal models, so that they could be seen both as biomarkers for the early diagnosis of AD and as novel therapeutic targets for its treatment. For example, the dysregulation of blood-based lncRNA BACE1-AS can be detected in AD patients (Fotuhi et al. 2019; Li et al. 2020a). As for the diagnosis, there is still no further investigation of the lncRNA involvement in AD.

The lncRNAs that are experimentally supported on LncRNADisease database version 2.0 are listed in the following Table 5.

Further to this, lncRNAs may become a new therapeutic target for AD treatment, in particular, ASO-based lncRNA knockdown has been demonstrated to have efficacy in the treatment of AD by targeting MAPT (TAU) (Scoles et al. 2019; Zhao et al. 2019).

Among ncRNAs, in addition to miRNAs and lncRNAs, circRNAs also deserve a mention. For example, it has been reported the existence in AD mouse models of circRNA-associated-competing endogenous RNA (ceRNA) networks involved in Aβ clearance and myelin function (Zhang et al. 2017). Significant associations were identified between circRNAs expression and AD diagnosis: it was demonstrated that most circRNA–AD associations are independent of changes in cognate linear messenger RNA expression or estimated brain cell-type proportions. Evidence was provided for circRNAs expression changes occurring early in presymptomatic AD and in autosomal dominant AD. It was also observed that AD-associated circRNAs

Table 5 Principal lncRNAs dysregulated in AD, type of their deregulation, and their mechanism in AD

Name of lncRNA	Deregulation in AD	Mechanism in AD	References
BACE1-AS	Upregulation	The antisense transcript BACE1 (BACE1-AS) regulates the mRNA of BACE1 and consequently the expression of the BACE1 protein both in vitro and in vivo. Increased beta-amyloid 1–42 causes elevated levels of BACE1-AS, which causes the stability of BACE1 mRNA, thereby generating more beta-amyloid 1–42	Faghihi et al. (2008, 2010)
MIAT	Downregulation	The lncRNA MIAT is involved in amyloid clearance via regulating the low-density lipoprotein receptor-related protein 1 (LRP1) expression, and miR-150-5p/VEGF mediated fibrillogenesis. The MIAT knockdown increases Aβ40-42 levels, neuronal loss while decreasing the expression and the number of tight junction proteins	Jiang et al. (2016a)
EBF3-AS	Upregulation	The lncRNA EBF3-AS regulates EBF3 (early B cell factor 3) expression and induces neuron apoptosis in AD	Gu et al. (2018)

are co-expressed with known AD genes. Finally, potential miRNA-binding sites were identified in AD-associated circRNAs for miRNAs predicted to target AD genes. Altogether, these results highlight the importance of analyzing non-linear RNAs and support future studies exploring the potential roles of circRNAs in AD pathogenesis (Dube et al. 2019). For example, by means of microarray, a recent study has found that 112 circRNAs were upregulated and 51 circRNAs were downregulated in AD patients compared with control subjects. In particular, circ-AXL, circ-GPHN, and circ-PCCA could have clinical implications for guiding disease management in AD patients (Li et al. 2020b).

In particular, it was reported on LncRNADisease database version 2.0 has shown evidence for a novel and significantly misregulated ciRS-7-miRNA-7-UBE2A signaling circuit in sporadic AD neocortex (Brodmann A22) and hippocampal CA1 (Zhao et al. 2016). These data suggest that the alteration of circRNA-miRNA-mRNA regulatory signaling represents an important state of epigenetic control over pathogenic gene expression programs in sporadic AD (Zhao et al. 2016).

In conclusion, much evidence has shown that ncRNAs are implicated in the central nervous system and drive neurodegeneration. Finally, the use of a combination of multiple ncRNAs as markers in biofluid can help the diagnosis, differentiation, and prediction of AD.

As for the treatments of AD, large sample trials are instead still necessary in order to draw a robust conclusion (Wei et al. 2020).

3.3 Epitranscriptome and Huntington's Disease

Huntington's disease (HD) is a dominantly inherited neurodegenerative pathology characterized by dementia, psychiatric disorders, and motor dysfunctions such as choreiform movements (Kumar et al. 2020; Tabrizi et al. 2020). HD is caused by abnormal Cytosine–Adenine–Guanine (CAG) triplet expansion within the exon 1 (N-terminal region) of the huntingtin (HTT) gene located on chromosome 4p16.3 (Kumar et al. 2020; Tabrizi et al. 2020). The number of CAG repeats is unstable and is strictly related to HD manifestation: non-pathogenic HTT alleles contain up to 35 CAG repeats (Kumar et al. 2020; Tabrizi et al. 2020), alleles with 36–39 repeats are considered mutant but with low penetrance (advanced age-onset and mild phenotype), alleles with more than 40 repeats show full penetrance with the general age of onset being after 60 years of age (late-onset HD), alleles with over 60 repeats are generally associated with a young age of onset (also known as Juvenile Huntington, JHD) (Fusilli et al. 2018), and markedly expanded alleles (>80 repeats) are related to pediatric HD (Migliore et al. 2019).

The CAG triplet encodes for Glutamine and its expansion is common to a group of diseases, named Polyglutamine (PolyQ) diseases, which are characterized by proteins with PolyQ tracts. The latter tend to aggregate intracellularly and are thus considered to have a key role in these diseases' pathogenesis (Bogomazova et al. 2019).

The physiological HTT gene has two predominant alternative mRNA transcripts, 10.3 kb and 13.7 kb, that differ in the 3′UTR sequence (The second is hypothesized to be enriched in non-dividing cells such as neurons) (Saudou and Humbert 2016; Romo et al. 2017). Mutant HTT is thought to be involved in the impairment of different pathways, such as autophagy, neurotransmission, and mitochondria, and can produce a pathogenic aberrant truncated transcript, as the abnormal CAG repeat promotes early polyadenylation and suppression of splicing between exon 1 and 2. This behavior appears to be mediated by the binding of Serine and Arginine Rich Splicing Factor 6 (SRSF6) to the mutant HTT mRNA with higher affinity compared to its physiological counterpart (Bogomazova et al. 2019; Nourse et al. 2020), resulting in the promotion of polyadenylation from a cryptic polyadenylation signal within intron 1 and the formation of a strongly pathogenic truncated transcript (Nourse et al. 2020). The abovementioned different HTT transcripts have different tissue specificity, expression, and stability, in addition to different sites for RNA-binding proteins and miRNAs (Romo et al. 2017). As a matter of fact, the presence of pathological extra CAG copies is accompanied by the formation of secondary RNA structures and the impairment of normal transcriptional and translational processes that can generate non-functional or misfolded proteins. In particular, the CAG repeats form a long hairpin that contributes to mutant HTT mRNA accumulation in the nucleus, as seen in other PolyQ diseases (Neueder and Bates 2018; Bogomazova et al. 2019). Mutant HTT mRNA binds different proteins involved in splicing, such as muscle blind like splicing regulator 1 (MBNL1),

which results in the alteration of several splicing targets (Neueder and Bates 2018; Angelbello et al. 2020).

As for dysregulations in translational processes, some studies highlighted that the pathological CAG repeat contributes to the depletion of the charged $tRNA^{Gln}$ (CUG), especially in the striatum, which causes a more frequent ribosome frameshifting that may worsen the disease phenotype (Leighton and Bredy 2018; Lant et al. 2019).

Moreover, the expanded CAG repeat mediates the binding of proteins involved in translational processes, such as midline 1 (MID1), protein phosphatase 2A (PPP2A), and ribosomal protein S6 kinase (S6K), which enhance the translation of mutant HTT (Neueder and Bates 2018; Bogomazova et al. 2019; Lontay et al. 2020). The CAG expansion of HTT mRNA has been suggested to be involved in the repeat-associated non-ATG (RAN) translation, a process in which ribosomes start translation without the presence of the canonical ATG sequence, consequently generating proteins with different amino acidic expansion (e.g., polyAlanine, polySerine, polyCysteine), although the contribution of this mechanism to HD pathogenesis has not been fully elucidated yet (Rodriguez and Todd 2019; Rudich et al. 2020).

It is well established that epigenetic, transcriptional, and translational alterations are all implicated in HD (Bassi et al. 2017). Not only is the transcriptome in HD regulated by mRNA expression and splicing but also by altered non-coding RNA expression. Several lncRNAs have indeed been investigated for their role in HD pathogenesis. For example, the antisense transcript of the HTT gene (HTT-AS) is thought to downregulate mutant HTT expression and was found to be reduced in HD brains compared to controls (Salvatori et al. 2020; Wu and Kuo 2020). Interestingly enough, mutant HTT modulates the nuclear translocation of RE1 silencing transcription factor (REST), which has a particularly central role for neurons development and homeostasis, thus altering the expression of REST target genes (Orozco-Díaz et al. 2019; Garcia-Manteiga et al. 2019). For example, the lncRNA DiGeorge critical region 5 (DGCR5) and Maternally Expressed 3 (MEG3) are a REST target and were proposed to be deregulated in HD (Wang et al. 2018a; Vieira et al. 2018). The lncRNA NEAT1 (nuclear paraspeckle assembly transcript 1) is considered to have a crucial role in several neurodegenerative diseases and its expression in HD CNS was found to be altered in cell and animal models, as well as human postmortem brain samples (Chanda et al. 2018; Cheng et al. 2018). Although the NEAT1 role in HD pathogenesis still needs to be clearly understood, its overexpression was proposed to have a neuroprotective function in CAG repeat expansion diseases (Cheng et al. 2018; An et al. 2018). Other ncRNAs that were proposed to be altered in HD are the antisense transcripts of the Brain-derived neurotrophic factor (BDNF-AS), the Human accelerated region 1 (HAR1), Taurine UpRegulated 1 (TUG1), and Tcl1 Upstream Neuron-Associated lincRNA (TUNA), as confirmed by the database LncRNADisease v2.0 (Zimmer-Bensch 2019; Salvatori et al. 2020). The role of another type of lncRNA was recently investigated the deregulation of enhancer RNAs (eRNAs) was observed in HD mice striatum compared to controls. In particular, their data indicated that there is a decreased eRNAs expression in HD mice, which is exacerbated by the loss of RNA Polymerase II (RNAPII) binding

Table 6 Principal miRNAs dysregulated in HD, type of their deregulation, and their mechanism in HD

Name of miRNAs	Deregulation in HD	Mechanism in HD	References
hsa-miR-125b-1/2	Downregulation	miR-125b-1/2 targets the HTT gene, influences the HTT aggregate formation/toxicity and the BDNF expression	Sinha et al. (2011)
hsa-miR-146a	Downregulation	miR-146a targets the HTT gene, influences the HTT aggregate formation/toxicity and the BDNF expression	Sinha et al. (2011)
	Upregulation in the striatum of patients	miR-146a plays a role in neuroinflammatory processes related to neurodegenerative diseases	Fan et al. (2020)
hsa-miR-150	Downregulation	miR-150 targets the HTT gene, influences the HTT aggregate formation/toxicity and the BDNF expression	Sinha et al. (2011)
hsa-miR-214	Upregulation	miR-214 targets the HTT gene, influences the HTT aggregate formation/toxicity and the BDNF expression	Sinha et al. (2011)
		miR-214 targets and downregulates Mitofusin2, alters mitochondrial morphology, and deregulates cell cycle	Bucha et al. (2015)
hsa-miR-22	Downregulation	miR-22 regulates genes implicated in HD (histone deacetylase 4, REST corepressor 1 and regulator of G-protein signaling 2)	Jovicic et al. (2013)

sites at enhancers. This finding correlates with a downregulation of the associated genes (Le Gras et al. 2017).

In light of their pivotal role as key post-transcriptional regulators and their involvement in fundamental processes of cell biology (e.g., cell growth, apoptosis), miRNAs deregulation has been proposed to contribute to neurodegenerative diseases. In recent years many miRNAs have been found dysregulated in HD patients and HD cell/animal models (John et al. 2020; Catanesi et al. 2020). Among these, we report in Table 6 the miRNAs deposited in the Human microRNA Disease Database (Huang et al. 2019b), a database that curates experiment-supported evidence for human miRNA and disease associations.

Among the therapeutic strategies for HD, RNA-targeted approaches focused on inhibiting mutant HTT transcriptional and translational processes have recently shown to be promising in counteracting the disease progression. HTT mRNA can be targeted by using diverse strategies such as antisense oligonucleotides (ASOs) or RNA interference (RNAi) (Shannon 2020; Smith and Tabrizi 2020; Dash and Mestre 2020).

ASOs that have entered the clinical trial phase are RG6042/tominersen (Phase III, NCT03761849), allele-specific WVE-120101, and WVE-120102 (Phase Ib/2a, NCT03225833, NCT03225846). Concerning RNAi-based strategies, the only product currently in Phase I/II trial is the nonallele-specific miRNA AMT-130 coupled to an AAV5 vector (NCT04120493). An additional recent strategy is the modulation of

mutant HTT RNA splicing that led to the development of two orally available agents, PTC-CHDI NOS and Skyhawk-Novartis NOS Preclinical, which are still being studied only in preclinical models (Dash and Mestre 2020).

In conclusion, it has now become clear that the advances in the understanding of HD pathogenesis rely on the comprehension of the RNA dysregulation role. In particular, mRNA toxicity and ncRNAs deregulation appear to be deeply involved in the molecular mechanism underlying this disease, and, at the same time, they could help the assessment of the disease progression and become a promising target for therapeutic strategies.

3.4 Epitranscriptome and Parkinson's Disease

Parkinson's disease (PD) is a neurodegenerative disease that impairs movement skills. It is characterized by the degeneration of dopaminergic (DA) neurons in the substantia nigra-striatum system (Chen et al. 2019b; Qin et al. 2020). Due to the depletion of the nigrostriatal pathway, the most severe symptomatology that characterizes this disease is represented by bradykinesia, hypokinesia, rigidity, resting, and postural tremor stability (Hanan et al. 2020). The hallmarks of PD are the loss of DA neurons in the Substantia Nigra (SN) and the accumulation of misfolded α-synuclein, found in intra-cytoplasmic inclusions called Lewy bodies (LBs) (Balestrino and Schapira 2020). This neurodegenerative disease affects approximately seven million people worldwide (Quan et al. 2017; Wei et al. 2018).

The diagnosis of PD is essentially clinical, and the diagnostic criteria for PD were recently updated by the Movement Disorders Society (MDS) (Postuma et al. 2015). PD can be distinguished in a familial form, due mostly to genetic factors, and a sporadic form, which can be caused by exposure to various toxic substances and may involve gene-environment interactions (Balestrino and Schapira 2020).

To date, the etiology of the disease in most patients is unknown, although several associated genetic mutations and some risk factors have been identified. Among these, age is the most important cause for the outbreak of the disease, but the male gender also represents a moderate risk (Gillies et al. 2014). In addition, some environmental agents have also been linked to the possible manifestation of PD, including exposure to some pesticides (Lee and Gilbert 2016) and to some substances, such as 1-methyl-4-phenyl tetrahydropyridine (MPTP) and Annonacin, which can cause nigrostriatal cell death (Höglinger et al. 2005).

Family history is, however, a very high-risk factor for the development of PD, and the relative risk in first-degree relatives of PD cases increases approximately two to three times compared to controls (Nalls et al. 2019). Familial forms of PD represent in fact only 5–15% of total cases (Balestrino and Schapira 2020).

Recent research has shown that both *epigenetics* and *epitranscriptomics* occupy a central role in the development of neurodegenerative diseases such as PD (Bicchi et al. 2013). In general, what emerged is that the initiation of this type of pathology is attributable to both a genetic predisposition and environmental triggers (Noyce et al.

2012). Recent studies have indeed shown that dysregulation of some lncRNAs is closely correlated with the onset of PD (Wei et al. 2018; Zhang et al. 2020). For example, 5 lncRNAs have been identified to be differently expressed in PD patients' neurons (H19, lincRNA-p21, MALAT1, SNHG1, and NEAT1), and this dysregulation is already present in the early stages of the pathology; in particular, the lncRNA H19 is strongly downregulated, while the others are upregulated. All these lncRNAs are naturally linked to cell proliferation, synaptogenesis, and apoptosis in CNS; their dysregulation may therefore be precisely connected with the development of neurological diseases such as PD (Wu et al. 2013; Kraus et al. 2017). The dysregulation of numerous miRNAs and lncRNAs has indeed been found to correlate with the molecular mechanisms of PD, as reported in Table 7.

Oxidative damage affecting RNAs during cellular aging is also correlated with the onset and progression of PD (Liu et al. 2020c). It has been shown that the damage induced by ROSs on RNAs mainly leads to the formation of 8-OHG, which accumulates in the nucleus and mitochondria of senescent cells, and in particular in SN-DA neurons of PD patients, resulting in the loss of these neurons (Liu et al. 2020c). Further to this, other studies have shown how oxidative damage, which involves RNAs through the formation of 8-OHG, is found in the early stages of PD and how it is correlated with the course and severity of the disease (Liu et al. 2020c). A recent study aimed at understanding whether a dysregulation in m^6A RNA methylation was also implicated in PD (Chen et al. 2019b). As previously described, N6-methyladenosine (m^6A) is the main modification that can affect RNA, playing a key role in the nervous system and the onset of neurodegenerative diseases; however, to date, there is still no clear evidence of its involvement in the development of PD. The relationship between m^6A mRNA methylation and the molecular mechanism underlying PD was explored both in vitro and in vivo models of PD induced by 6-hydroxydopamine (6-OHDA). In both cases, the level of m^6A in DA cells was decreased by overexpression of the demethylase FTO (fat mass and obesity-associated protein) and by using an inhibitor of m^6A, i.e., Cycloleucine. The results showed that m^6A reduction could induce *N*-methyl-D-aspartate receptor 1 (NMDAR1) expression, increase oxidative stress and Ca^{2+} influx, resulting in apoptosis of DA neurons (Fig. 5) (Chen et al. 2019b).

Thanks to this study, it was generally understood how the m^6A modification carries out a decisive action in the death of the DA neurons and how the *epitranscriptome* is therefore involved in the development of PD. Undoubtedly, in-depth studies will have to be conducted to understand better the relationship between the modifications involving the transcriptome and the molecular mechanisms underlying neurodegenerative diseases, such as PD. Doing so would be feasible to identify useful molecular targets for the development of effective therapies.

Table 7 Principal lncRNAs and miRNAs dysregulated in PD, type of their deregulation, and their mechanism in PD

Name	Type of RNA	Deregulation in PD	Mechanism in PD	References
hsa-let-7d	miRNA	Downregulation	The downregulation of let-7d induces neuronal damage with loss of DA neurons, while normal expression attenuates neurotoxin 6-hydroxydopamine (6-OHDA)-induced damage in PD mice by targeting caspase-3 in MN9D cells	Li et al. (2017)
hsa-miR-126	miRNA	Upregulation	The miR-126 upregulation in DA neuronal phenotype impairs the insulin/IGF-1/PI3K signaling pathway and negatively affects cell survival to neurotoxic insult and protection by IGF-1. The dysregulation of insulin signaling by miR-126 may be a contributing factor in PD pathogenesis	Kim et al. (2014)
hsa-miR-16-1	miRNA	Upregulation	The miR-16-1 upregulation promotes aberrant α-synuclein accumulation in PD via targeting heat shock protein 70 (HSP70)	Zhang and Cheng (2014)
hsa-miR-205	miRNA	Downregulation	The miR-205 downregulation may contribute to the potential pathogenic elevation of leucine-rich repeat kinase 2 (LRRK2) protein in the brain of patients with sporadic PD	Cho et al. (2013)
hsa-miR-214	miRNA	Downregulation	The miR-214 downregulation in PD results in an increase of α-synuclein expression, which is strongly associated with the pathogenesis of PD.	Wang et al. (2015)
hsa-miR-22	miRNA	Downregulation	The miR-22 upregulation in 6-OHDA-treated PC12 cells promotes the survival and proliferation of the cells, whereas the miR-22 inhibitor reverses this effect	Yang et al. (2016)
hsa-miR-221	miRNA	Downregulation	miR-221 plays a protective role in PD via regulating PC12 cell viability and apoptosis by targeting PTEN (Phosphatase and Tensin homolog)	Li et al. (2018c)
hsa-miR-30e	miRNA	Downregulation	The downregulation of miR-30e induces neuroinflammation and the loss of DA neurons, a pathological hallmark of PD	Li et al. (2018a)
hsa-miR-342	miRNA	Upregulation	The miR-342-3p upregulation inhibits proliferation and promotes apoptosis of DA neurons in PD	Wu et al. (2019)

(continued)

Table 7 (continued)

Name	Type of RNA	Deregulation in PD	Mechanism in PD	References
hsa-miR-34b	miRNA	Downregulation	The miR-34b/c downregulation induces an increase of striatal adenosine A2A receptor levels, which is an early event in PD-related pathology	Villar-Menéndez et al. (2014)
hsa-miR-7	miRNA	Downregulation	The upregulation of miR-7 promotes cell viability and protects from 1-methyl-4-phenylpyridinium iodide [MPP(+)]-induced cell apoptosis in SH-SY5Y, by directly targeting KLF4. Furthermore, miR-7 inhibits neuronal apoptosis in cellular PD models by targeting Bax and Sirt2	Li et al. (2016), Kong et al. (2016)
H19	lncRNA	Downregulation	The lncRNA H19 may attenuate neuronal apoptosis in 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-induced PD mice	Zhang et al. (2020)
lincRNA-p21	lncRNA	Upregulation	The long intergenic non-coding RNA-p21 (lincRNA-p21) may increase neuronal apoptosis in PD mice	Ge et al. (2020)
HOTAIR	lncRNA	Upregulation	The lncRNA HOTAIR promotes PD by upregulating the leucine-rich repeat kinase 2 (LRRK2) expression	Liu et al. (2016)
MALAT 1	lncRNA	Upregulation	The lncRNA MALAT1 promotes the apoptosis of DA neurons in PD	Liu et al. (2017b)
SNHG1	lncRNA	Upregulation	The lncRNA SNHG1 promotes α-synuclein aggregation and toxicity by targeting miR-15b-5p to activate SIAH1 (Seven In Absentia Homolog 1) in SH-SY5Y cells, promoting the loss of DA neurons	Chen et al. (2018)
NEAT1	lncRNA	Upregulation	The lncRNA NEAT1 mediates the toxicity in PD induced by MPTP/MPP + via regulation of gene expression	Liu and Lu (2018)

3.5 *Epitranscriptome and Amyotrophic Lateral Sclerosis*

Amyotrophic Lateral Sclerosis (ALS) is a progressive and lethal neurodegenerative disease characterized by the selective destruction of motor neurons. In particular, when the motor neurons of the brain, brainstem, and spinal cord weaken, the skeletal muscular atrophy propagates in the patient (van Es et al. 2017; Burk and Pasterkamp 2019).

In fact, the state of muscular weakness is followed by paralysis, trouble swallowing, respiratory muscle weakness, and respiratory impairment, which ultimately leads to the death of the patient. ALS has a different clinical course based on the type of motor neurons affected by the disease: ALS can present a limbs-onset,

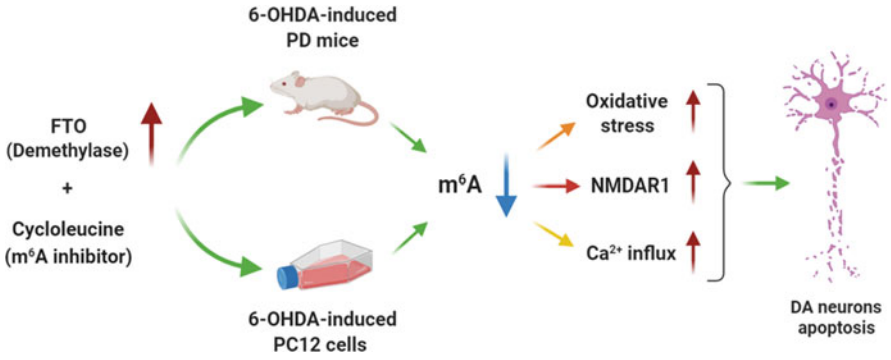


Fig. 5 DA neurons' apoptosis promoted by the reduction of m^6A in 6-OHDA-induced PC12 cells and PD mice through overexpression of the demethylase FTO and by using Cycloleucine, an inhibitor of m^6A (Chen et al. 2019b)

manifesting as muscle weakness, or a bulbar-onset, linked to speech impediments. Prognosis is difficult, and most patients die within 3 years from diagnosis. To date, there is no treatment for ALS, and the only two drugs approved by the Food and Drug Administration, Radicava (edaravone) and Rilutek (riluzole), fail to stop the progression of the disease (Chia et al. 2018; Oskarsson et al. 2018). ALS is divided into two categories: familial and sporadic. Familial ALS, where a specific gene mutation can be attributed to the disease, accounts for about 10% of total cases. On the other hand, sporadic cases, for which there is no family history, represent the remaining 90%. Numerous genes have been associated with this disease, including superoxide dismutase 1 (SOD1), Fused in Sarcoma (FUS), chromosome 9 open reading frame 72 (C9orf72), and TAR DNA-binding protein 43 (TDP-43) (Mathis et al. 2019). The course of the pathology is similar in both familial and sporadic forms: the abovementioned mutations are not exclusively present in neuronal cells, and neuronal TDP-43 protein aggregates are found in the vast majority of cases of this neurodegenerative disease (van Es et al. 2017; Bennett et al. 2019).

Recently, genome-wide association studies discovered differentially methylated genes involved in several pathways important to ALS. In particular, the CNS dysregulation of miRNA-related pathways is linked with significant neuronal damage and cell death, which may contribute to the progression of neurodegenerative diseases such as ALS (Bicchi et al. 2013; Rinchetti et al. 2018). In a protein complex with RNAse III DORSHA, the ALS genes TDP-43 and FUS were identified and found to play a role in miRNA biogenesis. In particular, TDP-43 has been shown to interact with proteins involved in pre-miRNA cytoplasmic cleavage mediated by the DICER enzyme, so it is no wonder that miRNA dysregulation has been spotted in ALS (Butti and Patten 2018). Interestingly, several neurodegenerative diseases including ALS share numerous deregulated miRNAs (Ferrante and Conti 2017; Quinlan et al. 2017). These deregulated miRNAs are essential for the function and maintenance of motor neurons, axonal growth and synaptic transmission and miRNA modifications. Hence, they probably contribute to the pathological

phenotype of ALS. In this regard, numerous studies have been published on miRNAs as potential circulating biomarkers for this neurodegenerative disease (Butti and Patten 2018).

The data indeed showed that 38 miRNAs were substantially downregulated for sporadic ALS patients in peripheral blood samples, as opposed to healthy controls, including miR-103a-3p, miR-106b-3p, miR-128-3p, miR-130a-3p, miR-130b-3p, miR-144-5p, miR-148a-3p, miR-148b-3p, etc. It has also been observed that the bulbar/spinal onset and the progressive rate of the disease are characterized by different miRNAs (Liguori et al. 2018). Moreover, De Felice et al., when comparing ALS blood samples with healthy controls, identified 696 already known and 49 new miRNAs differently expressed in ALS tissues. Among these, the most upregulated miRNAs were found to be miR-1, miR-10b-5p, miR-153-3p, miR-224-3p, miR-224-5p, miR-326, miR-338-3p, miR-877-3p, miR-1296-5p, miR-4695-3p, miR-3194-3p, and miR-5684, and the most downregulated were miR-143-3p, miR-144-3p, miR-144-5p, miR-190a-5p, miR-193a-5p, miR-199b-5p, miR-338-5p, miR-218-5p, miR-542-5p, miR-618, miR-4423-3p, and hsa-miR-125a-3p (De Felice et al. 2018; Ravnik-Glavač and Glavač 2020).

Another type of ncRNA that can be investigated as circulating biomarkers in ALS is lncRNAs. In fact, lncRNAs can silence or promote proximal gene expression on an epigenetic, transcriptional, and post-transcriptional level (Fernandes et al. 2019). Thus far, only one study has documented lncRNAs in ALS patients with different expressions in peripheral blood mononuclear cells (PBMCs) relatively to healthy controls (Gagliardi et al. 2018). A total of 293 lncRNAs in sporadic patients with ALS without the detected mutation has been found. The majority of these lncRNAs, 184/293 transcripts, are antisense and mostly unknown. Seven among these antisense lncRNAs have been detected in TDP-43 mutated patients, and only one, i.e., SNAP25-AS, had already been described. Two new antisense RNAs have instead been revealed in ALS patients with SOD1 mutation, e.g., the mitochondrial 2 (CKMT2) antisense (Gagliardi et al. 2018). However, much remains to be done in order to evaluate and understand the role of lncRNAs in ALS (Ravnik-Glavač and Glavač 2020).

CircRNAs are yet another class of ncRNAs with regulatory function resulting from splicing events during precursor mRNA processing and may thus impact the regulation of genes. Each circRNA binds several miRNAs competitively and reduces the mRNA silencing capability (Xie et al. 2017; Rong et al. 2017). Data analysis of circRNAs in leukocyte samples from sporadic ALS patients and healthy controls has allowed 425 differentially expressed circRNAs to be obtained. Among these, four showed a clinical relevance (Dolinar et al. 2019). For example, hsa_circ_0000567, located in the SETD3 gene, regulates muscle differentiation in mouse (Eom et al. 2011); the hsa_circ_0023919 sequence presents two binding sites for hsa-miR-9, and the upregulation of this miRNA was confirmed in both mouse models of ALS and in the blood of ALS patients (Vrabec et al. 2018); hsa_circ_0063411 binds the miR-647, which has already been found in the spinal cord of ALS samples and absent in the controls (Campos-Melo et al. 2013); finally, the hsa_circ_0088036, located in the sushi domain-containing one gene, is

Table 8 Principal lncRNAs and miRNAs dysregulated in ALS, type of their deregulation, and their mechanism in ALS

Name	Type of RNA	Deregulation in ALS	Mechanism in ALS	References
hsa-miR-125b	miRNA	Upregulation	miR-125b is significantly upregulated in ALS upon stimuli such as TNFα or 2′-3′-*O*-(benzoyl-benzoyl) ATP (BzATP) acting on P2X7r and sustaining inflammatory signaling in microglia	Parisi et al. (2016)
hsa-miR-206	miRNA	Upregulation	In vitro and in vivo studies have shown that miR-206 is involved in the formation of myofibers, differentiation of satellite cells, and formation of new neuromuscular junctions following nerve injury	Valdez et al. (2014)
NEAT1_2	lncRNA	Upregulation	NEAT1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis	Nishimoto et al. (2013)
MALAT1	lncRNA	Upregulation	Recruits splicing factors to paraspeckles	Riva et al. (2016)

potentially associated with ALS (Schymick et al. 2007). In conclusion, the hsa_circ_0023919, hsa_circ_0063411, and hsa_circ_0088036 can be used as potential diagnostic biomarkers of ALS thanks to their sensitivity and specificity for the optimal threshold point (>90%) (Dolinar et al. 2019; Ravnik-Glavač and Glavač 2020).

However, up to now, among the different ncRNAs implicated in the ALS pathology, those recognized and confirmed by two different databases as implicated in the progression of this neurodegenerative pathology are reported in Table 8 (Chen et al. 2013; Huang et al. 2019b).

3.6 Epitranscriptome and Lysosomal Storage Diseases

Lysosomal storage diseases (LSDs) are a group of pathologies caused by a dysfunction in the activity of lysosomal proteins, with the consequent accumulation of non-degraded metabolites within the lysosome and loss of cell functions (Urbanelli et al. 2011; Platt et al. 2012; Martino 2015). The LSD family consists of about 50 diseases caused by inherited genetic mutations that occur within genes that would normally code for lysosomal enzymes. Most outbreaks of LSDs occur during childhood, with a few exceptions that may manifest in adulthood. However, 75% of LSDs involve severe neurological implications leading to physical deterioration, functional impairment, progressive neurodegeneration, and death (Calzoni et al. 2019; Morena et al. 2020). The dysfunctions affecting lysosomes have important consequences on cellular metabolism as a whole, as it has been demonstrated that these organelles play crucial roles in vesicle trafficking, autophagy, and cell growth

and signaling. Furthermore, lysosomal alterations have been associated with pathogenic mechanisms in other neurodegenerative diseases such as Huntington's, Parkinson's, and Alzheimer's (Morena et al. 2017; Argentati et al. 2020).

It has only recently been discovered that lysosomal functions are regulated by a gene network called Coordinated Lysosomal Expression and Regulation (CLEAR), which is activated by a transcription factor denominate transcription factor EB (TFEB). The latter is in turn part of the protein complex Lysosome Nutrient Sensing (LYNUS), which also includes the mammalian target of the rapamycin complex 1 (mTORC1), and represents machinery attached to the lysosomal membrane. Therefore, alterations in the TFEB-mediated regulatory mechanism and CLEAR network genes are closely associated with the onset of LSDs (Queiroz et al. 2016). Recent studies have shown that the gene network that modulates lysosomal function is regulated by various ncRNAs, such as miRNAs, siRNAs, piRNAs, and lncRNAs. Alterations affecting these important regulators may therefore lead to the onset of LSDs and tune their severity.

To date, all the studies concerning the association of ncRNAs and LSDs available in the literature have focused only on miRNAs (Morena et al. 2019). A first study was conducted on fibroblasts of patients with Niemann-Pick type C (NPC), a lipid storage disease caused by mutations in the NPC1 or NPC2 genes: among the 365 miRNAs identified, three were found to be upregulated and other 38 downregulated (Ozsait et al. 2010; Queiroz et al. 2016). Similarly, it was observed in mouse models of Gaucher disease that miR-29b and miR-142 were upregulated, while miR-let7b was downregulated. This dysregulation has been associated with the onset of the inflammatory response characteristic of this disease (Ginns et al. 2014).

As of today, no further studies have been published referring to the association between *epitranscriptome* changes and the development of LSDs. However, as is the case with other neurological diseases, it is clear how alterations in the ncRNA expression may be involved in the modulation of the disease development. In addition, ncRNAs could be used as a biomarker and potential therapeutic targets of the disease (Queiroz et al. 2016).

4 Conclusions

In conclusion, the role of *epitranscriptome* in the pathophysiology of CNS diseases was examined, turning special attention to the role of ncRNAs. In particular, the changes that can affect the various types of RNAs together with the dysregulation of specific ncRNAs (especially lncRNAs and miRNAs) have proved to be of fundamental importance in the onset of a number of diseases namely GBM, AD, HD, PD, and ALS.

From the many studies reported here, the enormous potential of the *epitranscriptome* stands out: The association between the aberrant modifications of RNAs, along with the occurrence of their dysregulation, and neurological disorders

leads to a thorough understanding of the pathophysiology of the considered diseases, as it allows the identification of specific molecular markers potentially usable for highly efficient targeting in the possible treatment of these pathologies.

However, the role of RNA modifications in these processes is so far not entirely clear, and, although the potential impact of *epitranscriptome* on the CNS is extremely promising, some issues still need to be addressed before these tools could be extensively used as biomarkers or therapeutic targets.

References

Alvarez-Dominguez JR, Lodish HF (2017) Emerging mechanisms of long noncoding RNA function during normal and malignant hematopoiesis. Blood 130:1965–1975

Alzheimer's Association (2020) Alzheimer's disease facts and figures. Alzheimers Dement. https://doi.org/10.1002/alz.12068

Amlie-Wolf A, Tang M, Way J et al (2019) Inferring the molecular mechanisms of noncoding Alzheimer's disease-associated genetic variants. J Alzheimers Dis 72:301–318

An H, Williams NG, Shelkovnikova TA (2018) NEAT1 and paraspeckles in neurodegenerative diseases: A missing lnc found? Noncoding RNA Res 3:243–252

Angelbello AJ, Chen JL, Disney MD (2020) Small molecule targeting of RNA structures in neurological disorders. Ann NY Acad Sci 1471:57–71

Angelova MT, Dimitrova DG, Dinges N et al (2018) The emerging field of epitranscriptomics in neurodevelopmental and neuronal disorders. Front Bioeng Biotechnol 6:46

Arena A, Iyer AM, Milenkovic I et al (2017) Developmental expression and dysregulation of miR-146a and miR-155 in Down's syndrome and mouse models of Down's syndrome and Alzheimer's disease. Curr Alzheimer Res 14:1305–1317

Argentati C, Tortorella I, Bazzucchi M et al (2020) The other side of Alzheimer's disease: influence of metabolic disorder features for novel diagnostic biomarkers. J Pers Med 10:115

Armstrong RA (2019) Risk factors for Alzheimer's disease. Folia Neuropathol 57:87–105

Asano K, Suzuki T, Saito A et al (2018) Metabolic and chemical regulation of tRNA modification associated with taurine deficiency and human disease. Nucleic Acids Res 46:1565–1583

Bagyinszky E, Giau VV, An SA (2020) Transcriptomics in Alzheimer's disease: aspects and challenges. Int J Mol Sci 21. https://doi.org/10.3390/ijms21103517

Balestrino R, Schapira AHV (2020) Parkinson disease. Eur J Neurol 27:27–42

Banelli B, Forlani A, Allemanni G et al (2017) MicroRNA in glioblastoma: an overview. Int J Genomics 2017:7639084

Barbagallo D, Condorelli A, Ragusa M et al (2016) Dysregulated miR-671-5p / CDR1-AS / CDR1 / VSNL1 axis is involved in glioblastoma multiforme. Oncotarget 7:4746–4759

Barbagallo D, Caponnetto A, Cirnigliaro M et al (2018) CircSMARCA5 inhibits migration of glioblastoma multiforme cells by regulating a molecular axis involving splicing factors SRSF1/SRSF3/PTB. Int J Mol Sci 19:480

Barbieri I, Kouzarides T (2020) Role of RNA modifications in cancer. Nat Rev Cancer 20:303–322

Bassi S, Tripathi T, Monziani A et al (2017) Epigenetics of Huntington's disease. Adv Exp Med Biol 978:277–299

Bennett SA, Tanaz R, Cobos SN et al (2019) Epigenetics in amyotrophic lateral sclerosis: a role for histone post-translational modifications in neurodegenerative disease. Transl Res 204:19–30

Bertram L, Tanzi RE (2019) Alzheimer disease risk genes: 29 and counting. Nat Rev Neurol 15:191–192

Biamonti G, Amato A, Belloni E et al (2019) Alternative splicing in Alzheimer's disease. Aging Clin Exp Res. https://doi.org/10.1007/s40520-019-01360-x

Bicchi I, Morena F, Montesano S et al (2013) MicroRNAs and molecular mechanisms of neurodegeneration. Genes (Basel) 4:244–263
Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307
Bogomazova AN, Eremeev AV, Pozmogova GE et al (2019) The role of mutant RNA in the pathogenesis of Huntington's disease and other polyglutamine diseases. Mol Biol (Mosk) 53:954–967
Bohnsack MT, Sloan KE (2018) The mitochondrial epitranscriptome: the roles of RNA modifications in mitochondrial translation and human disease. Cell Mol Life Sci 75:241–260
Botti V, Cesaretti A, Ban Ž et al (2019) Fine structural tuning of styryl-based dyes for fluorescence and CD-based sensing of various ds-DNA/RNA sequences. Org Biomol Chem 17:8243–8258
Botti V, Urbanelli L, Sagini K et al (2020) Quaternized styryl-azinium fluorophores as cellular RNA-binders. Photochem Photobiol Sci 19:362–370
Braggin JE, Bucks SA, Course MM et al (2019) Alternative splicing in a presenilin 2 variant associated with Alzheimer disease. Ann Clin Transl Neurol 6:762–777
Briggs JA, Wolvetang EJ, Mattick JS et al (2015) Mechanisms of long non-coding RNAs in mammalian nervous system development, plasticity, disease, and evolution. Neuron 88:861–877
Brodie S, Lee HK, Jiang W et al (2017) The novel long non-coding RNA TALNEC2, regulates tumor cell growth and the stemness and radiation response of glioma stem cells. Oncotarget 8:31785–31801
Bucha S, Mukhopadhyay D, Bhattacharyya NP (2015) Regulation of mitochondrial morphology and cell cycle by microRNA-214 targeting Mitofusin2. Biochem Biophys Res Commun 465:797–802
Burk K, Pasterkamp RJ (2019) Disrupted neuronal trafficking in amyotrophic lateral sclerosis. Acta Neuropathol 137:859–877
Butti Z, Patten SA (2018) RNA dysregulation in amyotrophic lateral sclerosis. Front Genet 9:712
Cai H, Liu X, Zheng J et al (2017) Long non-coding RNA taurine upregulated 1 enhances tumor-induced angiogenesis through inhibiting microRNA-299 in human glioblastoma. Oncogene 36:318–331
Calzoni E, Cesaretti A, Polchi A et al (2019) Biocompatible polymer nanoparticles for drug delivery applications in cancer and neurodegenerative disorder therapies. J Funct Biomater 10:4
Campos-Melo D, Droppelmann CA, He Z et al (2013) Altered microRNA expression profile in amyotrophic lateral sclerosis: a role in the regulation of NFL mRNA levels. Mol Brain 6:26
Cao M, Li H, Zhao J et al (2019) Identification of age- and gender-associated long noncoding RNAs in the human brain with Alzheimer's disease. Neurobiol Aging 81:116–126
Catanesi M, d'Angelo M, Tupone MG et al (2020) MicroRNAs dysregulation and mitochondrial dysfunction in neurodegenerative diseases. Int J Mol Sci 21:5986
Ceyzériat K, Zilli T, Millet P et al (2020) Learning from the past: a review of clinical trials targeting amyloid, tau and neuroinflammation in Alzheimer's disease. Curr Alzheimer Res 17:112–125
Chanda K, Das S, Chakraborty J et al (2018) Altered levels of long NcRNAs Meg3 and Neat1 in cell and animal models of Huntington's disease. RNA Biol 15:1348–1363
Chen X-Q, Mobley WC (2019) Alzheimer disease pathogenesis: insights from molecular and cellular biology studies of oligomeric Aβ and Tau species. Front Neurosci 13:659
Chen G, Wang Z, Wang D et al (2013) LncRNADisease: a database for long-non-coding RNA-associated diseases. Nucleic Acids Res 41:D983–D986
Chen Y, Lian Y-J, Ma Y-Q et al (2018) LncRNA SNHG1 promotes α-synuclein aggregation and toxicity by targeting miR-15b-5p to activate SIAH1 in human neuroblastoma SH-SY5Y cells. Neurotoxicology 68:212–221
Chen F-Z, Zhao Y, Chen H-Z (2019a) MicroRNA-98 reduces amyloid β-protein production and improves oxidative stress and mitochondrial dysfunction through the Notch signaling pathway via HEY2 in Alzheimer's disease mice. Int J Mol Med 43:91–102

Chen X, Yu C, Guo M et al (2019b) Down-regulation of m6A mRNA methylation is involved in dopaminergic neuronal death. ACS Chem Neurosci 10:2355–2363
Cheng C, Spengler RM, Keiser MS et al (2018) The long non-coding RNA NEAT1 is elevated in polyglutamine repeat expansion diseases and protects from disease gene-dependent toxicities. Hum Mol Genet 27:4303–4314
Chia R, Chiò A, Traynor BJ (2018) Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. Lancet Neurol 17:94–102
Cho HJ, Liu G, Jin SM et al (2013) MicroRNA-205 regulates the expression of Parkinson's disease-related leucine-rich repeat kinase 2 protein. Hum Mol Genet 22:608–620
Cui Q, Shi H, Ye P et al (2017) m6A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. Cell Rep 18:2622–2634
Dahariya S, Paddibhatla I, Kumar S et al (2019) Long non-coding RNA: classification, biogenesis and functions in blood cells. Mol Immunol 112:82–92
Dai D, Wang H, Zhu L et al (2018) N6-methyladenosine links RNA metabolism to cancer progression. Cell Death Dis 9:124
Das U, Wang L, Ganguly A et al (2016) Visualizing APP and BACE-1 approximation in neurons yields insight into the amyloidogenic pathway. Nat Neurosci 19:55–64
Dash D, Mestre TA (2020) Therapeutic update on Huntington's disease: symptomatic treatments and emerging disease-modifying therapies. Neurotherapeutics. https://doi.org/10.1007/s13311-020-00891-w
De Felice B, Manfellotto F, Fiorentino G et al (2018) Wide-ranging analysis of microRNA profiles in sporadic amyotrophic lateral sclerosis using next-generation sequencing. Front Genet 9:310
Delaidelli A, Jan A, Herms J et al (2019) Translational control in brain pathologies: biological significance and therapeutic opportunities. Acta Neuropathol 137:535–555
DeVos SL, Miller RL, Schoch KM et al (2017) Tau reduction prevents neuronal loss and reverses pathological tau deposition and seeding in mice with tauopathy. Sci Transl Med 9:eaag0481
Dolinar A, Koritnik B, Glavač D et al (2019) Circular RNAs as potential blood biomarkers in amyotrophic lateral sclerosis. Mol Neurobiol 56:8052–8062
Dong Z, Cui H (2020) The emerging roles of RNA modifications in glioblastoma. Cancers (Basel) 12:736
Dorszewska J, Prendecki M, Oczkowska A et al (2016) Molecular basis of familial and sporadic Alzheimer's disease. Curr Alzheimer Res 13:952–963
Dube U, Del-Aguila JL, Li Z et al (2019) An atlas of cortical circular RNA expression in Alzheimer disease brains demonstrates clinical and pathological associations. Nat Neurosci 22:1903–1912
Dykes IM, Emanueli C (2017) Transcriptional and post-transcriptional gene regulation by long non-coding RNA. Genomics Proteomics Bioinformatics 15:177–186
Eom GH, Kim K-B, Kim JH et al (2011) Histone methyltransferase SETD3 regulates muscle differentiation. J Biol Chem 286:34733–34742
Essack M, Salhi A, Van Neste C et al (2020) DES-ROD: exploring literature to develop new links between RNA oxidation and human diseases. Oxid Med Cell Longev 2020:5904315
Esteller M (2011) Non-coding RNAs in human disease. Nat Rev Genet 12:861–874
Esteve-Puig R, Bueno-Costa A, Esteller M (2020) Writers, readers and erasers of RNA modifications in cancer. Cancer Lett 474:127–137
Faghihi MA, Modarresi F, Khalil AM et al (2008) Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. Nat Med 14:723–730
Faghihi MA, Zhang M, Huang J et al (2010) Evidence for natural antisense transcript-mediated inhibition of microRNA function. Genome Biol 11:R56
Fan W, Liang C, Ou M et al (2020) MicroRNA-146a is a wide-reaching neuroinflammatory regulator and potential treatment target in neurological diseases. Front Mol Neurosci 13:90
Fang M, Wang J, Zhang X et al (2012) The miR-124 regulates the expression of BACE1/β-secretase correlated with cell death in Alzheimer's disease. Toxicol Lett 209:94–105

Farr SA, Ripley JL, Sultana R et al (2014) Antisense oligonucleotide against GSK-3β in brain of SAMP8 mice improves learning and memory and decreases oxidative stress: involvement of transcription factor Nrf2 and implications for Alzheimer disease. Free Radic Biol Med 67:387–395

Fernandes JCR, Acuña SM, Aoki JI et al (2019) Long non-coding RNAs in the regulation of gene expression: physiology and disease. Noncoding RNA 5:17

Ferrante M, Conti GO (2017) Environment and neurodegenerative diseases: an update on miRNA role. Microrna 6:157–165

Fisher AJ, Beal PA (2018) Structural basis for eukaryotic mRNA modification. Curr Opin Struct Biol 53:59–68

Flamand MN, Meyer KD (2019) The epitranscriptome and synaptic plasticity. Curr Opin Neurobiol 59:41–48

Fotuhi SN, Khalaj-Kondori M, Hoseinpour Feizi MA et al (2019) Long non-coding RNA BACE1-AS may serve as an Alzheimer's disease blood-based biomarker. J Mol Neurosci 69:351–359

Fransquet PD, Ryan J (2018) Micro RNA as a potential blood-based epigenetic biomarker for Alzheimer's disease. Clin Biochem 58:5–14

Fusilli C, Migliore S, Mazza T et al (2018) Biological and clinical manifestations of juvenile Huntington's disease: a retrospective analysis. Lancet Neurol 17:986–993

Gagliardi D, Dziembowski A (2018) 5' and 3' modifications controlling RNA degradation: from safeguards to executioners. Philos Trans R Soc Lond B Biol Sci 373:20180160

Gagliardi S, Zucca S, Pandini C et al (2018) Long non-coding and coding RNAs characterization in peripheral blood mononuclear cells and spinal cord from amyotrophic lateral sclerosis patients. Sci Rep 8:2378

Galardi S, Michienzi A, Ciafrè SA (2020) Insights into the regulatory role of m6A epitranscriptome in glioblastoma. Int J Mol Sci 21:2816

Garcia HG, Berrocal A, Kim YJ et al (2020) Lighting up the central dogma for predictive developmental biology. Curr Top Dev Biol 137:1–35

Garcia-Manteiga JM, D'Alessandro R, Meldolesi J (2019) News about the role of the transcription factor REST in neurons: from physiology to pathology. Int J Mol Sci 21:235

Ge J, Yu Y-T (2013) RNA pseudouridylation: new insights into an old modification. Trends Biochem Sci 38:210–218

Ge B, Li S-L, Li F-R (2020) Astragaloside-IV regulates endoplasmic reticulum stress-mediated neuronal apoptosis in a murine model of Parkinson's disease via the lincRNA-p21/CHOP pathway. Exp Mol Pathol 115:104478

Geng L, Zhang T, Liu W et al (2018) Inhibition of miR-128 abates Aβ-mediated cytotoxicity by targeting PPAR-γ via NF-κB inactivation in primary mouse cortical neurons and Neuro2a cells. Yonsei Med J 59:1096–1106

Ghosh A, Mizuno K, Tiwari SS et al (2020) Alzheimer's disease-related dysregulation of mRNA translation causes key pathological features with ageing. Transl Psychiatry 10:192

Gillies GE, Pienaar IS, Vohra S et al (2014) Sex differences in Parkinson's disease. Front Neuroendocrinol 35:370–384

Ginns EI, Mak SK-K, Ko N et al (2014) Neuroinflammation and α-synuclein accumulation in response to glucocerebrosidase deficiency are accompanied by synaptic dysfunction. Mol Genet Metab 111:152–162

Gu C, Chen C, Wu R et al (2018) Long noncoding RNA EBF3-AS promotes neuron apoptosis in Alzheimer's disease. DNA Cell Biol 37:220–226

Guo R, Fan G, Zhang J et al (2017) A 9-microRNA signature in serum serves as a noninvasive biomarker in early diagnosis of Alzheimer's disease. J Alzheimers Dis 60:1365–1377

Guo T, Zhang D, Zeng Y et al (2020) Molecular and cellular mechanisms underlying the pathogenesis of Alzheimer's disease. Mol Neurodegener 15:40

Han S, Miller JE, Byun S et al (2019) Identification of exon skipping events associated with Alzheimer's disease in the human hippocampus. BMC Med Genomics 12:13

Han M, Liu Z, Xu Y et al (2020) Abnormality of m6A mRNA methylation is involved in Alzheimer's disease. Front Neurosci 14:98
Hanan M, Simchovitz A, Yayon N et al (2020) A Parkinson's disease CircRNAs resource reveals a link between circSLC8A1 and oxidative stress. EMBO Mol Med 12:e11942
Hartford CCR, Lal A (2020) When long noncoding becomes protein coding. Mol Cell Biol 40: e00528–e00519
Hébert SS, Horré K, Nicolaï L et al (2008) Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/beta-secretase expression. Proc Natl Acad Sci USA 105:6415–6420
Helm M, Motorin Y (2017) Detecting RNA modifications in the epitranscriptome: predict and validate. Nat Rev Genet 18:275–291
Höglinger GU, Michel PP, Champy P et al (2005) Experimental evidence for a toxic etiology of tropical parkinsonism. Mov Disord 20:118–119
Hon C-C, Ramilowski JA, Harshbarger J et al (2017) An atlas of human long non-coding RNAs with accurate 5' ends. Nature 543:199–204
Hsu PJ, Shi H, He C (2017) Epitranscriptomic influences on development and disease. Genome Biol 18:197
Hu S, Wang H, Chen K et al (2015) MicroRNA-34c downregulation ameliorates amyloid–βinduced synaptic failure and memory deficits by targeting VAMP2. J Alzheimers Dis 48:673–686
Huang T, Chen W, Liu J et al (2019a) Genome-wide identification of mRNA 5-methylcytosine in mammals. Nat Struct Mol Biol 26:380–388
Huang Z, Shi J, Gao Y et al (2019b) HMDD v3.0: a database for experimentally supported human microRNA-disease associations. Nucleic Acids Res 47:D1013–D1017
Huang AZ, Delaidelli A, Sorensen PH (2020a) RNA modifications in brain tumorigenesis. Acta Neuropathol Commun 8:64
Huang H, Camats-Perna J, Medeiros R et al (2020b) Altered expression of the m6A methyltransferase METTL3 in Alzheimer's disease. eNeuro 7:ENEURO.0125-20.2020
Jacob R, Zander S, Gutschner T (2017) The dark side of the epitranscriptome: chemical modifications in long non-coding RNAs. Int J Mol Sci 18:2387
Jain G, Stuendl A, Rao P et al (2019) A combined miRNA-piRNA signature to detect Alzheimer's disease. Transl Psychiatry 9:250
Jakubauskienė E, Vilys L, Pečiulienė I et al (2020) The role of hypoxia on Alzheimer's disease-related APP and Tau mRNA formation. Gene 766:145146
Janin M, Ortiz-Barahona V, de Moura MC et al (2019) Epigenetic loss of RNA-methyltransferase NSUN5 in glioma targets ribosomes to drive a stress adaptive translational program. Acta Neuropathol 138:1053–1074
Jiang Q, Shan K, Qun-Wang X et al (2016a) Long non-coding RNA-MIAT promotes neurovascular remodeling in the eye and brain. Oncotarget 7:49688–49698
Jiang X, Yan Y, Hu M et al (2016b) Increased level of H19 long noncoding RNA promotes invasion, angiogenesis, and stemness of glioblastoma cells. J Neurosurg 124:129–136
Jin Y, Tu Q, Liu M (2018) MicroRNA-125b regulates Alzheimer's disease through SphK1 regulation. Mol Med Rep 18:2373–2380
John A, Kubosumi A, Reddy PH (2020) Mitochondrial microRNAs in aging and neurodegenerative diseases. Cells 9:1345
Jonkhout N, Tran J, Smith MA et al (2017) The RNA modification landscape in human disease. RNA 23:1754–1769
Jovicic A, Zaldivar Jolissaint JF, Moser R et al (2013) MicroRNA-22 (miR-22) overexpression is neuroprotective via general anti-apoptotic effects and may also target specific Huntington's disease-related mechanisms. PLoS One 8:e54222
Jung Y, Goldman D (2018) Role of RNA modifications in brain and behavior. Genes Brain Behav 17:e12444
Juźwik CA, Drake S, Zhang Y et al (2019) microRNA dysregulation in neurodegenerative diseases: a systematic review. Prog Neurobiol 182:101664

Kawarada L, Suzuki T, Ohira T et al (2017) ALKBH1 is an RNA dioxygenase responsible for cytoplasmic and mitochondrial tRNA modifications. Nucleic Acids Res 45:7401–7415
Kenny A, Jiménez-Mateos EM, Zea-Sevilla MA et al (2019) Proteins and microRNAs are differentially expressed in tear fluid from patients with Alzheimer's disease. Sci Rep 9:15437
Kilchert C, Sträßer K, Kunetsky V et al (2020) From parts lists to functional significance-RNA-protein interactions in gene regulation. Wiley Interdiscip Rev RNA 11:e1582
Kim W, Lee Y, McKenna ND et al (2014) miR-126 contributes to Parkinson's disease by dysregulating the insulin-like growth factor/phosphoinositide 3-kinase signaling. Neurobiol Aging 35:1712–1721
Kirchner S, Ignatova Z (2015) Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. Nat Rev Genet 16:98–112
Kong Y, Wu J, Zhang D et al (2015) The role of miR-124 in drosophila Alzheimer's disease model by targeting delta in notch signaling pathway. Curr Mol Med 15:980–989
Kong B, Wu P-C, Chen L et al (2016) microRNA-7 protects against 1-methyl-4-phenylpyridinium iodide-induced cell apoptosis in SH-SY5Y cells by directly targeting Krüpple-like factor 4. DNA Cell Biol 35:217–225
Kraus TFJ, Haider M, Spanner J et al (2017) Altered long noncoding RNA expression precedes the course of Parkinson's disease-a preliminary report. Mol Neurobiol 54:2869–2877
Kumar A, Kumar V, Singh K et al (2020) Therapeutic advances for Huntington's disease. Brain Sci 10:43
Lafontaine DLJ (2015) Noncoding RNAs in eukaryotic ribosome biogenesis and function. Nat Struct Mol Biol 22:11–19
Lant JT, Berg MD, Heinemann IU et al (2019) Pathways to disease from natural variations in human cytoplasmic tRNAs. J Biol Chem 294:5294–5308
Le Gras S, Keime C, Anthony A et al (2017) Altered enhancer transcription underlies Huntington's disease striatal transcriptional signature. Sci Rep 7:42875
Lee A, Gilbert RM (2016) Epidemiology of Parkinson disease. Neurol Clin 34:955–965
Leidinger P, Backes C, Deutscher S et al (2013) A blood based 12-miRNA signature of Alzheimer disease patients. Genome Biol 14:R78
Leighton LJ, Bredy TW (2018) Functional interplay between small non-coding RNAs and RNA modification in the brain. Noncoding RNA 4:15
Lentini JM, Ramos J, Fu D (2018) Monitoring the 5-methoxycarbonylmethyl-2-thiouridine (mcm5s2U) modification in eukaryotic tRNAs via the γ-toxin endonuclease. RNA 24:749–758
Lewis CJT, Pan T, Kalsotra A (2017) RNA modifications and structures cooperate to guide RNA-protein interactions. Nat Rev Mol Cell Biol 18:202–210
Li J, Liu C (2019) Coding or noncoding, the converging concepts of RNAs. Front Genet 10:496
Li S, Lv X, Zhai K et al (2016) MicroRNA-7 inhibits neuronal apoptosis in a cellular Parkinson's disease model by targeting Bax and Sirt2. Am J Transl Res 8:993–1004
Li L, Liu H, Song H et al (2017) Let-7d microRNA attenuates 6-OHDA-induced injury by targeting caspase-3 in MN9D cells. J Mol Neurosci 63:403–411
Li D, Yang H, Ma J et al (2018a) MicroRNA-30e regulates neuroinflammation in MPTP model of Parkinson's disease by targeting Nlrp3. Hum Cell 31:106–115
Li H, Ren Y, Mao K et al (2018b) FTO is involved in Alzheimer's disease by targeting TSC1-mTOR-Tau signaling. Biochem Biophys Res Commun 498:234–239
Li L, Xu J, Wu M et al (2018c) Protective role of microRNA-221 in Parkinson's disease. Bratisl Lek Listy 119:22–27
Li J, Chen W, Yi Y et al (2019) miR-219-5p inhibits tau phosphorylation by targeting TTBK1 and GSK-3β in Alzheimer's disease. J Cell Biochem 120:9936–9946
Li D, Zhang J, Li X et al (2020a) Insights into lncRNAs in Alzheimer's disease mechanisms. RNA Biol 1–11
Li Y, Fan H, Sun J et al (2020b) Circular RNA expression profile of Alzheimer's disease and its clinical significance as biomarkers for the disease risk and progression. Int J Biochem Cell Biol 123:105747

Liguori M, Nuzziello N, Introna A et al (2018) Dysregulation of microRNAs and target genes networks in peripheral blood of patients with sporadic amyotrophic lateral sclerosis. Front Mol Neurosci 11:288

Liu Y, Lu Z (2018) Long non-coding RNA NEAT1 mediates the toxic of Parkinson's disease induced by MPTP/MPP+ via regulation of gene expression. Clin Exp Pharmacol Physiol 45:841–848

Liu Z, Wang C, Wang X, Xu S (2015) Therapeutic effects of transplantation of As-MiR-937-expressing mesenchymal stem cells in murine model of Alzheimer's disease. Cell Physiol Biochem 37:321–330

Liu S, Cui B, Dai Z et al (2016) Long non-coding RNA HOTAIR promotes Parkinson's disease induced by MPTP through up-regulating the expression of LRRK2. Curr Neurovasc Res 13:115–120

Liu D, Tang H, Li X-Y et al (2017a) Targeting the HDAC2/HNF-4A/miR-101b/AMPK pathway rescues tauopathy and dendritic abnormalities in Alzheimer's disease. Mol Ther 25:752–764

Liu W, Zhang Q, Zhang J et al (2017b) Long non-coding RNA MALAT1 contributes to cell apoptosis by sponging miR-124 in Parkinson disease. Cell Biosci 7:19

Liu H, Xu Y, Yao B et al (2020a) A novel N6-methyladenosine (m6A)-dependent fate decision for the lncRNA THOR. Cell Death Dis 11:613

Liu L, Wang Y, Wu J et al (2020b) N6-methyladenosine: a potential breakthrough for human cancer. Mol Ther Nucleic Acids 19:804–813

Liu Z, Chen X, Li Z et al (2020c) Role of RNA oxidation in neurodegenerative diseases. Int J Mol Sci 21:5022

Livneh I, Moshitch-Moshkovitz S, Amariglio N et al (2020) The m6A epitranscriptome: transcriptome plasticity in brain development and function. Nat Rev Neurosci 21:36–51

Lockhart J, Canfield J, Mong EF et al (2019) Nucleotide modification alters microRNA-dependent silencing of microRNA switches. Mol Ther Nucleic Acids 14:339–350

Long JM, Holtzman DM (2019) Alzheimer disease: an update on pathobiology and treatment strategies. Cell 179:312–339

Long JM, Ray B, Lahiri DK (2014) MicroRNA-339-5p down-regulates protein expression of β-site amyloid precursor protein-cleaving enzyme 1 (BACE1) in human primary brain cultures and is reduced in brain tissue specimens of Alzheimer disease subjects. J Biol Chem 289:5184–5198

Long JM, Maloney B, Rogers JT et al (2019) Novel upregulation of amyloid-β precursor protein (APP) by microRNA-346 via targeting of APP mRNA 5'-untranslated region: implications in Alzheimer's disease. Mol Psychiat 24:345–363

Lontay B, Kiss A, Virág L et al (2020) How do post-translational modifications influence the pathomechanistic landscape of Huntington's disease? A comprehensive review. Int J Mol Sci 21:4282

Lukiw WJ, Zhao Y, Cui JG (2008) An NF-kappaB-sensitive micro RNA-146a-mediated inflammatory circuit in Alzheimer disease and in stressed human brain cells. J Biol Chem 283:31315–31322

Magini A, Polchi A, Tozzi A et al (2015) Abnormal cortical lysosomal β-hexosaminidase and β-galactosidase activity at post-synaptic sites during Alzheimer's disease progression. Int J Biochem Cell Biol 58:62–70

Marsh J, Alifragis P (2018) Synaptic dysfunction in Alzheimer's disease: the effects of amyloid beta on synaptic vesicle dynamics as a novel target for therapeutic intervention. Neural Regen Res 13:616–623

Martino S (2015) Therapy for lysosomal storage disorders: a matter for stem cells. Int J Stem Cell Res Transplant 3:1–2

Martino S, Morena F, Barola C et al (2014) Proteomics and epigenetic mechanisms in stem cells. Current Proteomics 11:193–209

Mathis S, Goizet C, Soulages A et al (2019) Genetics of amyotrophic lateral sclerosis: a review. J Neurol Sci 399:217–226

Meng S, Zhou H, Feng Z et al (2019) Epigenetics in neurodevelopment: emerging role of circular RNA. Front Cell Neurosci 13:327
Migliore S, Jankovic J, Squitieri F (2019) Genetic counseling in Huntington's disease: potential new challenges on horizon? Front Neurol 10:453
Millan MJ (2017) Linking deregulation of non-coding RNA to the core pathophysiology of Alzheimer's disease: an integrative review. Prog Neurobiol 156:1–68
Modi PK, Jaiswal S, Sharma P (2016) Regulation of neuronal cell cycle and apoptosis by microRNA 34a. Mol Cell Biol 36:84–94
Mongelli A, Atlante S, Bachetti T et al (2020) Epigenetic signaling and RNA regulation in cardiovascular diseases. Int J Mol Sci 21:509
Morena F, Argentati C, Trotta R et al (2017) A comparison of lysosomal enzymes expression levels in peripheral blood of mild- and severe-Alzheimer's disease and MCI patients: implications for regenerative medicine approaches. Int J Mol Sci 18:1806
Morena F, Argentati C, Bazzucchi M et al (2018) Above the epitranscriptome: RNA modifications and stem cell identity. Genes (Basel) 9:329
Morena F, Oikonomou V, Argentati C et al (2019) Integrated computational analysis highlights unique miRNA signatures in the subventricular zone and striatum of GM2 gangliosidosis animal models. Int J Mol Sci 20:3179
Morena F, Argentati C, Acquati S et al (2020) Toward reference intervals of ARSA activity in the cerebrospinal fluid: implication for the clinical practice of metachromatic leukodystrophy. J Appl Lab Med. https://doi.org/10.1093/jalm/jfaa108
Nalls MA, Blauwendraat C, Vallerga CL et al (2019) Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet Neurol 18:1091–1102
Nan Y, Han L, Zhang A et al (2010) MiRNA-451 plays a role as tumor suppressor in human glioma cells. Brain Res 1359:14–21
Neueder A (2019) RNA-mediated disease mechanisms in neurodegenerative disorders. J Mol Biol 431:1780–1791
Neueder A, Bates GP (2018) RNA related pathology in Huntington's disease. Adv Exp Med Biol 1049:85–101
Nishimoto Y, Nakagawa S, Hirose T et al (2013) The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. Mol Brain 6:31
Nourse J, Spada S, Danckwardt S (2020) Emerging roles of RNA 3'-end cleavage and polyadenylation in pathogenesis, diagnosis and therapy of human disorders. Biomolecules 10:915
Noyce AJ, Bestwick JP, Silveira-Moriyama L et al (2012) Meta-analysis of early nonmotor features and risk factors for Parkinson disease. Ann Neurol 72:893–901
Oerum S, Dégut C, Barraud P et al (2017) m1A post-transcriptional modification in tRNAs. Biomolecules 7:20
Orlacchio A, Bernardi G, Orlacchio A et al (2007) RNA interference as a tool for Alzheimer's disease therapy. Mini Rev Med Chem 7:1166–1176
Orlacchio A, Bernardi G, Orlacchio A et al (2008) Patented therapeutic RNAi strategies for neurodegenerative diseases of the CNS. Exp Opin Therap Patents 18:1161–1174
Orozco-Díaz R, Sánchez-Álvarez A, Hernández-Hernández JM et al (2019) The interaction between RE1-silencing transcription factor (REST) and heat shock protein 90 as new therapeutic target against Huntington's disease. PLoS One 14:e0220393
Oskarsson B, Gendron TF, Staff NP (2018) Amyotrophic lateral sclerosis: an update for 2018. Mayo Clin Proc 93:1617–1628
Ozsait B, Komurcu-Bayrak E, Levula M et al (2010) Niemann-Pick type C fibroblasts have a distinct microRNA profile related to lipid metabolism and certain cellular components. Biochem Biophys Res Commun 403:316–321
Pan T (2018) Modifications and functional genomics of human transfer RNA. Cell Res 28:395–404

Parisi C, Napoli G, Amadio S et al (2016) MicroRNA-125b regulates microglia activation and motor neuron death in ALS. Cell Death Differ 23:531–541
Penke B, Szűcs M, Bogár F (2020) Oligomerization and conformational change turn monomeric β-amyloid and tau proteins toxic: their role in Alzheimer's pathogenesis. Molecules 25:1659
Platt FM, Boland B, van der Spoel AC (2012) The cell biology of disease: lysosomal storage disorders: the cellular impact of lysosomal dysfunction. J Cell Biol 199:723–734
Postuma RB, Berg D, Stern M et al (2015) MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord 30:1591–1601
Qin L, Min S, Shu L et al (2020) Genetic analysis of N6-methyladenosine modification genes in Parkinson's disease. Neurobiol Aging 93:143.e9–143.e13
Quan Z, Zheng D, Qing H (2017) Regulatory roles of long non-coding RNAs in the central nervous system and associated neurodegenerative diseases. Front Cell Neurosci 11:175
Queiroz MT, Pereira VG, do Nascimento CC et al (2016) The underexploited role of non-coding RNAs in lysosomal storage diseases. Front Endocrinol (Lausanne) 7:133
Quinlan S, Kenny A, Medina M et al (2017) MicroRNAs in neurodegenerative diseases. Int Rev Cell Mol Biol 334:309–343
Quintavalle C, Garofalo M, Zanca C et al (2012) miR-221/222 overexpession in human glioblastoma increases invasiveness by targeting the protein phosphate PTPμ. Oncogene 31:858–868
Rafels-Ybern À, Torres AG, Grau-Bove X et al (2018) Codon adaptation to tRNAs with Inosine modification at position 34 is widespread among eukaryotes and present in two bacterial phyla. RNA Biol 15:500–507
Ravnik-Glavač M, Glavač D (2020) Circulating RNAs as potential biomarkers in amyotrophic lateral sclerosis. Int J Mol Sci 21:1714
Ren J, Yang Y, Xue J et al (2018) Long noncoding RNA SNHG7 promotes the progression and growth of glioblastoma via inhibition of miR-5095. Biochem Biophys Res Commun 496:712–718
Rinchetti P, Rizzuti M, Faravelli I et al (2018) MicroRNA metabolism and dysregulation in amyotrophic lateral sclerosis. Mol Neurobiol 55:2617–2630
Riva P, Ratti A, Venturin M (2016) The long non-coding RNAs in neurodegenerative diseases: novel mechanisms of pathogenesis. Curr Alzheimer Res 13:1219–1231
Rodriguez CM, Todd PK (2019) New pathologic mechanisms in nucleotide repeat expansion disorders. Neurobiol Dis 130:104515
Romano G, Veneziano D, Nigita G et al (2018) RNA methylation in ncRNA: classes, detection, and molecular associations. Front Genet 9:243
Romo L, Ashar-Patel A, Pfister E et al (2017) Alterations in mRNA 3' UTR isoform abundance accompany gene expression changes in human Huntington's disease brains. Cell Rep 20:3057–3070
Rong D, Sun H, Li Z et al (2017) An emerging function of circRNA-miRNAs-mRNA axis in human diseases. Oncotarget 8:73271–73281
Roundtree IA, Evans ME, Pan T et al (2017) Dynamic RNA modifications in gene expression regulation. Cell 169:1187–1200
Rudich P, Watkins S, Lamitina T (2020) PolyQ-independent toxicity associated with novel translational products from CAG repeat expansions. PLoS One 15:e0227464
Salvatori B, Biscarini S, Morlando M (2020) Non-coding RNAs in nervous system development and disease. Front Cell Dev Biol 8:273
Santa-Maria I, Alaniz ME, Renwick N et al (2015) Dysregulation of microRNA-219 promotes neurodegeneration through post-transcriptional regulation of tau. J Clin Invest 125:681–686
Sarkar S, Jun S, Rellick S et al (2016) Expression of microRNA-34a in Alzheimer's disease brain targets genes linked to synaptic plasticity, energy metabolism, and resting state network activity. Brain Res 1646:139–151
Saudou F, Humbert S (2016) The biology of huntingtin. Neuron 89:910–926

Schymick JC, Scholz SW, Fung H-C et al (2007) Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. Lancet Neurol 6:322–328
Scoles DR, Minikel EV, Pulst SM (2019) Antisense oligonucleotides: a primer. Neurol Genet 5: e323
Selkoe DJ, Hardy J (2016) The amyloid hypothesis of Alzheimer's disease at 25 years. EMBO Mol Med 8:595–608
Sergiev PV, Aleksashin NA, Chugunova AA et al (2018) Structural and evolutionary insights into ribosomal RNA methylation. Nat Chem Biol 14:226–235
Shafik A, Schumann U, Evers M et al (2016) The emerging epitranscriptomics of long noncoding RNAs. Biochim Biophys Acta 1859:59–70
Shannon KM (2020) Recent advances in the treatment of Huntington's disease: targeting DNA and RNA. CNS Drugs 34:219–228
Shao W, Peng D, Wang X (2017) Genetics of Alzheimer's disease: from pathogenesis to clinical usage. J Clin Neurosci 45:1–8
Shen J, Hodges TR, Song R et al (2018) Serum HOTAIR and GAS5 levels as predictors of survival in patients with glioblastoma. Mol Carcinog 57:137–141
Shi J, Wang Y-J, Sun C-R et al (2017) Long noncoding RNA lncHERG promotes cell proliferation, migration and invasion in glioblastoma. Oncotarget 8:108031–108041
Shu M, Zheng X, Wu S et al (2011) Targeting oncogenic miR-335 inhibits growth and invasion of malignant astrocytoma cells. Mol Cancer 10:59
Sierksma A, Lu A, Salta E et al (2018) Deregulation of neuronal miRNAs induced by amyloid-β or TAU pathology. Mol Neurodegener 13:54
Sinha M, Ghose J, Bhattarcharyya NP (2011) Micro RNA -214,-150,-146a and-125b target huntingtin gene. RNA Biol 8:1005–1021
Smith AV, Tabrizi SJ (2020) Therapeutic antisense targeting of huntingtin. DNA Cell Biol 39:154–158
Smith PY, Hernandez-Rapp J, Jolivette F et al (2015) miR-132/212 deficiency impairs tau metabolism and promotes pathological aggregation in vivo. Hum Mol Genet 24:6721–6735
Song J, Yi C (2019) Reading chemical modifications in the transcriptome. J Mol Biol. S0022-2836 (19):30598–30594
Sun W, Samimi H, Gamez M et al (2018) Pathogenic tau-induced piRNA depletion promotes neuronal death through transposable element dysregulation in neurodegenerative tauopathies. Nat Neurosci 21:1038–1048
Swarbrick S, Wragg N, Ghosh S et al (2019) Systematic Review of miRNA as Biomarkers in Alzheimer's disease. Mol Neurobiol 56:6156–6167
Tabrizi SJ, Flower MD, Ross CA et al (2020) Huntington disease: new insights into molecular pathogenesis and therapeutic opportunities. Nat Rev Neurol 16:529–546
Tang Y, Bao JS, Su JH et al (2017) MicroRNA-139 modulates Alzheimer's-associated pathogenesis in SAMP8 mice by targeting cannabinoid receptor type 2. Genet Mol Res 16(1). https://doi.org/10.4238/gmr16019166
Torres AG, Batlle E, Ribas de Pouplana L (2014) Role of tRNA modifications in human diseases. Trends Mol Med 20:306–314
Trixl L, Lusser A (2019) The dynamic RNA modification 5-methylcytosine and its emerging role as an epitranscriptomic mark. Wiley Interdiscip Rev RNA 10:e1510
Urbanelli L, Emiliani C, Massini C et al (2008) Cathepsin D expression is decreased in Alzheimer's disease fibroblasts. Neurobiol Aging 29:12–22
Urbanelli L, Magini A, Polchi A et al (2011) Recent developments in therapeutic approaches for lysosomal storage diseases. Recent Pat CNS Drug Discov 6:1–19
Ursu A, Vézina-Dawod S, Disney MD (2019) Methods to identify and optimize small molecules interacting with RNA (SMIRNAs). Drug Discov Today 24:2002–2016
Valdez G, Heyer MP, Feng G et al (2014) The role of muscle microRNAs in repairing the neuromuscular junction. PLoS One 9:e93140

van den Berg MMJ, Krauskopf J, Ramaekers JG et al (2020) Circulating microRNAs as potential biomarkers for psychiatric and neurodegenerative disorders. Prog Neurobiol 185:101732
van der Lee SJ, Wolters FJ, Ikram MK et al (2018) The effect of APOE and other common genetic variants on the onset of Alzheimer's disease and dementia: a community-based cohort study. Lancet Neurol 17:434–444
van Es MA, Hardiman O, Chio A et al (2017) Amyotrophic lateral sclerosis. Lancet 390:2084–2098
Vandivier LE, Gregory BD (2017) Reading the epitranscriptome: new techniques and perspectives. Enzymes 41:269–298
Varghese S, Cotter M, Chevot F et al (2017) In vivo modification of tRNA with an artificial nucleobase leads to full disease remission in an animal model of multiple sclerosis. Nucleic Acids Res 45:2029–2039
Vassallo I, Zinn P, Lai M et al (2016) WIF1 re-expression in glioblastoma inhibits migration through attenuation of non-canonical WNT signaling by downregulating the lncRNA MALAT1. Oncogene 35:12–21
Vieira AS, Dogini DB, Lopes-Cendes I (2018) Role of non-coding RNAs in non-aging-related neurological disorders. Braz J Med Biol Res 51:e7566
Vilardo E, Barbato C, Ciotti M et al (2010) MicroRNA-101 regulates amyloid precursor protein expression in hippocampal neurons. J Biol Chem 285:18344–18351
Villar-Menéndez I, Porta S, Buira SP et al (2014) Increased striatal adenosine A2A receptor levels is an early event in Parkinson's disease-related pathology and it is potentially regulated by miR-34b. Neurobiol Dis 69:206–214
Visvanathan A, Patil V, Arora A et al (2018) Essential role of METTL3-mediated m6A modification in glioma stem-like cells maintenance and radioresistance. Oncogene 37:522–533
Vrabec K, Boštjančič E, Koritnik B et al (2018) Differential expression of several miRNAs and the host genes AATK and DNM2 in leukocytes of sporadic ALS patients. Front Mol Neurosci 11:106
Wan G, Zhou W, Hu Y et al (2016) Transcriptional regulation of lncRNA genes by histone modification in Alzheimer's disease. Biomed Res Int 2016:3164238
Wan Y-W, Al-Ouran R, Mangleburg CG et al (2020) Meta-analysis of the Alzheimer's disease human brain transcriptome and functional dissection in mouse models. Cell Rep 32:107908
Wang Z-H, Zhang J-L, Duan Y-L et al (2015) MicroRNA-214 participates in the neuroprotective effect of Resveratrol via inhibiting α-synuclein expression in MPTP-induced Parkinson's disease mouse. Biomed Pharmacother 74:252–256
Wang D-Q, Fu P, Yao C et al (2018a) Long non-coding RNAs, novel culprits, or bodyguards in neurodegenerative diseases. Mol Ther Nucleic Acids 10:269–276
Wang Z, Xu P, Chen B et al (2018b) Identifying circRNA-associated-ceRNA networks in the hippocampus of Aβ1-42-induced Alzheimer's disease-like rats using microarray analysis. Aging (Albany NY) 10:775–788
Wei C-W, Luo T, Zou S-S et al (2018) The role of long noncoding RNAs in central nervous system and neurodegenerative diseases. Front Behav Neurosci 12:175
Wei W, Wang Z-Y, Ma L-N et al (2020) MicroRNAs in Alzheimer's disease: function and potential applications as diagnostic biomarkers. Front Mol Neurosci 13:160
Wen Y, Yu WH, Maloney B et al (2008) Transcriptional regulation of beta-secretase by p25/cdk5 leads to enhanced amyloidogenic processing. Neuron 57:680–690
Wu Y-Y, Kuo H-C (2020) Functional roles and networks of non-coding RNAs in the pathogenesis of neurodegenerative diseases. J Biomed Sci 27:49
Wu P, Zuo X, Deng H et al (2013) Roles of long noncoding RNAs in brain development, functional diversification and neurodegenerative diseases. Brain Res Bull 97:69–80
Wu D-M, Wen X, Wang Y-J et al (2018) Effect of microRNA-186 on oxidative stress injury of neuron by targeting interleukin 2 through the Janus kinase-signal transducer and activator of transcription pathway in a rat model of Alzheimer's disease. J Cell Physiol 233:9488–9502
Wu D-M, Wang S, Wen X et al (2019) Suppression of microRNA-342-3p increases glutamate transporters and prevents dopaminergic neuron loss through activating the Wnt signaling

pathway via p21-activated kinase 1 in mice with Parkinson's disease. J Cell Physiol 234:9033–9044
Xie L, Mao M, Xiong K et al (2017) Circular RNAs: a novel player in development and disease of the central nervous system. Front Cell Neurosci 11:354
Xing H, Guo S, Zhang Y et al (2016) Upregulation of microRNA-206 enhances lipopolysaccharide-induced inflammation and release of amyloid-β by targeting insulin-like growth factor 1 in microglia. Mol Med Rep 14:1357–1364
Xiong X, Li X, Yi C (2018) N1-methyladenosine methylome in messenger RNA and non-coding RNA. Curr Opin Chem Biol 45:179–186
Xu X, Wang Y, Liang H (2018) The role of A-to-I RNA editing in cancer development. Curr Opin Genet Dev 48:51–56
Xylaki M, Atzler B, Outeiro TF (2019) Epigenetics of the synapse in neurodegeneration. Curr Neurol Neurosci Rep 19:72
Yan C, Wan R, Shi Y (2019) Molecular mechanisms of pre-mRNA splicing through structural biology of the spliceosome. Cold Spring Harb Perspect Biol 11:a032409
Yang CH, Yue J, Pfeffer SR et al (2014) MicroRNA-21 promotes glioblastoma tumorigenesis by down-regulating insulin-like growth factor-binding protein-3 (IGFBP3). J Biol Chem 289:25079–25087
Yang G, Song Y, Zhou X et al (2015) MicroRNA-29c targets β-site amyloid precursor protein-cleaving enzyme 1 and has a neuroprotective role in vitro and in vivo. Mol Med Rep 12:3081–3088
Yang CP, Zhang ZH, Zhang LH et al (2016) Neuroprotective role of microRNA-22 in a 6-hydroxydopamine-induced cell model of Parkinson's disease via regulation of its target gene TRPM7. J Mol Neurosci 60:445–452
Yang X, Liu M, Li M et al (2020) Epigenetic modulations of noncoding RNA: a novel dimension of Cancer biology. Mol Cancer 19:64
Yeasmin F, Yada T, Akimitsu N (2018) Micropeptides encoded in transcripts previously identified as long noncoding RNAs: a new chapter in transcriptomics and proteomics. Front Genet 9:144
Zendjabil M (2018) Circulating microRNAs as novel biomarkers of Alzheimer's disease. Clin Chim Acta 484:99–104
Zetterberg H, Burnham SC (2019) Blood-based molecular biomarkers for Alzheimer's disease. Mol Brain 12:26
Zhang Z, Cheng Y (2014) miR-16-1 promotes the aberrant α-synuclein accumulation in Parkinson disease via targeting heat shock protein 70. Sci World J 2014:938348
Zhang C, Jia G (2018) Reversible RNA modification N1-methyladenosine (m1A) in mRNA and tRNA. Genomics Proteomics Bioinf 16:155–161
Zhang C-Z, Zhang J-X, Zhang A-L et al (2010) MiR-221 and miR-222 target PUMA to induce cell survival in glioblastoma. Mol Cancer 9:229
Zhang J, Hu M, Teng Z et al (2014) Synaptic and cognitive improvements by inhibition of 2-AG metabolism are through upregulation of microRNA-188-3p in a mouse model of Alzheimer's disease. J Neurosci 34:14919–14933
Zhang B, Chen C-F, Wang A-H et al (2015a) MiR-16 regulates cell death in Alzheimer's disease by targeting amyloid precursor protein. Eur Rev Med Pharmacol Sci 19:4020–4027
Zhang Y, Han D, Wei W et al (2015b) MiR-218 inhibited growth and metabolism of human glioblastoma cells by directly targeting E2F2. Cell Mol Neurobiol 35:1165–1173
Zhang B, Wang L-L, Ren R-J et al (2016) MicroRNA-146a represses LRP2 translation and leads to cell apoptosis in Alzheimer's disease. FEBS Lett 590:2190–2200
Zhang S, Zhu D, Li H et al (2017) Characterization of circRNA-associated-ceRNA networks in a senescence-accelerated mouse prone 8 brain. Mol Ther 25:2053–2061
Zhang Y-H, Bai S-F, Yan J-Q (2019) Blood circulating miRNAs as biomarkers of Alzheimer's disease: a systematic review and meta-analysis. Biomark Med 13:1045–1054
Zhang Y, Xia Q, Lin J (2020) LncRNA H19 attenuates apoptosis in MPTP-Induced Parkinson's disease through regulating miR-585-3p/PIK3R3. Neurochem Res 45:1700–1710

Zhao Z-B, Wu L, Xiong R et al (2014) MicroRNA-922 promotes tau phosphorylation by downregulating ubiquitin carboxy-terminal hydrolase L1 (UCHL1) expression in the pathogenesis of Alzheimer's disease. Neuroscience 275:232–237

Zhao Y, Alexandrov PN, Jaber V et al (2016) Deficiency in the ubiquitin conjugating enzyme UBE2A in Alzheimer's disease (AD) is linked to deficits in a natural circular miRNA-7 sponge (circRNA; ciRS-7). Genes (Basel) 7:116

Zhao Y, Dunker W, Yu Y-T et al (2018) The role of noncoding RNA pseudouridylation in nuclear gene expression events. Front Bioeng Biotechnol 6:8

Zhao M-Y, Wang G-Q, Wang N-N et al (2019) The long-non-coding RNA NEAT1 is a novel target for Alzheimer's disease progression via miR-124/BACE1 axis. Neurol Res 41:489–497

Zhou C, Molinie B, Daneshvar K et al (2017) Genome-wide maps of m6A circRNAs identify widespread and cell-type-specific methylation patterns that are distinct from mRNAs. Cell Rep 20:2262–2276

Zhou Y, Wang Z-F, Li W et al (2018) Protective effects of microRNA-330 on amyloid β-protein production, oxidative stress, and mitochondrial dysfunction in Alzheimer's disease by targeting VAV1 via the MAPK signaling pathway. J Cell Biochem 119:5437–5448

Zhou Y, Deng J, Chu X et al (2019) Role of post-transcriptional control of calpain by miR-124-3p in the development of Alzheimer's disease. J Alzheimers Dis 67:571–581

Zhu Z-M, Huo F-C, Pei D-S (2020) Function and evolution of RNA N6-methyladenosine modification. Int J Biol Sci 16:1929–1940

Zimmer-Bensch G (2019) Emerging roles of long non-coding RNAs as drivers of brain evolution. Cells 8:1399

Epitranscriptomic Signatures in Neural Development and Disease

Shikha Sharma and Nibedita Lenka

Contents

Abstract Epitranscriptomics, reminiscent of RNA epigenetics, involves post-transcriptional modifications occurring across coding and non-coding RNAs. Emergence of epitranscriptomic regulatory machinery underlying various developmental processes has facilitated exploring the intricacies involved during the same. Recent

S. Sharma · N. Lenka (✉)
National Centre for Cell Science, Pune, Maharashtra, India
e-mail: nibedita@nccs.res.in

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_3

advancements in the field have led to the identification of more than 160 RNA modifications reflecting epitranscriptome dynamism. Undoubtedly, this has added another layer of fine-tuned regulation to the existing gene regulatory machinery. Despite the progress being made in identifying the epitranscriptomic signature and the players involved, a concrete understanding on their distribution and functional relevance during mammalian development and cell fate decision still remains elusive. In fact, evidences are also emerging pertaining to aberrant RNA modifications and their association with various developmental defects and diseases, including that of cancer. Brain, among the vital organs has been reported to be sensitive to epitranscriptomic modifications, where RNA modifications do have a key role to play in modulating various developmental and physiological aspects ranging generation of neurons and glial cells, synaptic transmission, learning, and memory. In addition, those have also been shown to be contributing to tumour heterogeneity, tumour progression, and suppression. In this chapter, we shall be discussing about the current status on epitranscriptomic machinery operational during neural development and its deregulation manifesting neurological disorders and cancer with special emphasis on methylation of adenine and cytosine bases at position 6 and 5, respectively.

Keywords RNA modification · Neuro-epitranscriptomics · m^6A · m^5C · m^1A · 2′-OMe · Pseudouridine

1 Introduction

Epitranscriptomics or RNA epigenetics is a post-transcriptional chemical modification present on coding and non-coding RNAs important for RNA fate regulation. Indeed, a large number of chemical modifications of RNA have been identified till date across all types of RNAs. While more than 160 modifications have been found in human ribosomal RNA (rRNA) that function as quality control checkpoints in ribosome assembly, transfer RNA (tRNA) in human encompasses ~14 modifications contributing to tRNA stability, tRNA folding, decoding capacity, and decoding accuracy (El Yacoubi et al. 2012; Liu and Pan 2015a). High-throughput sequencing methods have also revealed the dynamic RNA epigenetic nature of various mRNA and tRNA modifications such as 5-methylcytosine (m^5C), N6-methyladenosine (m^6A), 2′-*O* methylation (2′-OMe), and pseudouridine (ψ) (Angelova et al. 2018; Ramos and Fu 2019). Among these, m^6A is one of the best studied and most abundant mRNA modifications involved in gene regulation (Zheng et al. 2013; Liu and Pan 2015a; Angelova et al. 2018). Recent reports have also suggested mRNA modifications playing a crucial role in mediating mRNA metabolism, including transport, decay, translation, and splicing (Zhao et al. 2017). Although various reports have shed light on the understanding of the role of these modifications in RNA metabolism, yet this field remains largely unexplored (Hsu et al.

2017a). Moreover, errors in RNA modification can also have detrimental effects in various processes such as stem cell maintenance, angiogenesis, development, and gametogenesis and can also lead to cancer (Niu et al. 2013; Miao et al. 2019). In a recent review, we have consolidated the RNA modifications and their structure based on the reports by various groups, specifying the role of prominent RNA modifications in cell fate transition (Haran and Lenka 2019).

2 RNA Modifications and Neurogenesis

The brain is one of the most complex and vital organs in the body controlling almost all the necessary functions in an organism. Its development commences with neuroectoderm specification during gastrulation that further proceeds in a hierarchical manner with generation of neural stem cells (NSCs), progenitors, and their mature derivatives such as neurons and glia. In the adult mammalian brain too, NSCs are present in discrete pockets in the subventricular zone and subgranular zone of the hippocampus possessing the characteristics of self-renewal and multipotent differentiation capacity generating neurons, astrocytes, and oligodendrocytes (Kase et al. 2020). These neural cells integrate themselves into existing neural circuitry to perform various functions of the brain, such as learning and memory (Kase et al. 2020). Brain development encompasses precise and fine-tuned regulatory cascade, alteration/defect in which leads to developmental anomaly and cessation in vital body functions. Various studies have stated that epigenetic modifications pertaining to histone modifications, DNA methylation, and non-coding RNAs play a crucial role in mediating neuronal development and neurogenesis (Keverne et al. 2015; Yao et al. 2016). In recent years another layer of the regulatory module has been depicted, showing RNA modifications such as m^6A, m^5C occurring widely in mRNAs of the mammalian nervous system (Meyer et al. 2012; Li et al. 2017). The same is true in the case of tRNAs as well (Ramos and Fu 2019). Undoubtedly, these modifications do influence the regulation of gene expression during both embryonic and postnatal neural development. Moreover, some of the key players such as YTHDF1 involved in mediating these modifications have also been implicated in playing key roles in neurological development, learning, and memory (Shi et al. 2018). Hence understanding the epitranscriptomic machinery would shed light on the critical regulatory cascade underlying brain development and its physiology.

2.1 m^6A Modification

m^6A, is usually found on prokaryotic DNA, but is rare in eukaryotic DNA and controversial in mammals (Luo et al. 2015). In contrast, mRNAs and long non-coding RNAs (lncRNAs) harbour abundant m^6A modification in both prokaryotes and eukaryotes and also across species ranging from rodents to humans

(Desrosiers et al. 1975; Wei et al. 2018). m^6A exhibits conserved pattern with its prevalence noted mostly within long exons, transcription termination sites, and 3′ UTR region and to a lesser extent in 5′ UTRs (Dominissini et al. 2012; Meyer et al. 2012, 2015; Schwartz et al. 2014; García-Campos et al. 2019; Zhang et al. 2019). Transcriptome-wide mapping of m^6A modification has demonstrated that, m^6A is found on the adenine of DRACH consensus motif where D can be adenine, guanine, or Uracil, and H refers to adenine, uracil, or cytosine (Noack and Calegari 2018). Although m^6A modification is seen across multiple tissues in mammals, including that in cancer, it varies with developmental stages and has been shown to be upregulated during embryonic and postnatal development with its higher preponderance seen in the adult brain and ovary as compared with the other tissues (Meyer et al. 2012; Noack and Calegari 2018; Yoon et al. 2017). Incidentally, increased m^6A methylation promotes reprogramming of mouse embryonic fibroblast (MEFs) to pluripotent stem cells (PSCs), while reduced m^6A methylation impedes the same process (Chen et al. 2015). Moreover, m^6A methylation pattern also changes upon environmental stimuli such as stress (Dominissini et al. 2012; Zhou et al. 2015). Together, these studies suggest the significance of m^6A during organismal development and alteration in which might cause anomaly during the same, including that of cancer (Meyer and Jaffrey 2014; Geula et al. 2015; Haran and Lenka 2019).

2.1.1 Regulators of m^6A Modification

m^6A modification is reversible and mediated by various players designated as writer, eraser, and reader proteins. While writers facilitate carving the modification through their methyltransferase activity, the readers recognize the structural changes in RNA, and the erasers remove the same (Roundtree et al. 2017a, b). The regulators of m^6A modification are described below and are summarized in Table 1.

Writers

m^6A modification is co-transcriptionally executed by writers, a large complex of methyltransferases, having three components: (1) METTL3 (methyltransferase like-3), (2) METTL14 (methyltransferase like-14), and (3) WTAP (Wilms' tumour 1-associating protein) that transfer the methyl group of *S*-adenosylmethionine (SAM) to adenine (Noack and Calegari 2018). While METTL3 possesses catalytic function and acts as SAM binding component, METTL14 and Wilms' tumour 1-associating protein (WTAP) serve as RNA binding scaffold. In fact, METTL14, despite being homologous to METTL3, does not bind to the SAM domain and hence does not contribute independently to methyltransferase function. Never-the-less, biochemical studies have revealed that both METTL3 and METTL14 proteins interact with each other at a stoichiometric ratio of 1:1 and that eventually leads to the enhancement of methylation activity rather than METTL3 acting alone (Liu et al. 2014; Wang et al. 2016). On the other hand, WTAP is a regulatory subunit of the

Table 1 Regulators of m^6A and m^5C RNA modifications and the functional attributes

RNA modification	Regulators/constitute			Role/function	References
m^6A	*Writer*	*Eraser*	*Reader*		
	METTL3			*S*-adenosyl methionine (SAM) binding component, mRNA stability, and splicing	Pendleton et al. (2017), Noack and Calegari (2018)
	Wilms' tumour 1-associating protein (WTAP)			Alternative splicing patterns	Ping et al. (2014)
	WTAP + co-factors KIAA1429 (VIRMA), ZC3H13, RBM15/ RBM15B			m^6A methylation specificity	Ping et al. (2014), Knuckles et al. (2018)
	METTL14 and WTAP			RNA binding scaffold	Liu et al. (2014), Wang et al. (2016)
	METTL16			m^6A mRNA Methylation RNA stability and splicing	Pendleton et al. (2017)
		Fat mass and obesity-associated protein (FTO)		Oxidization of m^6A and generation of metastable hm^6A and f^6A that further gets converted to adenosine	Fu et al. (2013, 2014b)
		ALKBH5		mRNA export and stability	Zheng et al. (2013)
			YTHDF1	Modulation of mRNA translation efficiency through interaction with eukaryotic initiation factor 3 (eIF3)	Liao et al. (2018)
			YTHDF2	Targeting mRNAs decay by recruiting CCR4-NOT deadenylase complex	Liao et al. (2018)
			YTHDC1	Splicing, epigenetic silencing, and RNA nuclear export	Patil et al. (2017), Xiang et al. (2017), Zhao et al. (2019)

(continued)

Table 1 (continued)

RNA modification	Regulators/constitute			Role/function	References
			YTHDC2	RNA stability	Kretschmer et al. (2018)
			HNRNPC	Pre-mRNA processing, Alternative splicing, RNA maturation	Liu et al. (2015)
			HNRNPG	Pre-mRNA processing, RNA Polymerase-II occupancy, Alternative splicing	Liu et al. (2017), Zhou et al. (2019)
			HNRNPA2B1	Alternative splicing Pri-miRNA processing	Liu et al. (2017)
			IGBP2	mRNA stability	Huang et al. (2018)
m^5C and hm^5C	NSUN1, NSUN5			rRNA methylation and ribosome biogenesis	Sharma et al. (2013), Gigova et al. (2014)
	NSUN2			Mitotic spindle stability, Cell Cycle progression, Protein synthesis	Gkatza et al. (2019), Xue et al. (2020)
	NSUN3			Promote mitochondrial tRNA methylation ($mt\text{-}tRNA^{Met}$) and alter translation, ESCs differentiation	Van Haute et al. (2016), Trixl et al. (2018)
	NSUN4			Promote mitochondrial ribosome assembly through interaction with MTERF4	Camara et al. (2011)
	NSUN6			Methylation of $tRNA^{Cys}$ and $tRNA^{Thr}$ and ribosome biogenesis	Haag et al. (2015)
	DNMT2			rRNA and tRNA stability, function and biogenesis, Translation	Goll et al. (2006), Tuorto et al. (2012)

		10–11 translocation (TET)		Translational efficiency	Fu et al. (2014a), Delatte et al. (2016)
			ALYREF	mRNA export	Shi et al. (2017), Yang et al. (2017)
			YBF1	mRNA stability, modulation of mRNA metabolism through ELAVL1 interaction	Chen et al. (2019b), Yang et al. (2019)

methyltransferase complex that interacts with both METTL3 and METTL14. Incidentally, deletion of *Wtap* is associated with the reduction in the level of m^6A methylation more efficiently than *Mettl3* and *Mettl14* knockdown. Moreover, WTAP bound genes exhibit alteration in alternative splicing patterns, and WTAP along with its co-factors KIAA1429 (VIRMA), ZC3H13, and RBM15/RBM15B mediate the recruitment of methylase transferase complex in nuclear speckles and determine m^6A methylation specificity for transcripts (Ping et al. 2014; Knuckles et al. 2018). Unlike METTL3 and METTL14, METTL16 is another SAM-dependent m^6A mRNA methyltransferase that methylates adenosine outside the DRACH sequence motif prominently at the longer conserved sequence of UACAGAGAA, with a neighbouring GU motif as its consensus sequence (Pendleton et al. 2017; Warda et al. 2017). METTL16 binds to mRNA, rRNA, U6 spliceosomal RNA, MAT2A mRNA encoding SAM synthase, lncRNAs including 3′ triple helix of metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) and X-inactive specific transcript and contributes to mRNA stability and splicing (Pendleton et al. 2017; Warda et al. 2017). Interestingly, METTL16 functions as both writer and reader. As a writer, it methylates MAT2A mRNA in the presence of SAM, resulting in the retention of intron and nuclear degradation. On the contrary, during low SAM levels, METTL16 occupies MAT2A mRNA for longer period and hence promotes the splicing of retained introns (Pendleton et al. 2017).

A recent report has stated the association between m^6A modifications and histone H3 trimethylation at Lys36 (H3K36me3) with m^6A modification seen in the vicinity of H3K36me3 peaks (Huang et al. 2019). Indeed the frequency of m^6A occurrence was seen decreasing upon H3K36me3 hypomethylation/depletion and hence suggesting histone modification might guide m^6A modification co-transcriptionally. H3K36me3 is regarded as the marker for transcription elongation that is enriched in the CDS region near the 3′end. METTL14, which is a known m^6A writer, functions as a reader for H3K36me3, and the interaction between the two, in turn, facilitates the recruitment of m^6A methyltransferase complex on newly synthesized RNAs and further deposition of m^6A in those (Huang et al. 2019). Similarly, another study has also shown m^6A regulating histone modifications such as histone acetylation (H3K27ac) and histone methylation (H3K4me3, H3K27me3) in NSCs by modulating the stability of transcripts of histone modifiers (Wang et al. 2018). These studies do unequivocally emphasize an interesting cross-talk existing between histone modification and RNA modification in the gene regulatory landscape and reveal a mechanism for the methylation of mammalian transcriptomes.

Concerning the regulation of m^6A methylation, miRNAs have been suggested to modulate the same via sequence pairing mechanism by altering the binding of METTL3, the writer, to the transcripts containing miRNA binding sites (Chen et al. 2015). Moreover, m^6A modification of 3′UTR prevents the binding of miRNA to their targets. While ablation of *Mettl3* in epiblast and naïve embryonic stem cells (ESCs) leads to embryonic lethality in mice post-implantation (Geula et al. 2015), *Mettl14* conditional knockout in mouse embryos remains associated with a decrease in body size and postnatal lethality (Yoon et al. 2017). METTL3 was also shown to regulate the expression of RNA editing enzymes ADAR and APOBEC3A

and altered A-to-I and C-to-U editing events (Visvanathan et al. 2019). In addition, m^6A methylation is crucial for the expression of long intergenic non-coding RNAs (lincRNAs) too.

Erasers

These are the players which modulate RNA modifications by counterbalancing the effect of writers. Demethylases that include fat mass and obesity-associated protein (FTO) and AlkB homologue 5 (ALKBH5) function as erasers and participate in removing the epitranscriptomic marks (Jia et al. 2011; Zheng et al. 2013). While FTO oxidizes m^6A to produce metastable N6-hydroxymethyladenosine (hm^6A) and N6-formyladenosine (f^6A), which further gets converted to adenosine (Fu et al. 2013, 2014b), ALKBH5 reverses m^6A directly to adenine (Zheng et al. 2013). Recently, various reports have indicated FTO independent m^6A modification, which are basically based on the study showing lack of increase in m^6A methylation in FTO knockout mouse embryos (Hess et al. 2013; Mauer et al. 2017; García-Campos et al. 2019). Instead, FTO was found to preferentially demethylate N^6,2′-*O*-dimethyladenosine (m^6Am) in an m^7G cap-dependent manner and regulate the stability of m^6Am mRNA by conferring resistance to decapping enzyme DCP2 (Mauer et al. 2017). In a subsequent report by the same group, it has been demonstrated that FTO controls the central step of small nuclear RNAs (snRNAs) processing by selectively demethylating its m^2Am and dimethylated m2 snRNA, m^6Am methylating isoforms in m^7G cap-dependent manner (Mauer et al. 2019). FTO binds to multiple RNA species, including mRNA, tRNA, and snRNA, and can demethylate cap m^6Am and internal m^6A methylation in mRNA, cap and internal m^6Am methylation in snRNA, internal m^6A methylation in U6 RNA, and N^1-methyladenosine (m^1A) methylation in tRNA (Wei et al. 2018). Interestingly though, FTO has been shown to primarily demethylate m^6Am without any effect on m^6A levels throughout the transcriptome, seen in response to either its overexpression or depletion in human ESCs and HEK293T cells (García-Campos et al. 2019). In contrast, the overexpression of ALKBH5 has been shown to decrease the m^6A methylation level, thereby indicating ALKBH5 as primary demethylating agent for m^6A modification. In fact, ALKBH5 affects mRNA export by colocalizing with mRNA-processing factors including phosphorylated SC35 (SC35-pi) (serine/arginine-rich splicing factor 2), ASF/SF2 (alternative splicing factor/splicing factor 2), and SM (Smith antigen) in nuclear speckles, where some of these proteins function as both splicing factors and export adaptor proteins (Zheng et al. 2013). Further, ALKBH5 has also been shown to be important for maintaining RNA stability, suggesting the importance of demethylation activity for RNA metabolism (Zheng et al. 2013). The same group has also demonstrated the association of ALKBH5 with male infertility by regulating spermatogenesis, where ALKBH5 mediated demethylation was seen as essential for maintaining the correct splicing and production of longer 3′ UTR mRNAs in the nuclei of spermatocytes and round spermatids (Zheng et al. 2013). Hence, the stated players do contribute to cell

specific and developmental event specific co- and post-transcriptional modulations in a precise manner.

Readers

Similar to epigenetics, m^6A modification also regulates the RNA molecules in either cis- or trans modes. While in cis mode, m^6A exerts its effect on the RNA structure by destabilizing Watson–Crick A: U base pairing similar to the epigenetic marks on the nucleosome, m^6A in trans mode mediates RNA function through the recruitment of m^6A reader protein complexes characterized mostly by YT-Homology (YTH) domain (Liao et al. 2018). Several YTH domain-containing proteins have been found in humans and that include YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2. Among these YTHDF2 was the first reader protein identified and well-studied so far. In fact, all YTH domain proteins are ubiquitously expressed except YTHDC2, which is found in testis (Hsu et al. 2017b). Both YTHDF1 and YTHDF2 binds to m^6A single-stranded RNA stronger than DNA, while YTHDC1 exerts opposite effect by binding DNA at m^6A more strongly than RNA (Woodcock et al. 2020). Concerning the function, YTH-domain-containing proteins control the fate of RNA at post-transcriptional level by regulating translation, splicing, localization, and stability (Zhao et al. 2019). Between YTHDF1 and YTHDF2, both exert contrasting influence on the transcribed mRNA. While binding of YTHDF1 to m^6A sites on mRNA promotes translation efficiency by directly interacting with eukaryotic initiation factor 3 (eIF3), YTHDF2 participates in the decay of targeted mRNAs by recruiting CCR4-NOT deadenylase complex (Liao et al. 2018). This indicates an interesting interactor-dependent contrasting influence of YTH-domain-containing proteins on RNA fate. YTHDF2 is implicated in the degradation of m^6A containing R-loops, which leads to growth retardation and increase in the level of γH2AX (Abakir et al. 2020). Indeed, the stability of the YTHDF2 protein depends on CDK1 activity. Together, CDK1, YTHDF2, and WEE1 constitute feedforward regulatory loop to enhance the mitotic entry. However, components of E3 ligase such as Cullin 1 (CUL1), Cullin 4A (CUL4A), damaged DNA-binding protein 1 (DDB1), and S-phase kinase-associated protein 2 (SKP2) mediate proteolysis of YTHDF2 (Fei et al. 2020).

The m^6A reader YTHDC1 in the nucleus has been implicated in various processes such as splicing, epigenetic silencing by non-coding RNA XIST and in RNA nuclear export (Patil et al. 2017; Xiang et al. 2017; Liao et al. 2018; Zhao et al. 2019). As seen in HeLa cells, YTHDC1 mediates methylated mRNA export from the nucleus to the cytoplasm by interacting with the splicing factor and nuclear export adaptor protein SRSF3 and facilitating its binding to RNA (Roundtree et al. 2017b). Similarly, YTHDC2 is a putative RNA helicase that associates with meiosis-specific coiled-coil domain-containing protein (MEIOC) and participates in controlling RNA levels during meiosis, and it also facilitates translation (Hsu et al. 2017b). It has also been shown to interact with 5′-3′exoribonuclease XRN1 via its ankyrin repeat domains, indicating the role of YTHDC2 in RNA stability (Kretschmer et al. 2018).

In recent years several other m^6A readers other than the YTH family of proteins have been identified. In fact, m^6A methylation impacts alternative splicing by promoting the binding of YTH and HNRNP (heterogeneous nuclear ribonucleoprotein proteins) to the RNA (Dominissini et al. 2012; Liu et al. 2015, 2017; Noack and Calegari 2018; Zhou et al. 2019). The RNA binding protein HNRNPC is an m^6A reader that binds to m^6A modified RNA by m^6A switch mechanism that in turn destabilizes RNA hairpin and results in the exposure of HNRNPC single-stranded binding motif (Liu et al. 2015). In addition, ribonucleoproteins HNRNPA2B1 and HNRNPG have been reported to mediate alternative splicing in an m^6A dependent mechanism (Liu et al. 2017; Zhou et al. 2019). Indeed, HNRNPA2B1 binds to nuclear transcripts and exerts similar alternative splicing effects as that of METTL3, the m^6A writer, and also mediates pre-miRNA processing by interacting with DGCR8, a miRNA Microprocessor complex member protein (Liu et al. 2017). Similarly, HNRNPG has been reported to bind RNA and phosphorylated carboxy-terminal domain of RNA Polymerase-II through Arg-Arg-Methionine (RRM) and Arg-Gly-Gly (RGG) motifs to facilitate RNA Polymerase-II occupancy, pre-mRNA processing, and has been linked with neural development and neuromuscular diseases (Liu et al. 2017; Zhou et al. 2019). The insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs) have also been identified as m^6A reader proteins that promote the stability and storage of the mRNAs under normal and stress conditions (Huang et al. 2018). Moreover, ribosomes can also act as m^6A readers, and single-molecule ribosome translocation assay has demonstrated that ribosomes can stall mRNA at m^6A site influencing RNA stability and translation (Choi et al. 2016). However, it remains unclear regarding what extent it can influence translation and mRNA stability. Furthermore, few anti-readers of m^6A methylation, such as Ras GTPase-activating protein-binding protein 1 (G3BP1) and G3BP2 have also been identified (Arguello et al. 2017; Edupuganti et al. 2017). These anti-readers proteins are repelled by the presence of m^6A modification and thus preferentially bind to unmethylated sequences. The target sequence of G3BPs within RNAs shows a positive correlation with the binding sites of YTHDF2 and YTHDC1 (Arguello et al. 2017; Edupuganti et al. 2017). Taken together, it can be inferred that the m^6A readers can exert their afore-stated multifaceted, including contrasting influence, based on their interaction/association with various factors mentioned.

2.2 *m^5C and 5-Hydroxymethylcytosine (hm^5C) Modifications*

m^5C in RNA represents another important post-transcriptional modification involving cytosine base at carbon-5 position. Recent technological advancement has facilitated the detection of over 10,000 m^5C modification sites in human transcriptome (Xue et al. 2020). m^5C and hm^5C are found abundantly in rRNA and tRNA and are shown to be essential for their maintenance, stability, function, and biogenesis (Flores et al. 2017; Xue et al. 2020). In addition, m^5C modification is also prevalent in mRNAs and lncRNAs, and that is catalyzed by the NOP2/Sun

(NSUN) domain of RNA methyltransferase family member NSUN2. These are evolutionary conserved and are found predominantly in mRNAs at CG dinucleotides near translation initiation site and at untranslated regions near argonaute binding sites (Xue et al. 2020). Moreover, NSUN2 modulates the function of mRNA export adapter ALYREF, its RNA binding affinity, and nuclear-cytoplasmic shuttling of mRNA, indicating the role of m^5C modification in RNA export and post-transcriptional regulation (Khoddami and Cairns, 2013; Yang et al. 2017; Xue et al. 2020).

m^5C modification can be present at varied levels among different tissues and is implicated in common metabolic processes and cell-type-specific functions (Amort et al. 2017; Yang et al. 2017). During early embryogenesis in zebrafish, the maternal-to-zygotic transition has been reported to be associated with m^5C in mRNA, and thereby it facilitates mRNA stability during the same (Yang et al. 2019). Moreover, m^5C modification has also been reported to be essential for the maintenance of cortical, hippocampal, and striatal neurons in vitro as well as in mouse model in vivo (Tuorto et al. 2012; Blanco et al. 2014), where the loss of m^5C has been reported to be leading to impairment in neural differentiation and neural cell migration (Flores et al. 2017). Even using *Drosophila melanogaster*, the fruit fly, as a model system, it was found that deletion of NSUN2 ortholog might lead to short-term memory loss (Abbasi-Moheb et al. 2012). Similarly, the mutation in NSUN2 in humans has been shown to be associated with microcephaly, intellectual disability, and facial dysmorphism reflecting Dubowitz syndrome, and hence suggesting the evolutionarily conserved role of RNA modification in neural development and cognitive function (Abbasi-Moheb et al. 2012; Khan et al. 2012; Martinez et al. 2012).

miRNAs are the common substrate for methylation at adenosine, guanosine, and cytosine residues. A recent study has revealed that the significant pool of miRNAs do contain abundant m^5C modification (Cheray et al. 2020). They have demonstrated that DNMT3A/AGO4 miRNAs exhibit m^5C methylation, suppress miRNA/mRNA duplex formation that eventually leads to alteration in their gene repression function. Further, they have observed that m^5C methylation of miRNA-181a-5p abolishes its tumour suppressor function and results in poor prognosis in glioblastoma multiforme (GBM) (Cheray et al. 2020). Similarly, using transcriptomic wide mapping studies in Drosophila as a model system, it has been reported that hm^5C marks are present in the transcripts of many genes preferably in the coding sequences (Delatte et al. 2016). Importantly, hm^5C levels have been reported to be higher in the brain compared to other tissues (Delatte et al. 2016; Noack and Calegari 2018). However, studies on hm^5C are lacking in the mammalian system due to the synthesis of hm^5C by the very same TET enzymes targeting both DNA and RNA (Lian et al. 2016). However, the studies on hm^5C are available in flies, since they lack DNA methyltransferases leading to negligible levels of m^5C and hm^5C in DNA as compared to that in RNA. In fact, the levels of hm^5C are associated positively with translational efficiency (Delatte et al. 2016). Collectively, it underscores the importance of m^5C and hm^5C RNA modifications as post-transcriptional gene regulatory machinery and their importance during development and disease.

2.2.1 Regulators of m^5C and hm^5C Modifications

Similar to m^6A, m^5C modification also involves writers, readers, and erasers for its regulation (Table 1). m^5C and hm^5C modifications are catalyzed by proteins encompassing NSUN family members and DNMT2 as writers, RNA binding proteins ALYREF and YBF1 as readers, and demethylases including Ten-eleven translocation (TET) family of enzymes as erasers (Goll et al. 2006; Tuorto et al. 2012; Khoddami and Cairns 2013; Xue et al. 2020).

Writers

The two writers of m^5C modification include DNMT2 and NSUN. Both proteins possess Rossman-fold catalytic domain and an *S*-adenosyl methionine (SAM) binding site, while exhibiting complementary target specificities (Xue et al. 2020). DNMT2 primarily contributes to the m^5C modification of tRNA, which is important for its stability and is implicated in the regulation of translation (Tuorto et al. 2012). DNMT2 that is associated with tRNA, also known as tRNA methyltransferase 1 (TRMDT1) exhibits sequence and structural similarity to DNA methyltransferase (Dong et al. 2001). It mediates the methylation of $tRNA^{AspGUC}$, $tRNA^{GlyGCC}$, and $tRNA^{ValAAC}$ having 3′GCG and 5′-CA sequences within the anticodon loop at methylated position C38 (Tuorto et al. 2012). This DNMT2 is localized in the cytoplasm and its knockdown does not render any alteration in the DNA methylation level (Goll et al. 2006). However, its depletion results in the reduction of m^5C methylation in mRNA, suggesting m^5C to be an additional substrate of DNMT2 apart from tRNA (Xue et al. 2019). Similarly, the NSUN family proteins include 7 members (NSUN1 to NSUN7) involved in catalyzing methylation of cytosine to m^5C. Among these, NSUN2 was the first one to be identified and widely studied protein (Abbasi-Moheb et al. 2012; Xue et al. 2020). It was found to be associated with diverse RNA species, including mRNA, tRNA, rRNA, and mitochondrial RNA (mtRNA) (Xue et al. 2020). NSUN2 was shown to exert various vital functions such as cell cycle progression, mitochondrial oxidative phosphorylation, protein synthesis, and regulation of plant root development upon oxidative stress response (Gkatza et al. 2019). Similarly, NSUN1, NSUN4, and NSUN5 have been implicated in rRNA methylation and subsequent ribosome biogenesis and assembly (Xue et al. 2020). Even in yeast, NSUN1 homologue Nop2 was shown to mediate 25S rRNA methylation at position C2870 in domain V and alter the biogenesis of the 60S ribosome subunit (Sharma et al. 2013). Camara et al. (2011) had studied the function of NSUN4 and had shown its interaction with MTERF4 (mitochondrial regulatory factor) and its recruitment to the larger subunit of the mitochondrial ribosome. That, in turn, promoted the ribosome assembly by methylating 12S rRNA at position C911. Similarly, NSUN5 was demonstrated to methylate 25S rRNA at position C2278 within its conserved domain IV (Gigova et al. 2014). NSUN3 and NSUN6 also modify tRNAs predominately similar to that of DNMT2. NSUN3 is present in

mitochondria and plays a role in the methylation of mitochondrial tRNA methionine (mt-tRNA^{Met}) at C34 in the wobble position within the anticodon loop. Further, oxidation of methylated cytosine to 5-formylcytosine (f^5C) has also been implicated in driving the normal mitochondrial translation of the respiratory chain complex and oxidative phosphorylation (Van Haute et al. 2016). Interestingly, NSUN3 dysfunction promotes ESCs differentiation towards the meso- and endoderm lineages rather than neuroectoderm (Trixl et al. 2018). The other member NSUN6 is partially localized in the Golgi apparatus and pericentriolar matrix in the cytoplasm. It is implicated in the methylation of tRNA^{Cys} and tRNA^{Thr} at position C72 and it alters their biogenesis as well (Haag et al. 2015).

Erasers

TET enzymes include TET1, TET2, and TET3, which are Fe (II) and α-ketoglutarate (α-KG) dependent dioxygenases, and these have distinct cellular distributions. While TET1 and TET2 are found predominantly in the nucleus, TET3 is present in both the nucleus and cytoplasm (Traube and Carell 2017). Interestingly, the TET enzymes oxidize m^5C modification to hm^5C in both DNA and RNA (Fu et al. 2014a). Moreover, Huber et al. (2015) have shown the oxidation of m^5C to hm^5C and f^5C, and the latter has been proven to further oxidize to five carboxycytidine by TET enzymes (Basanta-Sanchez et al. 2017).

Readers

The key component of mRNA export complex includes Aly/REF export factor (ALYREF), the THO subcomplex and RNA helicase UAP56. ALYREF binds to 5′ end of mature RNA in a CBP80-dependent manner, whereas it binds to the 3′end region in a PABPN1-dependent manner and interacts with CstF64 (3′ processing factor), thereby enhancing its overall binding to mRNA (Xue et al. 2020). Y-box binding protein 1 (YBX1) is another reader protein implicated in controlling mRNA stability in the cytoplasm. YBX1 influences mRNA metabolism by stabilizing m^5C methylated mRNAs by interacting with ELAV-like protein 1 (ELAVL1) (Chen et al. 2019b; Yang et al. 2019).

2.3 N1-methyladenosine (m^1A)

m^1A methylation representing methylation of adenine residue at carbon 1 position is found abundantly in various transcripts as a single m^1A site of over 4000 genes in eukaryotic cells (Dominissini et al. 2016; Iyer et al. 2016). These m^1A methylation sites are preferentially enriched around the start codon that is upstream of the first splice site exhibiting high GC content and are highly conserved. These are capable of

influencing RNA structure and RNA–protein interactions, which can play role in translation initiation (Dominissini et al. 2016; Li et al. 2016a; Iyer et al. 2016; Zhao et al. 2017). m^1A was first noted in tRNAs at position 58 where it was implicated in stabilizing the tertiary structure (Schevitz et al. 1979). It is also found in 28S rRNA in humans and is known to play an important role in rRNA biogenesis (Peifer et al. 2013). A number of studies have demonstrated the occurrence of m^1A modification in mammalian mRNA and with m^1A/A ratios varying among different cell lines and tissues (Dominissini et al. 2016; Li et al. 2016a). The enzymes such as ALKB, ALKBH1, TRM6, TRM10, and TRM61 that have been reported to catalyze m^1A modification carry out SAM-dependent methyltransferase reaction in rRNA and tRNA. While TRMT6 and TRMT61A promote the methylation of cytoplasmic tRNA, TRMT61B modulates mitochondrial tRNA (Guy and Phizicky 2014). Similarly, 28S rRNA methylation is carried out by RRP8 (NML) (Peifer et al. 2013). m^1A modification can also be removed by demethylases belonging to ALKB family enzymes ALKBH1 and ALKBH3. However, specific readers and writers for m^1A modifications are largely unknown to elucidate the functional significance of this modification.

2.4 2′-O-Methylation (2′-OMe or Nm)

2′-OMe is co- or post-transcriptional RNA modification exhibited by the addition of methyl group to the 2′hydroxyl of the ribose moiety in any nucleotide. 2′-OMe is found essentially in all RNAs that include rRNAs, small nuclear RNAs, tRNAs, and mRNAs and plays a major role in the stabilization of higher order RNA secondary structures, protection of RNAs from nuclease attacks, and also increases the hydrophobicity of RNA (Angelova et al. 2018). Additionally, this modification also occurs in miRNAs, siRNAs, and piRNAs at their 3′end in both plants and animals (Angelova et al. 2018). FTSJ1, also known as MRX9, TRM7, and JM23 homologue in different species, is a tRNA 2′-O-methyltransferase that modifies the anticodon loop of $tRNA^{Phe}$ and $tRNA^{Trp}$ at position C32 and N34 (Somme et al. 2014; Dimitrova et al. 2019). FTSJ1 gene is present on the small arm of chromosome X, and its loss of function is associated with non-syndromic X-linked intellectual disability (NSXLID), suggesting a functional link between $tRNA^{Phe}$ and central nervous system development (Dimitrova et al. 2019). The 2′-OMe methylation in mRNA is predominantly seen in the cap region during transcription. This commences upon the addition of 7-methylguanosine to the 5′ end of the first transcribed nucleotide (N1), known as cap0 (m^7GpppN), followed by 2′-*O*-methylation on its ribose moiety, hence forming cap1 (m^7GpppNm). Additionally, 2′-OMe can also be present on the second transcribed nucleotide (N2) as cap2 (m^7GpppNmNm) (Dimitrova et al. 2019). In fact, cap2 is found in higher eukaryotes, whereas cap1 is seen in lower eukaryotes, and the 2′-OMe methylation at both cap1 and cap2 is carried out by CMTR1 and CMTR2, respectively (Werner et al. 2011; Byszewska et al. 2014). 2′-OMe modification in mRNA has also been reported in the 5′ and 3′

UTR, in CDS near the splice sites, and in introns. While the presence of 2′-OMe on the cap structure prevents the degradation of m^7GpppRNAs by DXO protein implicated in decapping and exoribonuclease activity, the same within the coding region impairs the codon reading during translation (Picard-Jean et al. 2018; Dimitrova et al. 2019). Moreover, 2′-OMe on rRNA has been reported to be modulated at key functional sites of translation in human ribosomes, implicating its role in controlling the translation of mRNA as well (Erales et al. 2017). Incidentally, CMTR1 expression has been found to be upregulated in the brain of murine model for Alzheimer's disease (AD) that is mediated by expression of genes associated with innate immune response (Dimitrova et al. 2019). Hence, it may be interesting to assess the link between 2′-OMe, protein translation, and Tau accumulation.

2.5 Pseudouridine (Ψ)

Pseudouridine (5-ribosyluracil or Ψ) is the first discovered most ubiquitous RNA modification found in both coding and non-coding RNAs. It is implicated in the maintenance of RNA secondary structure and its stability (Angelova et al. 2018). Pseudouridylation was also shown to influence various cellular processes, including splicing, translation efficiency, gene expression regulation, and telomere maintenance. The Pseudouridine synthase (Pus) family of enzymes catalyzes pseudouridylation of their substrates in both RNA-dependent and RNA-independent manner (Angelova et al. 2018). The former causes H/ACA box snoRNAs mediated pseudouridylation of target RNAs by sequence-specific interaction between snoRNAs and target RNA, where uridine modification is catalyzed by specific enzymes present in sno-ribonucleoprotein (snoRNP) such as dyskerin in human and cbf5 in yeast (Duan et al. 2009). Similarly, RNA-independent pseudouridylation is mediated by Pus that catalyzes Ψ formation at particular target RNA (Carlile et al. 2014; Angelova et al. 2018). Pus enzymes are evolutionary conserved and divided into six families according to their consensus sequences such as TruA, TruB, TruD, RluA, RsuA, and Pus10, while the latter predominately belongs to eukaryotes and archaea (Angelova et al. 2018). Pus1 is a member of the TruA family and is implicated in the Ψ of tRNA, rRNA, snRNA, and mRNA (Carlile et al. 2014; Schwartz et al. 2014). Pus3 also belongs to TruA family and exhibits sequence similarity with Pus1.

Various studies have revealed the importance of Ψ influencing neuronal development and functions both positively and negatively. Mutation in Pus1 has been found to be associated with mitochondrial myopathy, sideroblastic anaemia, mild cognitive impairment, and neuronal survival, suggesting its function in brain activity (Angelova et al. 2018). Similarly, the presence of Pus3 in the nervous system of mice embryos, detected through in situ hybridization studies, has suggested its plausible role in neural development. Moreover, Dyskerin has also been shown to be present abundantly in embryonic neural tissue and specific subsets of neurons in the cerebellum and the olfactory bulb of adult brains, indicating its crucial role in neural

development (Heiss et al. 2000). RluA-1 is another enzyme found in the dendrites of a subset of peripheral neurons in embryos, and it modulates uridines in rRNA and tRNA in *Drosophila melanogaster* and functions in the peripheral nervous system development during embryogenesis (Wang et al. 2011). In contrast, pseudouridylation has also been seen as associated with neural disorders. In patients with mild-to-moderate severity of AD, increased level of Ψ has been observed (Lee et al. 2007). Similarly, the patients with myotonic dystrophy type 2 (DM2), a neuronal disorder, show increase binding of Muscle blind-like 1 protein (MBNL1) to CCUG repeats in the intronic region of CNBP genes (Delorimier et al. 2017). While Ψ modification within CCUG repeats moderately affects the binding of MBNL1, the same within YGCY motif-containing structured RNA results in the drastic reduction in MBNL1 binding to CCUG repeats. Furthermore, pseudouridylation can be used as a direct indicator of oxidative stress, which is a causative agent for neurodegeneration (Uttara et al. 2009). Li et al. (2015) have also reported 40–50% upregulation in the mRNA of Ψ upon H_2O_2 treatment indicating the role of pseudouridylation in cellular stress.

2.6 7-Methylguanosine (m^7G)

The N7-position of guanosine 46 is methylated in various prokaryotic and eukaryotic tRNAs to exhibit m^7G methylation (Boccaletto et al. 2018). This reaction is catalyzed by heterodimeric complex exhibiting Trm8 as the catalytic subunit and Trm82 as associated subunit vital for enzymatic activity (Alexandrov et al. 2005). Mammalian homologues of both the subunits are METTL1 and WDR4 genes, respectively (Ramos and Fu 2019). Mutation in WDR4 is associated with the reduction of m^7G in $tRNA^{Phe}$ and results in the formation of microcephalic primordial dwarfism (Ramos and Fu 2019). Indeed, WDR4 mutation is implicated in growth delay and intellectual disability (Chen et al. 2018). This modification is important for the proper folding of tRNA and formation of the tertiary structure by mediating a hydrogen interaction between guanosine at position 46 in the T loop and position 22 in the D-stem loop. This interaction is facilitated by positive charge conferred by m^7G methylation (Lorenz et al. 2017). m^7G modification in combination with other tRNA modifications does mediate tRNA instability by inducing rapid tRNA decay (RTD) pathway. In particular, Trm8/Trm4-deficiency results in the de-aminoacylation and instability of a single tRNA, tRNA-Val-AAC, in yeast cells making them temperature sensitive (Alexandrov et al. 2005; Chernyakov et al. 2008). In addition, the RTD pathway interacts with the translation elongation factor 1A suggesting the role of m^7G modification in translation control (Dewe et al. 2012). Trm8 mediated m^7G tRNA modification is essential for proper codon occupancy by ribosomes as well (Chou et al. 2017). Moreover, Trm8- or Trm82-deficient yeast cells show enhanced P-site occupancy over valine codons GUC/GUG/GUU and changes in UGC and UGU codon occupancy, which are not decoded by Trm8/82 modified tRNAs. In fact, Trm82 deficient cells show global change in ribosome

occupancy irrespective of the change in mRNA expression (Chou et al. 2017; Ramos and Fu 2019). However, m^7G modification does not participate in GAAC activation, suggesting that it influences different pathways as compared to other tRNA modifications that are associated with the activation of GAAC (Chou et al. 2017; Ramos and Fu 2019). Indeed, WDR4 is involved in the regulation of METTL1 by controlling the phosphorylation of its catalytic subunit by insulin-mediated kinase, PKB, which might have implications in controlling the synthesis of certain proteins mediating insulin signalling (Cartlidge et al. 2005). Furthermore, METTL1 and WDR4 involved in m^7G tRNA modification exert a crucial influence on cell cycle control and brain development, and their deficiency leads to impairment in self-renewal and neural differentiation in mouse ESCs (Lin et al. 2018).

3 Methods for the Detection of RNA Modifications

Although RNA modifications have been detected since long, the lack of reliable and efficient detection methods and site-specific mapping have led the field growing at a relatively slower pace. Of late though, there has been growing interest seen considering the implications that these post-transcriptional modifications are associated with pertaining to development and disease. Epitranscriptomic sequencing methods mainly focus on the enrichment and purification of modified RNA and the improvement of bioinformatics analysis. Apparently, advancement in the next generation sequencing methods has led to the identification of RNA modification at single-nucleotide resolution. Moreover, AI and Machine Learning approaches have also been explored of late in obtaining error-free sequence identification. The various approaches that have been undertaken for the detection of different RNA modifications are described below and can be seen in Table 2.

3.1 m^6A-seq and MeRIP

In 2012, two breakthrough methods called m^6A-seq and MeRIP (m^6A-specific methylated immunoprecipitation or methylated RNA immunoprecipitation) were discovered to identify any type of RNA modifications, including m^6A, m^6Am, m^5C, hm^5C,and 2′-OMe (Dominissini et al. 2012; Meyer et al. 2012). Methylation at position 6 of adenine residue has no effect on its ability to base pair with thymidine or uracil, thus impeding its detection using standard sequencing or hybridization methods (Meyer et al. 2012). This modification is present abundantly in $polyA^+$ mRNA and other RNAs, including rRNA, tRNA, lncRNA, and snRNA. Initially, these methods were optimized to identify $polyA^+$ mRNA methylation. Subsequently however, the protocol could be adapted to profile all types of RNA. Incidentally, these methods have detected ~10,000 m^6A peaks in the mammalian transcriptome (Meyer et al. 2012; Dominissini et al. 2012). The strategy involves RNA

Table 2 Methods of detecting RNA modifications

RNA modification	Detection method	Principle	Nucleotide resolution	References
m^6A, m^6Am, m^5C, hm^5C, and 2′OMe	m6A-seq/MeRIP-seq	RNA–protein immunoprecipitation	100–150 nt	Dominissini et al. (2012), Meyer et al. (2012)
m^6A, m^6Am, m^5C, hm^5C	m^6A CLIP/miCLIP	RNA–protein immunoprecipitation	~20–80 nt	Ke et al. (2015), Linder et al. (2015)
m^6A	Photo-cross-linking assisted m^6A seq (PA-m6A seq)	RNA–protein immunoprecipitation	Enzyme specific nucleotide resolution	Schwartz et al. (2013), Li et al. (2016b)
m^6A, m^5C, 2′-OMe, Ψ	SCARLET	Site-specific cleavage and splint ligation	Single-nucleotide resolution	Liu and Pan (2015b)
m^6A RNA isoforms	m^6A-LAIC-seq	RNA immunoprecipitation	Enzyme specific nucleotide resolution	Li et al. (2016b), Molinie et al. (2016)
m^6A	Single-base mapping of m^6A by an antibody-independent method	m6A sensitive RNA endoribonucleases (MazF and ChpBk)	Single-nucleotide resolution	Zhang et al. (2019)
m^6A	Metabolic labelling method	Substitution of the methyl group of SAM with the allyl	Single-base resolution	Shu et al. (2020)
m^5C, hm^5C	Bisulphite sequencing	Sodium bisulphite mediated deamination of methylated and unmethylated cytosine	Single-nucleotide resolution	Schaefer et al. (2009), Chen et al. (2019c)
m^5C	Aza-IP-seq	RNA–protein immunoprecipitation	Single-nucleotide resolution	Khoddami and Cairns (2013), Li et al. (2016b)

fragmentation to smaller oligonucleotides followed by immunoprecipitation with anti-m^6A antibody, isolation of RNA, cDNA library preparation, adapter ligation, and finally sequencing. Indeed, MeRIP-seq method can increase the yield of m^6A RNA sequences by >130 fold (Meyer et al. 2012). However, the stated methods also do suffer from several drawbacks, including (1) requirement of large input material, (2) low resolution, and (3) difficulty in assessing false positives (Schwartz et al. 2013). Moreover, the current available bioinformatics methods are able to detect only single site per 100–200 nt wide peak resulting in the missing of the substantial amount of m^6A clusters, which may otherwise contain upto 15 such m^6A sites (Ke et al. 2015; Linder et al. 2015).

3.2 m^6A CLIP or miCLIP

m^6A CLIP (Cross-linking immunoprecipitation) or miCLIP (m^6A individual-nucleotide-resolution cross-linking and immunoprecipitation) or methylation induced cross-linking and immunoprecipitation detection method possesses the advantage of identification of methylation sites in RNA using sequencing method at single-nucleotide resolution (Ke et al. 2015; Linder et al. 2015). In this method, RNA is fragmented to ~20–80 nt followed by the formation of UV-induced covalent RNA-antibody complex. Subsequent recovery of these complexes is carried out by immunoprecipitation followed by isolation of RNA, reverse transcription, library construction, and sequencing, and these steps basically rely on the UV-induced mutations. After antibody removal, remnants of peptides on the RNA at cross-linking site cause truncations, insertions, and C to T mutations during the formation of cDNA from RNA at the position +1 to the m^6A site (5′ to the m^6A site) in sequencing reads. miCLIP has been used to detect other RNA modifications as well that includes m^6Am and m^5C (Ke et al. 2015; Linder et al. 2015). C271A induced mutation of NSUN2 inhibits its release from RNA and thus is used to detect m^5C modification in non-coding RNA such as tRNA using miCLIP method (Li et al. 2016b). Indeed, miCLIP exhibits high-resolution and low false discovery rate (Ke et al. 2015; Linder et al. 2015).

3.3 Photo-Cross-Linking Assisted m^6A seq (PA-m^6A seq)

This method utilizes 4-thiouridine (4SU) incorporation into the RNA by addition through growth media that facilitates its incorporation near m^6A sites followed by immunoprecipitation, UV cross-linking, competition elution of the RNA-antibody complex, isolation of RNA, library construction, and sequencing (Schwartz et al. 2013; Li et al. 2016a,b). Remnants of peptide fragments after antibody removal induce point mutation involving C to T at the 4SU cross-linking site during reverse transcription reaction. However, the challenge remains on precisely identifying m^6A sites using point mutation based on 4SU incorporation, as the latter position can vary with respect to any m^6A residue. Hence, the limitation of this technique could be the non-detection of the m^6A sites, which does not incorporate 4SU (Li et al. 2016b).

3.4 SCARLET

SCARLET (site-specific cleavage and radioactive-labelling followed by ligation-assisted extraction and thin-layer chromatography) is used to detect internal RNA modification in mRNA/lncRNA, which includes m^6A, m^5C, 2′-OMe, and ψ as well as other possible modifications. This method is based on the combination of two

previous techniques, such as site-specific cleavage and splint ligation for the identification of RNA modification in transcriptome and can analyze any modification which follow Watson–Crick base pairing (Liu and Pan 2015b). In this method, chimeric DNA oligonucleotide binds to the target RNA around the modification site, followed by RNase cleavage, radiolabelling with phosphorous-32, splint ligation to an 116 nt ssDNA oligonucleotide, isolation of the radio-labelled product, and identification of modified and unmodified adenosines by thin-layer chromatography. The advantage of this method is that the RNA modification in low abundant RNAs such as mRNA, tRNA, and lncRNA can also be detected at single-nucleotide resolution using this strategy. However, this method is not suitable for the quantification of large-scale m^6A sites (Linder et al. 2015).

3.5 *m^6A-LAIC-seq*

m^6A-LAIC-seq (m^6A-level and isoform-characterization sequencing) is a high-throughput sequencing approach used to characterize the methylation level of whole-transcriptome. This involves full-length RNAs being subjected to immunoprecipitation followed by elution of methylated RNA-antibody complex, RNA isolation, reverse transcription, library preparation, and sequencing. In this method, external RNA controls consortium (ERCC) spike ins are added to the eluent and supernatant, and m^6A levels per site or gene are determined by ERCC-normalized RNA abundance in different pools (Molinie et al. 2016; Li et al. 2016b). This method can be utilized to compare the alternative spliced isoforms of RNA modifications.

3.6 *Single-Base Mapping of m^6A by an Antibody-Independent Method*

This is a high-throughput antibody-independent m^6A identification method based on the m^6A sensitive RNA endoribonuclease, which distinguishes between methylated and unmethylated adenine residues through the ACA motif (Zhang et al. 2019). This group has identified two endoribonucleases; MazF and ChpBk sensitive to m^6A methylation within the ACA motif. Both the enzymes are reported to cleave unmethylated motif leaving m^6A methylated motifs intact, which are further identified by m^6A -REF-seq at single-nucleotide resolution.

3.7 Metabolic Labelling Method

Recently, Shu et al. (2020) have reported metabolic labelling method, also called m^6A label sequencing to identify m^6A modification at single-base resolution. In this method, mammalian cells are fed with a methionine analogue, Se-allyl-L-selenohomocysteine, which substitutes the methyl group of SAM with the allyl resulting in the modification of m^6A to N6-allyladenosine (a^6A). Further, reverse transcription leads to the iodination-induced misincorporation at the opposite site in complementary DNA, which is further detected by sequencing methods at single-nucleotide resolution (Shu et al. 2020).

3.8 Bisulphite Sequencing

Bisulphite sequencing is based on the principle that sodium bisulphite deaminates methylated and unmethylated cytosine at acidic pH to yield uracil sulfonate/5-methyluracil sulfonate, which can be further converted to uracil/thymine at basic pH (Xue et al. 2020). In 2009, this technique was modified based on the PCR method, which facilitated the amplification of cDNA from low levels of RNAs (Schaefer et al. 2009). In brief, sodium bisulphite can deaminate the unmethylated cytosines to uracils that can be subsequently replaced by thymines (Ts) during PCR amplification. However, the methylated cytosines remain unchanged during the process, thus allowing a differentiation between methylated and unmethylated cytosines in RNA (Schaefer et al. 2009). Coupled with NGS, this method is now regarded as the gold standard for mapping cytosine methylation at single-nucleotide resolution with high specificity (Chen et al. 2019c). However, the limitation that this method carries pertains to non-discrimination of m^5C from hm^5C modifications using the same (Huber et al. 2015).

3.9 Aza-IP-seq

5-Azacytidine-mediated RNA immunoprecipitation is designed for the detection of the targets of methyltransferases such as NSUN and DNMT2 responsible for m^5C marks. In this method, 5-aza-C is introduced into the cell, which gets incorporated into the nascent RNA in place of cytosine (Khoddami and Cairns 2013). This transition results in the trapping of m^5C methyltransferases to RNA even after methylation, which can then be immunoprecipitated with enzyme or tag target antibody. These samples are further subjected to sequencing, which allows the detection of m^5C sites at single-nucleotide resolution based on C to G transversion. However, this method suffers from certain limitations such as 5-aza-C toxicity,

random incorporation of 5-aza-C, and the possibility of RNA not replaced by 5-aza-C (Khoddami and Cairns 2013; Li et al. 2016a, b).

4 Functional Attributes

RNA modifications are essential post-transcriptional events. As per the classical gene expression process, DNA is transcribed to hnRNA, which further through splicing or alternative splicing events culminates in processed mRNA, the translation ready code. Hence the question arises pertaining to why at all RNA modification is required that adds another layer of complexity to the machinery? As discussed in the previous sections, RNA modifications are associated with multifaceted attributes. m^6A modification has been implicated in the regulation of wide gamut of activities associated with mRNA metabolism such as nuclear export, pre-micro RNA processing, microRNA (miRNA) mediated decay, polyadenylation, R-loop accumulation, RNA stability, and translation (Meyer et al. 2012; Zheng et al. 2013; Ke et al. 2015; Roundtree et al. 2017a, b; Angelova et al. 2018; Huang et al. 2018; Kretschmer et al. 2018; Abakir et al. 2020). In fact, m^6A modification commences at the pre-mRNA level during transcription and mRNA processing and culminates upon the release of mRNA from chromatin into the nucleoplasmic RNA pool (Huang et al. 2019). The presence of m^6A methylation in the 5′UTR region of mRNA promotes translation in a cap-independent manner by directly binding to the eukaryotic initiation factor 3 (elF3) (Meyer et al. 2015). Moreover, this modification controls alternative splicing for a certain pool of mRNAs and lncRNAs, rather than a complete population of RNAs (Roundtree et al. 2017b; Zhu et al. 2018). The functional relevance of m^6A RNA modification can be seen in Fig. 1.

m^5C and hm^5C are implicated in mediating rRNA and tRNA stability, their function and biogenesis, as well as mRNA export and post-transcriptional regulation (Khoddami and Cairns 2013; Flores et al. 2017; Yang et al. 2017; Xue et al. 2020) (Fig. 2). Similarly, 2′-OMe is also associated with the stability of tRNA structure, translation regulation, and splicing (Dimitrova et al. 2019), whereas Mettl1-mediated m^7G modifications are crucial for tRNA folding and formation of tertiary structure and mRNA translation (Lorenz et al. 2017; Lin et al. 2018). While m^1A methylation is important for the maintenance of RNA structure and RNA–protein interactions, which can influence translation initiation as well as rRNA biogenesis (Dominissini et al. 2016; Li et al. 2016a; Iyer et al. 2016; Peifer et al. 2013; Zhao et al. 2017), Ψ plays a vital role in the maintenance of RNA secondary structure and its stability, and is known to regulate several cellular processes including splicing, translation efficiency, gene expression regulation, and telomere maintenance (Angelova et al. 2018). In addition to the afore-stated attributes, RNA modifications also influence various developmental events (Haran and Lenka 2019). In the following sections, we shall be discussing the specific contribution of RNA modification during neurogenic progression, and also some of the associated neurological disorders in response to altered epitranscriptomic signatures.

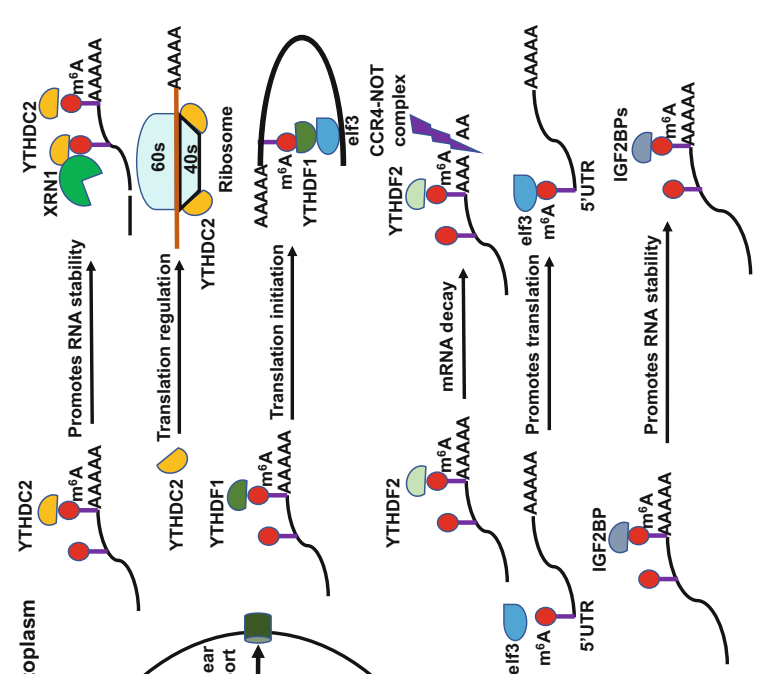


Fig. 1 Functional relevance of m^{6}A RNA modification and associated players. In nuclear compartment, m^{6}A modification is mediated by a large complex of RNA methyltransferase comprising of METTL13, METTL14, and WTAP along with co-factors. The same is removed by demethylases such as ALKBH5 and FTO, acting as erasers. The nuclear export of m^{6}A modified RNA commences upon binding of reader protein YTHDC1 to m^{6}A along with its interacting

partners; nuclear adaptor protein SRSF3 and splicing factor. Subsequently, in the cytoplasmic compartment, the m^6A reader proteins bind to RNA transcripts and mediate their translation, stability, and degradation. While the interaction of reader protein YTHDF1 with elf3 promotes translation, YTHDF2, through its interaction with CCR4-NOT deadenylase complex, contributes to mRNA decay. The interaction of elf3 with m^6A also promotes translation. YTHDC2 imparts RNA stability by interacting with exoribonuclease XRN1 and also participates in translation regulation by binding to the ribosome. Even IGFT2BPs also serve as m^6A readers and promote RNA stability

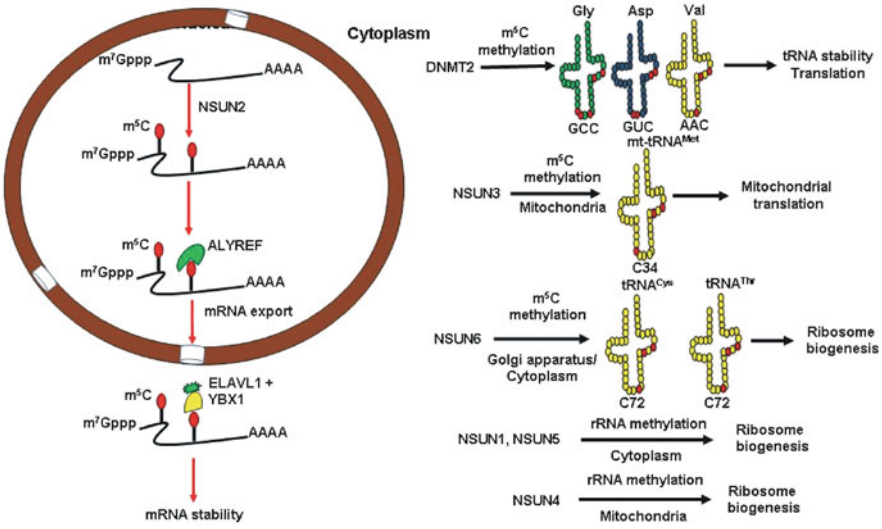


Fig. 2 Functional relevance of m^5C RNA modification and players involved. Inside nuclear compartment m^5C modification is catalyzed by NSUN2 RNA methyltransferase, and the m^5C reader protein ALYREF promotes the nuclear export of m^5C transcript. m^5C RNA methyltransferases mediate the m^5C modification of tRNAs in the cytoplasmic compartment and facilitate ribosome biogenesis, translation, and tRNA stability. While DNMT2 mediates the m^5C methylation on $tRNA^{Gly}$, $tRNA^{Asp}$ and $tRNA^{Val}$, NSUN3 causes m^5C methylation on mitochondrial $tRNA^{Met}$. Similarly, NSUN6 localized in Golgi apparatus or cytoplasm catalyzes the m^5C methylation on $tRNA^{Cys}$ and $tRNA^{Thr}$. NSUN1 and NSUN5 confer rRNA methylation in the cytoplasm, and NSUN4 does that in mitochondria. Reader protein YBX1 is implicated in mediating mRNA stability in the cytoplasm by binding to m^5C sites

4.1 Implication of RNA Modifications in Neurogenesis

Neurogenesis is a complex developmental process orchestrated through fine-tuned and temporo-spatial regulatory cascades involving various signalling cross-talks, genetic and epigenetic modulators, etc. The recent findings indicate the involvement of epitranscriptomic machinery underlying the same. Indeed, pluripotent stem cells have emerged as an elegant model system to study neural development and human neurogenesis, which would have otherwise been difficult to monitor due to technical obstacles in obtaining embryonic and foetal tissues. OCT4, Nanog, and SOX2, the key pioneering and pluripotency-associated factors, play a deterministic role in deciding the cell fate concerning whether these cells should remain unspecialized or would they undergo differentiation by altering the expression level of the stated factors during pluripotent stem cells differentiation in culture in vitro (Heng and Ng 2010). While the majority of mRNAs of core pluripotency factors such as *Nanog*, *Klf4*, *c-Myc*, *Lin28*, *Med1*, *Jarid2,* and *Eed* in mouse ESCs exhibit m^6A modification, Oct4, the crucial pioneering factor, interestingly lacks the same (Batista et al. 2014). These m^6A modified transcripts show reduced translational efficiency as

compared to unmodified transcripts. Hence it undoubtedly reflects that m^6A RNA modification machinery might be residing at the crossroad dictating the cell fate. Accordingly, it is plausible that the translation of pluripotency-associated genes that are required to be attenuated during gastrulation and further differentiation into various lineages might be regulated by these modifications. Indeed, the impairment in neural differentiation upon *Mettl3* depletion proves this (Batista et al. 2014). Using *Mettl3* knockout mouse ESCs it has been demonstrated that, when these cells are subjected to neural differentiation, less than 6% of the cells stain positive for TUJ1 (beta-3 tubulin-neuronal marker) as compared to 53% seen in WT cells, and this is due to the inability of the *Mettl3* KO cells to repress Nanog and activate *TuJ1* transcription during neural differentiation. Similarly, in human ESCs, *Mettl3* depletion also leads to impairment in neuronal differentiation, indicating the evolutionary conserved phenomenon of m^6A methylation in ESCs (Batista et al. 2014). These data unequivocally reflect a profound influence of METTL3 dependent m^6A modification during neurogenesis. Another study has also revealed that m^6A modification present on the transcript of pluripotency markers drives their stability and expression during differentiation (Geula et al. 2015).

Many of the existing studies have in fact shed light on the mechanism underlying the RNA modification and have delineated an interesting link between epigenetics and epitranscriptomics and both modulating together the key pluripotency-associated markers in ESCs and neural-specific ones in NSCs (Aguilo et al. 2015; Chen et al. 2019a; Wang et al. 2018). It has been shown that the chromatin-associated zinc finger protein 217 (ZFP217) interacts with several epigenetic markers, and hence it regulates the expression of key pluripotent genes by modulating the levels of m^6A methylation on their transcripts (Aguilo et al. 2015). They have also noted that ZFP217 depletion results in the upregulation of *Mettl3* that subsequently enhances the m^6A methylation level of *Sox2*, *Klf4*, *Nanog,* and *c-Myc* mRNAs and thereby promotes their degradation while activating ESCs differentiation (Aguilo et al. 2015). In fact, the YTHDF2 m^6A reader protein implicated in the mRNA degradation process is highly expressed in induced pluripotent stem cells (iPSCs), and it undergoes degradation upon neural differentiation, thereby inducing the expression of neural-specific transcripts (Heck et al. 2020). Figure 3 depicts the role of m^6A modification during pluripotent stem cells differentiation into the neural lineage.

Recent advancements concerning single-base resolution m^6A-CLIP have demonstrated that the depletion of *Mettl14* or *Ythdf1* in mouse dorsal root ganglion (DRG) leads to reduced translation and functional axon regeneration in the peripheral nervous system (Weng et al. 2018). Reduction in the level of m^6A methylation by conditional knockout of *Mettl3*/*Mettl14* in the developing mouse nervous system has also been shown to be associated with aberrant cell cycles with longer S and G2 phases in radial glia cells and cortical neural progenitor cells (NPCs). This in fact culminates in the slow development of late-born neurons and thereby extends the cortical neurogenesis into postnatal stages. However, the sciatic nerve lesion in the mouse dorsal root ganglion (DRG) region results in the upregulation of transcripts harbouring m^6A methylation sites implicated in regeneration-associated genes and

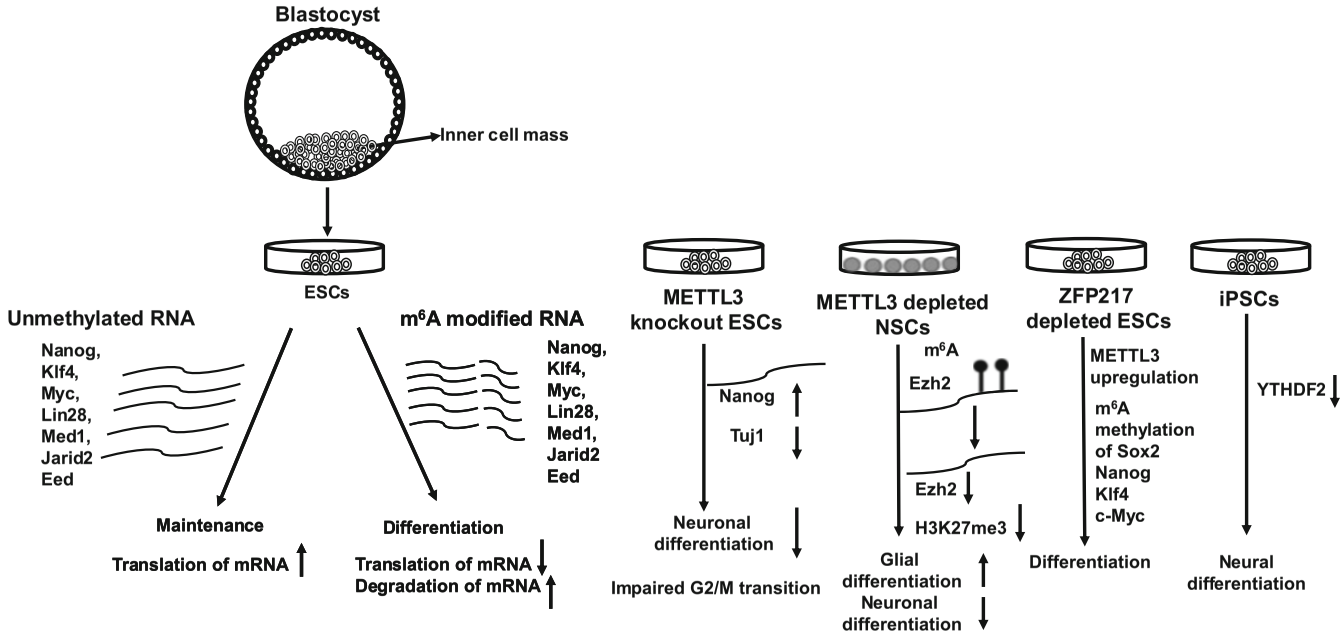


Fig. 3 Regulatory role of m^{6}A modification during differentiation of pluripotent stem cells into neural lineage. During exit from maintenance and entry into differentiation, the transcripts of core pluripotency markers in ESCs such as; *Nanog*, *Klf4*, *c-Myc*, *Lin28*, *Med1*, *Jarid2, Eed* etc., undergo m^{6}A modification and thereby suppress their translation. However, upon *Mettl3* depletion in ESCs *Nanog* expression is reinforced accompanied by differentiation impairment. *Mettl3* depletion in NSCs also leads to gliogenic differentiation at the expense of neuronal ones along with G2/M transition impairment. While knockout of ZFP217 in ESCs results in the upregulation of *Mettl3* that in turn promotes m^{6}A methylation of pluripotency markers and attenuates differentiation; downregulation of reader protein *Ythdf2* promotes neural differentiation

translational machinery (Weng et al. 2018). Prior to that with the help of m^{6}A-seq of mouse brain, it has been demonstrated that many transcription factors such as Sox1, Pax6, Sox2, Neurog/Neurogenin 2, and Emx2 as well as several genes related to stem cell, cell cycle, and cortical neurogenesis are m^{6}A tagged during cortical neurogenesis (Yoon et al. 2017). Another study has shown the depletion of *Mettl14* causing decreased proliferation in embryonic NSCs and its role in self-renewal and decreased generation of cortical neurons (Wang et al. 2018). Moreover, they have also stated the importance of m^{6}A modification in the regulation of histone modifications suggesting a mechanism for gene regulation. Recently Chen et al. (2019a) have reported the role of *Mettl3* in adult neural stem cells (aNSCs) where its depletion has been shown to shift the differentiation balance towards glial cells at the expense of neuronal ones, impairing thereby the neural development process. In addition, they have observed m^{6}A methylation playing a role in the regulation of the transcript level of histone methyltransferase *Ezh2* and subsequently *H3K27me3*, suggesting a cross-talk between RNA modification and histone modification and hence an epigenetic and epitranscriptomic link involvement in the regulation of neurogenesis and neurodevelopment. Similarly, depletion of FTO has been shown to affect reduced proliferation and neural differentiation of adult NSCs suggesting its role in neurogenesis (Li et al. 2017). Together these studies underscore the critical influence of post-transcriptional RNA modifications and the players underlying the same during embryonic and adult neurogenesis. Hence, alteration in their expression or modulation, as would be obvious might lead to neurodevelopmental disorders in developing embryos as well as other neurological complications in adults.

4.2 m^{6}A Modification in Synaptic Function and Behaviour

The brain performs various functions such as learning, emotion, cognition, memory, and motor control. These neuronal activities induce complex patterns of gene expression changes, which modulate neural circuits by altering synaptic development and connectivity (Cholewa-Waclaw et al. 2016). In fact, a number of recent studies have highlighted m^{6}A RNA modification playing a crucial role in regulating gene expression that is associated with the change in cognitive function, such as learning and memory, and hence affecting brain functions. *Fto* deficiency has been shown to increase m^{6}A in a subset of mRNA in *Fto* knockout mice, and that eventually has yielded in reduced neuronal firing rate accompanied by increased dopamine content in the synapse (Hess et al. 2013). This was further shown to result in activating presynaptic dopamine receptors, D2R and D3R, and impairment in dopaminergic signal transduction. Subsequent report has also demonstrated an association between decreased FTO expression in the mouse dorsal hippocampal CA1 neurons and contextual fear, signifying the role of FTO in memory formation (Walters et al. 2017). A couple of other groups have also reported that the mutation in FTO results in brain atrophy and psychological disorder in adulthood (Ho et al. 2010; Hess et al. 2013). Furthermore, knockout of m^{6}A demethylase FTO in mice

has been reported to result in postnatal growth retardation and reduction in adult neurogenesis leading to impaired learning and memory (Boissel et al. 2009; Li et al. 2017). In line with this, another group has also observed the association between m^6A methylation and memory formation in the medial prefrontal cortex (mPFC) and hence suggested its role in mRNA turnover during the same (Widagdo et al. 2016). During the central nervous system development, FTO has also been seen to exert its critical influence. A family carrying R316Q mutation that inactivates FTO enzymatic activity has displayed a range of developmental anomalies and postnatal growth retardation, possibly due to premature cellular senescence, structural and functional brain defects, microcephaly, and psychomotor delay, to name a few (Boissel et al. 2009). This unequivocally suggests the importance of FTO-demethylase in modulating neurogenic development and progression.

m^6A sequencing in a mouse model has also demonstrated the detection of m^6A and m^6Am modifications and their association with the transcripts related to neuronal and synaptic regulation in physiological brain function and under stress-induced conditions (Engel et al. 2018; Koranda et al. 2018). The deletion of *Mettl14* in striatonigral and striatopallidal neurons was seen decreasing the expression of striatal genes related to neurons and synapses and specific genes related to each neuronal subtype. In contrast, conditional knockouts for *Mettl3* and *Fto* did not show any alteration in anxiety and locomotion, but showed an increase in fear memory (Engel et al. 2018). Moreover, depletion of *Mettl14* was also found to be associated with impaired learning and locomotion (Koranda et al. 2018). Additionally, both *Mettl3* or *Mettl14* in flies have also been shown to be involved in sex determination and neuronal function with their knockouts causing impaired locomotion in them (Lence et al. 2016), whereas the same in Zebrafish was seen causing morphological, ectoderm, and haematopoietic defects (Ping et al. 2014; Zhang et al. 2017a). Overall, these studies depict a regulatory role of RNA modification in neural development and cognitive function. The subsequent section would highlight the association between altered epitranscriptome and some of the neural defects, including cancer.

4.3 Altered Epitranscriptome in Neurodevelopmental and Neurodegenerative Diseases

Only a couple of studies so far have revealed the association between RNA modification and neurodegenerative diseases. As discussed in the previous section**,** FTO seems to play a crucial role during the development of the nervous system in humans. Mouse embryos exhibiting deletion of *Fto* locus also display abnormalities in brain patterning, such as defective hypothalamus and telencephalon development (Jung and Goldman 2018). Additionally, subsequent studies conducted using meta-analysis along with pathway and enrichment analysis by the Genetics Core of the Alzheimer's disease Neuroimaging Initiative (ADNI) have demonstrated the association of reduced *Fto* gene expression with AD (Jung and Goldman 2018).

Quantification conducted on m^6A methylation using APP/PS1 transgenic AD mouse model has revealed a notable increase in m^6A methylation in the cortex and the hippocampus of the brain (Han et al. 2020). They have further observed upregulation of *Mettl3* and downregulation of *Fto* in AD mice suggesting that m^6A modification may promote the development of AD. However, concerning *sporadic Parkinson's disease (PD),* no such significant association of m^6A modification regulators includes METTL3, METTL14, WTAP, ALKBH, FTO, YTHDF1, YTHDF2, YTHDF3, ELAVL1, and HNRNPC with PD was noted (Qin et al. 2020). Traumatic brain injury (TBI), on the other hand, has been reported to involve downregulation in *Mettl14* and *Fto* genes, causing a significant alteration in the methylation pattern in 175 mRNAs. Furthermore, FTO has been shown to be essential for the repair of neurological damage caused by TBI while exerting no effect on spatial learning and memory of TBI rats (Yu et al. 2020). On the contrary, post-transcriptional modifications in tRNAs may result in brain dysfunction, and tRNA methyltransferase-1, which is required for catalyzing the demethylation of guanosines in tRNA causes cognitive disorder in human (Liu and Straby 2000; Dunin-Horkawicz et al. 2006; Najmabadi et al. 2011). Never-the-less, further studies are required to gain a concrete understanding of the role of epitranscriptomics in neurodevelopment and associated neurodegenerative disorders.

4.4 Altered Epitranscriptome and Glioblastoma

In recent years, various evidences have suggested the role of m^6A modification in human carcinogenesis. The m^6A regulators; the writers, erasers, and readers have been implicated to play varied role in carcinogenesis such as cancer stem cell formation, cancer metabolism, epithelial to mesenchymal transition (EMT), signal transduction, and apoptosis by controlling mRNA stability and translation (Lin et al. 2019; Zhou et al. 2020). Li et al. (2019) have identified widespread alteration in genetic signatures pertaining to m^6A regulators by analyzing more than 10,000 patient samples across 33 different cancer types. Those alterations were found to have a strong correlation with tumour subtype-specific activation and inhibition of pathways involved in cancer.

Glioblastoma multiforme (GBM) is one of the most aggressive and debilitating cancers of the central nervous system. However, the know-how on m^6A modification in glioma cells is relatively new. Only a couple of reports have appeared in the past years, and also with most of the studies being carried out using glioblastoma stem cells (GSCs) as the model system to explore the role of RNA modification in the progression of tumorigenesis (Fig. 4). Indeed, the very first study on m^6A modification in glioblastoma using the GSCs model has revealed the implication of m^6A methylation in GSCs self-renewal and tumorigenesis (Cui et al. 2017). They have seen knockdown of *Mettl3* or *Mettl14* enhancing GSCs proliferation and tumorigenesis, while either overexpression of *Mettl3* or knockdown of *Fto* has led to suppression in GSCs growth and tumour progression in vivo. Further, they have

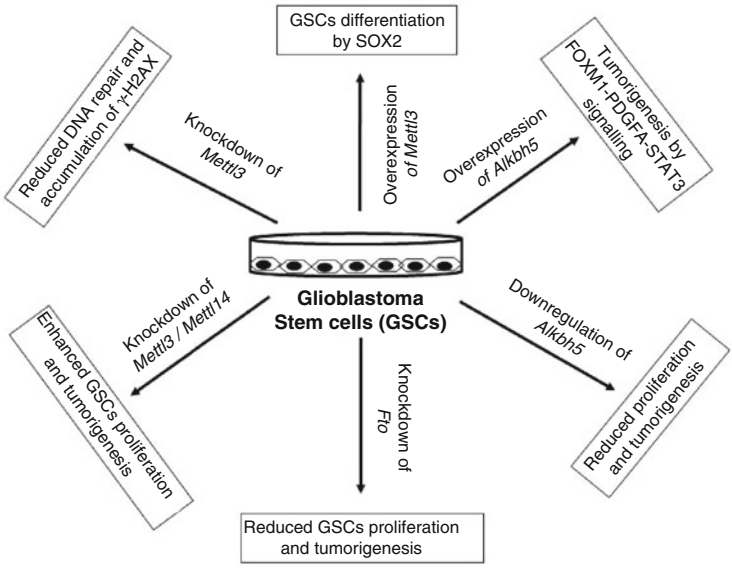


Fig. 4 Association of altered m^6A methylation status with glioblastoma. m^6A modification regulates proliferation, differentiation, and self-renewal of glioblastoma stem cells (GSCs). The loss of function of *Mettl3* and *Mettl14*, the m^6A writers, promotes proliferation and tumorigenesis in GSCs accompanied by reduced DNA repair and accumulation of γ-H2AX in them. However, the opposite is true in the case of gain of function of *Mettl3*. Similarly, the cessation in GSCs growth and tumorigenesis is seen in case of loss of function of *Fto or Alkbh5*, the m^6A readers, whereas gain of function of *Alkbh5* leads to enhanced tumorigenesis

observed upregulation of several oncogenes upon *Mettl3* or *Mettl14* depletion, and the same being downregulated upon *Fto* inhibition, thereby reinforcing the association of m^6A methylation status with glioblastoma onset and progression. In parallel, another group has revealed that m^6A demethylase ALKBH5 is highly expressed in GSCs as compared to healthy tissues or cells and glioblastoma cell lines (Zhang et al. 2017b). In fact, the downregulation of ALKBH5 has been shown to be associated with reduced proliferation of patient-derived GSCs. Further, they have provided insight into the mechanism for understanding the role of m^6A methylation in glioblastoma. Using m^6A sequencing analysis and integrated transcriptome, it has been demonstrated that ALKBH5 demethylates and causes increased expression of FOXM1, a known mediator of GSC self-renewal and tumorigenesis through the mediation of PDGF-A/STAT3 signalling (Gong et al. 2015; Zhang et al. 2017b). Accordingly, the depletion of the interaction between ALKBH5 and FOXM1 has potentiated reduction in GSCs tumorigenesis, indicating its therapeutic relevance and further exploration thereof. The influence of m^6A modification on glioblastoma has also been substantiated by observing downregulation of m^6A levels in glioma tissues and accordingly increasing the expression of m^6A in glioma cell line (U251 cells) has been proven to facilitate reduced migration and proliferation along with induction of apoptosis in these cells (Li et al. 2019b).

A recent study has shown that imidazobenzoxazin-5-thione MV1035, a sodium channel blocker, inhibits the m^6A demethylase ALKBH5 in U87-MG glioblastoma cell line (Malacrida et al. 2020). This results in reduced cellular migration and invasiveness and hence validates the role of RNA modification in glioblastoma tumorigenesis. Similarly, the mutation in the isocitrate dehydrogenase 1/2 (IDH1/2) enzymes has been reported to facilitate the production of oncometabolite R-2-hydroxyglutarate (R-2HG) that is shown to exert anti-tumour activity in glioma cells through increased m^6A methylation (Su et al. 2017). In leukaemia cells as well, R-2HG causes increased m^6A methylation by inhibiting FTO activity, and that further leads to decreased stability of *Myc/Cebpa* transcripts and reduced proliferation/survival of cancer cells (Su et al. 2017). Moreover, it has also been reported that WTAP component of RNA methyltransferase is highly expressed in glioblastoma cells making it more tumorigenic by regulating migration and invasion of cancer cells (Jin et al. 2012). A decreased expression of METTL3 during GSCs differentiation has also revealed the importance of METTL3 mediated m^6A methylation in GSCs neurosphere formation and dedifferentiation (Visvanathan et al. 2017). Further through RNA immunoprecipitation, they have identified SOX2 expressed in neural progenitors and differentiated neurons, as the target of METTL3 and both METTL3 and Human antigen R (HuR) have been found essential for maintaining the stability of SOX2. Considering the association of METTL3 with DNA repair and that is mediated through transcription factor SOX2, *Mettl3* deficient GSCs have been proven to be more sensitive to γ-irradiation accompanied by accumulation of γ-H2AX. In addition, METTL3 has also been found to be highly expressed in glioblastoma tumour, and silencing of *Mettl3* in U87/TIC has been shown to inhibit tumour growth in vivo as well (Visvanathan et al. 2017). They have further demonstrated METTL3 as essential for the expression of the actively transcribed genes important for the maintenance of GSCs, with its silencing in GSCs resulting in the global downregulation of tumorigenic pathways important for GSC maintenance and glioma progression (Visvanathan et al. 2019). Another group through integrated transcriptome and MeRIP-seq analyses has demonstrated the role of m^6A methyltransferase METTL3 in glioblastoma tumour growth and progression by modulating nonsense-mediated RNA decay (NMD) (Li et al. 2019a). They have observed that METTL3 modulates the expression of serine and arginine-rich splicing factors (SRSF) and promotes the YTHDC1 dependent NMD of SRSF transcript, which further leads to the alteration in splicing events. In addition, the phenotype mediated by *Mettl3* deficiency was found to be rescued by decreasing the expression of *Bcl-X* or *Ncor2* isoform. Another report has described the differential expression of 13 m^6A RNA methylation regulators in glioma samples (Chai et al. 2019). While there existed a direct correlation between the expression level of WTAP, RBM15, ALBKH5, YTHDF2 and WHO grade, an inverse correlation was noted with FTO and intriguingly no correlation with METTL3, METTL14, or METTL16. They further revealed FTO and YTHDC1 as protective genes, while WTAP, ALBKH5, YTHDF1, YTHDF2, and RBM15 were assigned as risky genes. Based on these signature genes, they revealed two subgroups of gliomas, i.e., RM1 and RM2. RM2 subgroup was found to be associated with a poorer prognosis, lower frequency of

IDH mutation, and higher WHO grade as compared to RM1 subgroup. Moreover, the markers of EMT and TNF-α signalling via NF-kβ were also present in RM2 subgroup, suggesting a close association between m^6A modification and the progression of glioblastoma (Chai et al. 2019). Similarly concerning m^5C modification, NSUN RNA methyltransferase modulating m^5C has been demonstrated to be associated with tumour suppressor function in several in vivo mouse models of glioma cells (Janin et al. 2019). They have found the C3782 position of 28S rRNA that is implicated in protein synthesis as the target site for NSUN5. In addition, the inactivation of NSUN5 was seen increasing the sensitivity of glioma cells towards bioactivable substrates of the stress-related enzyme NQO1 (Janin et al. 2019). Taken together, these studies do shed light on the key role that the RNA modification machinery may play in GSCs maintenance, tumour growth, and progression of glioblastoma.

Based on the afore-stated findings by various investigators, it may be very well inferred that RNA modifications and the associated regulators with their contextual expression and activities may exert a profound influence on the development and physiology of neural cells. Alteration or dysregulation in the same can lead to neurodevelopmental and neurodegenerative disorders, including that of cancer. Hence, a concrete understanding of the players and their precise mode of dynamic cell specific and balanced action may eventually help devising strategies to restore the normal epitranscriptomic signature in them and the associated function thereof.

5 Challenges and Future Perspective

The human brain increases in size with the expansion of NPCs during development and diversifies with its fate specification. The complex nature and heterogeneity of the brain pose a major challenge to study its development. So far, studies using mouse models have uncovered many aspects of brain development, yet we lack the knowledge on the mechanism underlying its complexity. Over the recent years, tremendous progress has been made to understand the gene regulatory network controlling cell fate transition from NPCs to neural cells during brain development. Single-cell transcriptomic analysis has revealed the transcriptional profile of gene expression for the generation of neural cells (Nowakowski et al. 2017; Zhong et al. 2018). However, the field of epitranscriptomics being still evolving and upcoming; the involvement of the same and the mechanism governing NSCs fate transition is not yet well understood. Table 1 has depicted the epitranscriptomic machinery identified so far and the implication in RNA metabolism and function. Nevertheless, there is also a possibility for the presence of additional regulators of RNA modification for each subtype of RNA. In fact, the investigation concerning neural fate specification and the role of these players in RNA modification during the same remains largely unexplored. Delineating the same may facilitate determining/defining the presence of each regulator in different brain regions and their functional attributes in neurogenesis and neurodevelopment. Furthermore, signalling pathways

implicated in RNA modification in driving neurogenesis during embryonic brain development in a temporo-spatial manner and how RNA modification is mediating the complexity of the brain in demarcating the different regions of brain for its proper function and repair process also remain to be explored. Albeit the rapid advancement in the high-throughput sequencing methods that has enabled us to understand the epitranscriptomic modification in axon regeneration, neural fate transition, memory formation, and synaptic function, the concrete mechanistic basis underlying the same is still unclear. Moreover, considering the species-specific differences noted concerning m^6A methylation and their association with genetic disorders pertaining to mental disorders, which seems to be more prominent in the human foetal cortex than that in mouse ones at comparable development stages, it signifies the species-specific developmental complexity and the difference in gene regulatory network module during brain development (Yoon et al. 2017). Therefore human ESCs/iPSCs derived 3D gastruloid and organoid model systems can serve as elegant platforms to study the nitty-gritty of human brain development in a dish and the mechanism underlying epitranscriptomic modulation during its formation and associated developmental disorders. Furthermore, the association of m^6A methylation with glioblastoma suppression presents another exciting avenue for exploration of epitranscriptomic drug candidates that can modulate the cell fate transition and thereby participate in the prevention of or as a curative option for addressing dreaded neurological disorders. The field is still evolving and further studies with refined RNA modification mapping tools and technologies such as miCLIP, meRIP-Seq, nanopore sequencing, mass-spectrometry, etc. alongside AI- and machine learning-based approaches for screening for small molecule inhibitors that can have target specificity for epitranscriptomic modifiers/proteins can eventually lead to identifying good drug targets. Undoubtedly, uncovering the epitranscriptomic landscape with additional specificity shall eventually shed light on the gene regulatory network underlying neurodevelopment, cognitive function, and associated disorders and finding the curative options thereof.

References

Abakir A, Giles TC, Cristini A et al (2020) N^6-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. Nat Genet 52:48–55

Abbasi-Moheb L, Mertel S, Gonsior M et al (2012) Mutations in NSUN2 cause autosomal-recessive intellectual disability. Am J Hum Genet 90:847–855

Aguilo F, Zhang F, Sancho A et al (2015) Coordination of m^6A mRNA methylation and gene transcription by ZFP217 regulates pluripotency and reprogramming. Cell Stem Cell 17:689–704

Alexandrov A, Grayhack EJ, Phizicky EM (2005) tRNA m^7G methyltransferase Trm8p/Trm82p: evidence linking activity to a growth phenotype and implicating Trm82p in maintaining levels of active Trm8p. RNA 11:821–830

Amort T, Rieder D, Wille A et al (2017) Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain. Genome Biol 18:1

Angelova MT, Dimitrova DG, Dinges N et al (2018) The emerging field of epitranscriptomics in neurodevelopmental and neuronal disorders. Front Bioeng Biotechnol 6:46

Arguello E, Deliberto AN, Kleiner RE (2017) RNA chemical proteomics reveals the N^6-methyladenosine (m^6A)-regulated protein–RNA interactome. J Am Chem Soc 139:17249–17252

Basanta-Sanchez M, Wang R, Liu Z et al (2017) TET1-mediated oxidation of 5-Formylcytosine (5fC) to 5-Carboxycytosine (5caC) in RNA. Chembiochem 18:72–76

Batista PJ, Molinie B, Wang J et al (2014) m^6A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell 15:707–719

Blanco S, Dietmann S, Flores JV et al (2014) Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. EMBO J 33:2020–2039

Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307

Boissel S, Reish O, Proulx K et al (2009) Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. Am J Hum Genet 85:106–111

Byszewska M, Śmietański M, Purta E et al (2014) RNA methyltransferases involved in 5′ cap biosynthesis. RNA Biol 11:1597–1607

Camara Y, Asin-Cayuela J, Park CB et al (2011) MTERF4 regulates translation by targeting the methyltransferase NSUN4 to the mammalian mitochondrial ribosome. Cell Metab 13:527–539

Carlile TM, Rojas-Duran MF, Zinshteyn B et al (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515:143–146

Cartlidge RA, Knebel A, Peggie M et al (2005) The tRNA methylase METTL1 is phosphorylated and inactivated by PKB and RSK in vitro and in cells. EMBO J 24:1696–1705

Chai RC, Wu F, Wang QX et al (2019) m^6A RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas. Aging 11:1204–1225

Chen T, Hao YJ, Zhang Y et al (2015) m^6A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. Cell Stem Cell 16:289–301

Chen X, Gao Y, Yang L et al (2018) Speech and language delay in a patient with WDR4 mutations. Eur J Med Genet 61:468–472

Chen J, Zhang YC, Huang C et al (2019a) m^6A regulates neurogenesis and neuronal development by modulating histone methyltransferase Ezh2. Genomics Proteomics Bioinformatics 17:154–168

Chen X, Li A, Sun BF et al (2019b) 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. Nat Cell Biol 21:978–990

Chen YS, Ma HL, Yang Y et al (2019c) 5-Methylcytosine analysis by RNA-BisSeq. Methods Mol Biol 1870:237–248

Cheray M, Etcheverry A, Jacques C et al (2020) Cytosine methylation of mature microRNAs inhibits their functions and is associated with poor prognosis in glioblastoma multiforme. Mol Cancer 19:36

Chernyakov I, Whipple JM, Kotelawala L et al (2008) Degradation of several hypomodified mature tRNA species in *Saccharomyces cerevisiae* is mediated by Met22 and the 5′-3′ exonucleases Rat1 and Xrn1. Genes Dev 22:1369–1380

Choi J, Ieong KW, Demirci H et al (2016) N^6-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. Nat Struct Mol Biol 23:110–115

Cholewa-Waclaw J, Bird A, von Schimmelmann M et al (2016) The role of epigenetic mechanisms in the regulation of gene expression in the nervous system. J Neurosci 36:11427–11434

Chou HJ, Donnard E, Gustafsson HT et al (2017) Transcriptome- wide analysis of roles for tRNA modifications in translational regulation. Mol Cell 68:978–992

Cui Q, Shi H, Ye P et al (2017) m^6A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. Cell Rep 18:2622–2634

Delatte B, Wang F, Ngoc LV et al (2016) Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. Science 351:282–285

Delorimier E, Hinman MN, Copperman J et al (2017) Pseudouridine modification inhibits muscleblind-like 1 (MBNL1) binding to CCUG repeats and minimally structured RNA through reduced RNA flexibility. J Biol Chem 292:4350–4357

Desrosiers RC, Friderici KH, Rottman FM (1975) Characterization of Novikoff hepatoma mRNA methylation and heterogeneity in the methylated 5′ terminus. Biochemist 14:4367–4374

Dewe JM, Whipple JM, Chernyakov I et al (2012) The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. RNA 18:1886–1896

Dimitrova DG, Teysset L, Carré C (2019) RNA 2′-O-methylation (Nm) modification in human diseases. Genes (Basel) 10:117

Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. Nature 485:201–206

Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S et al (2016) The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. Nature 530:441–446

Dong A, Yoder JA, Zhang X et al (2001) Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. Nucleic Acids Res 29:439–448

Duan J, Li L, Lu J et al (2009) Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. Mol Cell 34:427–439

Dunin-Horkawicz S, Czerwoniec A, Gajda MJ (2006) MODOMICS: a database of RNA modification pathways. Nucleic Acids Res 34:D145–D149

Edupuganti RR, Geiger S, Lindeboom R et al (2017) N^6-methyladenosine (m^6A) recruits and repels proteins to regulate mRNA homeostasis. Nat Struct Mol Biol 24:870–878

El Yacoubi BE, Bailly M, de Crécy-Lagard V (2012) Biosynthesis and function of posttranscriptional modifications of transfer RNAs. Annu Rev Genet 46:69–95

Engel M, Eggert C, Kaplick PM et al (2018) The role of m^6A/m-RNA methylation in stress response regulation. Neuron 99:389–403

Erales J, Marchand V, Panthu B et al (2017) Evidence for rRNA 2′-O-methylation plasticity: control of intrinsic translational capabilities of human ribosomes. Proc Natl Acad Sci USA 114:12934–12939

Fei Q, Zou Z, Roundtree IA et al (2020) YTHDF2 promotes mitotic entry and is regulated by cell cycle mediators. PLoS Biol 18:e3000664

Flores JV, Cordero-Espinoza L, Oeztuerk-Winder F et al (2017) Cytosine-5 RNA methylation regulates neural stem cell differentiation and motility. Stem Cell Rep 8:112–124

Fu Y, Jia G, Pang X et al (2013) FTO-mediated formation of N^6-hydroxymethyladenosine and N^6-formyladenosine in mammalian RNA. Nat Commun 4:1798

Fu L, Guerrero CR, Zhong N et al (2014a) Tet-mediated formation of 5-Hydroxymethylcytosine in RNA. J Am Chem Soc 136:11582–11585

Fu Y, Dominissini D, Rechavi G et al (2014b) Gene expression regulation mediated through reversible m^6A RNA methylation. Nat Rev Genet 15:293–306

García-Campos MA, Edelheit S, Toth U et al (2019) Deciphering the "m^6A code" via antibody-independent quantitative profiling. Cell 178:731–747

Geula S, Moshitch-Moshkovitz S, Dominissini D et al (2015) m^6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science 347:1002–1006

Gigova A, Duggimpudi S, Pollex T et al (2014) A cluster of methylations in the domain IV of 25S rRNA is required for ribosome stability. RNA 20:1632–1644

Gkatza NA, Castro C, Harvey RF et al (2019) Cytosine- 5 RNA methylation links protein synthesis to cell metabolism. PLoS Biol 17:e3000297

Goll MG, Kirpekar F, Maggert KA et al (2006) Methylation of $tRNA^{Asp}$ by the DNA methyltransferase homolog Dnmt2. Science 311:395–398

Gong AH, Wei P, Zhang S et al (2015) FoxM1 drives a feed-forward STAT3-activation signaling loop that promotes the self-renewal and tumorigenicity of glioblastoma stem-like cells. Cancer Res 75:2337–2348

Guy MP, Phizicky EM (2014) Two-subunit enzymes involved in eukaryotic post-transcriptional tRNA modification. RNA Biol 11:608–616

Haag S, Warda AS, Kretschmer J et al (2015) NSUN6 is a human RNA methyltransferase that catalyzes formation of m5C72 in specific tRNAs. RNA 21:1532–1543

Han M, Liu Z, Xu Y et al (2020) Abnormality of m^6A mRNA methylation is involved in Alzheimer's disease. Front Neurosci 14:98

Haran V, Lenka N (2019) Deciphering the epitranscriptomic signatures in cell fate determination and development. Stem Cell Rev Rep 15:474–496

Heck AM, Russo J, Wilusz J et al (2020) YTHDF2 destabilizes m^6A-modified neural-specific RNAs to restrain differentiation in induced pluripotent stem cells. RNA 26:739–755

Heiss NS, Bächner D, Salowsky R et al (2000) Gene structure and expression of the mouse dyskeratosis congenita gene, dkc1. Genomics 67:153–163

Heng JCD, Ng HH (2010) Transcriptional regulation in embryonic stem cells. Adv Exp Med Biol 695:76–91

Hess ME, Hess S, Meyer KD et al (2013) The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. Nat Neurosci 16:1042

Ho AJ, Stein JL, Hua X et al (2010) A commonly carried allele of the obesity-related FTO gene is associated with reduced brain volume in the healthy elderly. Proc Natl Acad Sci USA 107:8404–8409

Hsu PJ, Shi H, He C (2017a) Epitranscriptomic influences on development and disease. Genome Biol 18:197

Hsu PJ, Zhu Y, Ma H et al (2017b) Ythdc2 is an N^6-methyladenosine binding protein that regulates mammalian spermatogenesis. Cell Res 27:1115–1127

Huang H, Weng H, Sun W et al (2018) Recognition of RNA N^6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol 20:285–295

Huang H, Weng H, Zhou K et al (2019) Histone H3 trimethylation at lysine 36 guides m^6A RNA modification co-transcriptionally. Nature 567:414–419

Huber SM, van Delft P, Mendil L et al (2015) Formation and abundance of 5-hydroxymethylcytosine in RNA. Chembiochem 16:752–755

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

Janin M, Ortiz-Barahona V, De Moura MC et al (2019) Epigenetic loss of RNA-methyltransferase NSUN5 in glioma targets ribosomes to drive a stress adaptive translational program. Acta Neuropathol 138:1053–1074

Jia G, Fu Y, Zhao X et al (2011) N^6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887

Jin DI, Lee SW, Han ME et al (2012) Expression and roles of Wilms' tumor 1-associating protein in glioblastoma. Cancer Sci 103:2102–2109

Jung Y, Goldman D (2018) Role of RNA modifications in brain and behavior. Genes Brain Behav 17:e12444

Kase Y, Shimazaki T, Okano H (2020) Current understanding of adult neurogenesis in the mammalian brain: how does adult neurogenesis decrease with age? Inflamm Regen 40:10

Ke S, Alemu EA, Mertens C et al (2015) A majority of m^6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes Dev 29:2037–2053

Keverne EB, Pfaff DW, Tabansky I (2015) Epigenetic changes in the developing brain: effects on behavior. Proc Natl Acad Sci USA 112:6789–6795

Khan MA, Rafiq MA, Noor A et al (2012) Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. Am J Hum Genet 90:856–863

Khoddami V, Cairns BR (2013) Identification of direct targets and modified bases of RNA cytosine methyltransferases. Nat Biotechnol 31:458–464

Knuckles P, Lence T, Haussmann IU et al (2018) Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m^6A machinery component Wtap/Fl(2)d. Genes Dev 32:415–429

Koranda JL, Dore L, Shi H et al (2018) Mettl14 is essential for epitranscriptomic regulation of striatal function and learning. Neuron 99:283–292

Kretschmer J, Rao H, Hackert P et al (2018) The m^6A reader protein YTHDC2 interacts with the small ribosomal subunit and the 5′-3′ exoribonuclease XRN1. RNA 24:1339–1350

Lee SH, Kim I, Chung BC (2007) Increased urinary level of oxidized nucleosides in patients with mild-to-moderate Alzheimer's disease. Clin Biochem 40:936–938

Lence T, Akhtar J, Bayer M et al (2016) m^6A modulates neuronal functions and sex determination in Drosophila. Nature 540:242–247

Li X, Zhu P, Ma S et al (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol 11:592–597

Li X, Xiong X, Wang K et al (2016a) Transcriptome-wide mapping reveals reversible and dynamic N^1-methyladenosine methylome. Nat Chem Biol 12:311–316

Li X, Xiong X, Yi C (2016b) Epitranscriptome sequencing technologies: decoding RNA modifications. Nat Methods 14:23–31

Li L, Zang L, Zhang F et al (2017) Fat mass and obesity-associated (FTO) protein regulates adult neurogenesis. Hum Mol Genet 26:2398–2411

Li Y, Xiao J, Bai J et al (2019) Molecular characterization and clinical relevance of m^6A regulators across 33 cancer types. Mol Cancer 18:137

Li F, Yi Y, Miao Y et al (2019a) N^6-Methyladenosine modulates nonsense-mediated mRNA decay in human glioblastoma. Cancer Res 79:5785–5798

Li F, Zhang C, Zhang G (2019b) m^6A RNA methylation controls proliferation of human glioma cells by influencing cell apoptosis. Cytogenet Genome Res 159:119–125

Lian H, Li WB, Jin WL (2016) The emerging insights into catalytic or non-catalytic roles of TET proteins in tumors and neural development. Oncotarget 7:64512–64525

Liao S, Sun H, Xu C (2018) YTH domain: a family of N^6-methyladenosine (m^6A) readers. Genomics Proteomics Bioinformatics 16:99–107

Lin S, Liu Q, Lelyveld VS et al (2018) Mettl1/Wdr4-mediated m^7G tRNA methylome is required for normal mRNA translation and embryonic stem cell self-renewal and differentiation. Mol Cell 71:244–255

Lin X, Chai G, Wu Y et al (2019) RNA m^6A methylation regulates the epithelial mesenchymal transition of cancer cells and translation of snail. Nat Commun 10:2065

Linder B, Grozhik AV, Olarerin-George AO et al (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods 12:767–772

Liu N, Pan T (2015a) RNA epigenetics. Transl Res 165:28–35

Liu N, Pan T (2015b) Probing RNA modification status at single-nucleotide resolution in total RNA. Methods Enzymol 560:149–159

Liu J, Straby KB (2000) The human tRNA(m(2)(2)G(26))dimethyltransferase: functional expression and characterization of a cloned hTRM1 gene. Nucleic Acids Res 28:3445–3451

Liu J, Yue Y, Han D et al (2014) A METTL3-METTL14 complex mediates mammalian nuclear RNA N^6-adenosine methylation. Nat Chem Biol 10:93–95

Liu N, Dai Q, Zheng G et al (2015) N^6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature 518:560–564

Liu N, Zhou KI, Parisien M et al (2017) N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. Nucleic Acids Res 45:6051–6063

Lorenz C, Lunse CE, Morl M (2017) tRNA modifications: impact on structure and thermal adaptation. Biomol Ther 7:35

Luo GZ, Blanco MA, Greer EL et al (2015) DNA N^6-methyladenine: a new epigenetic mark in eukaryotes? Nat Rev Mol Cell Biol 16:705–710

Malacrida A, Rivara M, Di Domizio A et al (2020) 3D proteome-wide scale screening and activity evaluation of a new ALKBH5 inhibitor in U87 glioblastoma cell line. Bioorganic Med Chem 28:115300

Martinez FJ, Lee JH, Lee JE et al (2012) Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. J Med Genet 49:380–385

Mauer J, Luo X, Blanjoie A et al (2017) Reversible methylation of m^6A_m in the 5′ cap controls mRNA stability. Nature 541:371–375

Mauer J, Sindelar M, Despic V et al (2019) FTO controls reversible m^6Am RNA methylation during snRNA biogenesis. Nat Chem Biol 15:340–347

Meyer KD, Jaffrey SR (2014) The dynamic epitranscriptome: N^6-methyladenosine and gene expression control. Nat Rev Mol Cell Biol 15:313–326

Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′UTRs and near stop codons. Cell 149:1635–1646

Meyer KD, Patil DP, Zhou J et al (2015) 5′ UTR m^6A promotes cap-independent translation. Cell 163:999–1010

Miao W, Chen J, Jia L et al (2019) The m^6A methyltransferase METTL3 promotes osteosarcoma progression by regulating the m^6A level of LEF1. Biochem Biophys Res Commun 516:719–725

Molinie B, Wang J, Lim KS et al (2016) m^6A-LAIC-seq reveals the census and complexity of the m^6A epitranscriptome. Nat Methods 13(8):692–698

Najmabadi H, Hu H, Garshasbi M (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature 478:57–63

Niu Y, Zhao X, Wu YS et al (2013) N^6-methyl-adenosine (m^6A) in RNA: an old modification with a novel epigenetic function. Genomics Proteomics Bioinform 11:8–17

Noack F, Calegari F (2018) Epitranscriptomics: a new regulatory mechanism of brain development and function. Front Neurosci 12:85

Nowakowski TJ, Bhaduri A, Pollen AA et al (2017) Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. Science 358:1318–1323

Patil DP, Pickering BF, Jaffrey S (2017) Reading m^6A in the transcriptome: m^6A -binding proteins. Trends Cell Biol 28:113–127

Peifer C, Sharma S, Watzinger P et al (2013) Yeast Rrp8p, a novel methyltransferase responsible for m^1A 645 base modification of 25S rRNA. Nucleic Acids Res 41:1151–1163

Pendleton KE, Chen B, Liu K et al (2017) The U6 snRNA m^6A methyltransferase METTL16 regulates SAM synthetase intron retention. Cell 169:824–835

Picard-Jean F, Brand C, Tremblay-Létourneau M et al (2018) 2′-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. PLoS One 13:e0202308

Ping XL, Sun BF, Wang L et al (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res 24:177–189

Qin L, Min S, Shu L et al (2020) Genetic analysis of N6-methyladenosine modification genes in Parkinson's disease. Neurobiol Aging 93:143.e9–143.e13

Ramos J, Fu D (2019) The emerging impact of tRNA modifications in the brain and nervous system. Biochim Biophys Acta Gene Regul Mech 1862:412–428

Roundtree IA, Evans ME, Pan T et al (2017a) Dynamic RNA modifications in gene expression regulation. Cell 169:1187–1200

Roundtree IA, Luo GZ, Zhang Z et al (2017b) YTHDC1 mediates nuclear export of N^6-methyladenosine methylated mRNAs. elife 6:e31311

Schaefer M, Pollex T, Hanna K et al (2009) RNA cytosine methylation analysis by bisulfite sequencing. Nucleic Acids Res 37:e12

Schevitz RW, Podjarny AD, Krishnamachari N et al (1979) Crystal structure of a eukaryotic initiator tRNA. Nature 278:188–190

Schwartz S, Agarwala SD, Mumbach MR et al (2013) High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell 155:1409–1421

Schwartz S, Mumbach MR, Jovanovic M et al (2014) Perturbation of m^6A writers reveals two distinct classes of mRNA methylation at internal and 5′ sites. Cell Rep 8:284–296

Sharma S, Yang J, Watzinger P et al (2013) Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. Nucleic Acids Res 41:9062–9076

Shi M, Zhang H, Wu X et al (2017) ALYREF mainly binds to the 5′ and the 3′ regions of the mRNA in vivo. Nucleic Acids Res 45:9640–9653

Shi H, Zhang X, Weng YL et al (2018) m^6A facilitates hippocampus-dependent learning and memory through YTHDF1. Nature 563:249–253

Shu X, Cao J, Cheng M et al (2020) A metabolic labeling method detects m^6A transcriptome-wide at single base resolution. Nat Chem Biol 16:887–895

Somme J, Van Lae B, Roovers M et al (2014) Characterization of two homologous 2′-O-methyltransferases showing different specificities for their tRNA substrates. RNA 20:1257–1271

Su R, Dong L, Li C et al (2017) R-2HG exhibits anti-tumor activity by targeting FTO/m^6A/MYC/CEBPA signaling. Cell 172:90–105

Traube FR, Carell T (2017) The chemistries and consequences of DNA and RNA methylation and demethylation. RNA Biol 14:1099–1107

Trixl L, Amort T, Wille A et al (2018) RNA cytosine methyltransferase Nsun3 regulates embryonic stem cell differentiation by promoting mitochondrial activity. Cell Mol Life Sci 75:1483–1497

Tuorto F, Liebers R, Musch T et al (2012) RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat Struct Mol Biol 19:900–905

Uttara B, Singh AV, Zamboni P et al (2009) Oxidative stress and neurodegenerative diseases: a review of upstream and downstream antioxidant therapeutic options. Curr Neuropharmacol 7:65–74

Van Haute L, Dietmann S, Kremer L et al (2016) Deficient methylation and formylation of mt-tRNA (met) wobble cytosine in a patient carrying mutations in NSUN3. Nat Commun 7:12039

Visvanathan A, Patil V, Arora A et al (2017) Essential role of METTL3-mediated m^6A modification in glioma stem-like cells maintenance and radioresistance. Oncogene 37:522–533

Visvanathan A, Patil V, Abdulla S et al (2019) N^6-methyl adenosine landscape of glioma stem-like cells: METTL3 is essential for the expression of actively transcribed genes and sustenance of the oncogenic signalling. Genes (Basel) 10:141

Walters BJ, Mercaldo V, Gillon CJ et al (2017) The role of the RNA demethylase FTO (fat mass and obesity-associated) and mRNA methylation in hippocampal memory formation. Neuropsychopharmacology 42:1502–1510

Wang CC, Lo JC, Chien CT et al (2011) Spatially controlled expression of the Drosophila pseudouridine synthase RluA-1. Int J Dev Biol 55:223–227

Wang X, Feng J, Xue Y et al (2016) Structural basis of N^6-adenosine methylation by the METTL3-METTL14 complex. Nature 534:575–578

Wang Y, Li Y, Yue M et al (2018) N^6-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications. Nat Neurosci 21:195–206

Warda AS, Kretschmer J, Hackert P et al (2017) Human METTL16 is a N^6-methyladenosine (m^6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep 18:2004–2014

Wei J, Liu F, Lu Z et al (2018) Differential m^6A, m^6A_m, and m^1A demethylation mediated by FTO in the cell nucleus and cytoplasm. Mol Cell 71:973–985

Weng YL, Wang X, An R et al (2018) Epitranscriptomic m^6A regulation of axon regeneration in the adult mammalian nervous system. Neuron 97:313–325

Werner M, Purta E, Kaminska KH et al (2011) 2′-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. Nucleic Acids Res 39:4756–4768

Widagdo J, Zhao QY, Kempen MJ et al (2016) Experience-dependent accumulation of N^6-methyladenosine in the prefrontal cortex is associated with memory processes in mice. J Neurosci 36:6771–6777

Woodcock CB, Horton JR, Zhou J et al (2020) Biochemical and structural basis for YTH domain of human YTHDC1 binding to methylated adenine in DNA. Nucleic Acids Res 48:10329–10341
Xiang Y, Laurent B, Hsu CH et al (2017) RNA m^6A methylation regulates the ultraviolet-induced DNA damage response. Nature 543:573–576
Xue S, Xu H, Sun Z et al (2019) Depletion of TRDMT1 affects 5-methylcytosine modification of mRNA and inhibits HEK293 cell proliferation and migration. Biochem Biophys Res Commun 520:60–66
Xue C, Zhao Y, Li L (2020) Advances in RNA cytosine-5 methylation: detection, regulatory mechanisms, biological functions and links to cancer. Biomark Res 8:43
Yang X, Yang Y, Sun BF et al (2017) 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m^5C reader. Cell Res 27:606–625
Yang Y, Wang L, Han X et al (2019) RNA 5-methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA decay. Mol Cell 75:1188–1202
Yao B, Christian KM, He C et al (2016) Epigenetic mechanisms in neurogenesis. Nat Rev Neurosci 17:537–549
Yoon KJ, Ringeling FR, Vissers C et al (2017) Temporal control of mammalian cortical neurogenesis by m^6A methylation. Cell 171:877–889
Yu J, Zhang Y, Ma H et al (2020) Epitranscriptomic profiling of N6-methyladenosine-related RNA methylation in rat cerebral cortex following traumatic brain injury. Mol Brain 13:11
Zhang C, Chen Y, Sun B et al (2017a) m^6A modulates haematopoietic stem and progenitor cell specification. Nature 549:273–276
Zhang S, Zhao BS, Zhou A et al (2017b) m^6A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. Cancer Cell 31:591–606
Zhang Z, Chen LQ, Zhao YL et al (2019) Single-base mapping of m^6A by an antibody-independent method. Sci Adv 5:eaax0250
Zhao BS, Roundtree IA, He C (2017) Post-transcriptional gene regulation by mRNA modifications. Nat Rev Mol Cell Biol 18:31–42
Zhao YL, Liu YH, Wu RF et al (2019) Understanding m^6A function through uncovering the diversity roles of YTH domain-containing proteins. Mol Biotechnol 61:355–364
Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29
Zhong S, Zhang S, Fan X et al (2018) A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. Nature 555:524–528
Zhou J, Wan J, Gao X et al (2015) Dynamic m^6A mRNA methylation directs translational control of heat shock response. Nature 526:591–594
Zhou KI, Shi H, Lyu R et al (2019) Regulation of co-transcriptional pre-mRNA splicing by m^6A through the low-complexity protein hnRNPG. Mol Cell 76:70–81
Zhou Z, Lv J, Yu H et al (2020) Mechanism of RNA modification N6-methyladenosine in human cancer. Mol Cancer 19:104
Zhu LY, Zhu YR, Dai DJ et al (2018) Epigenetic regulation of alternative splicing. Am J Cancer Res 8:2346–2358

Epitranscriptomics and Diseases

Masamitsu Konno and Hideshi Ishii

Contents

M. Konno · H. Ishii (✉)
Osaka University Graduate School of Medicine, Center of Medical Innovation and Translational Research, Osaka, Japan
e-mail: hishii@gesurg.med.osaka-u.ac.jp

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_4

Abstract The role of epitranscriptomics, i.e., RNA base modification, as a component of a "methylome" in disease has emerged as a result of the development of next generation sequencing and other related state-of-art technologies. Epitranscriptomic state is controlled by writing, erasing, and reading methylation, which is mediated by enzymatic reactions. More than 20 methyltransferases have been identified so far. Epitranscriptomic mechanisms are involved in the development of diseases such as cancer of mice and humans. In this review article, we discuss recent developments in epitranscriptomics for the further application of epitranscriptomic knowledge for use as diagnostic tools and therapeutic approaches.

Keywords Epitranscriptomics · RNA · Methylation · Metabolism · Human diseases · Cancer

1 Introduction

The epigenome, an external or additional moiety of the genome that is without nucleotide sequence changing, is essential for sequentially occurring each cell-specific expression throughout living organisms and maintaining cells in a functionally appropriate state. The epigenetic residues are the results of chemical modification of nucleotides, i.e., DNA or RNA, by methylation, and histone modification through reactions including methylation, acetylation, and phosphorylation. Since RNA is transcribed from DNA in the nucleus, the process of RNA modification is known as epitranscriptomics (Konno et al. 2019a). Thus, the epitranscriptome can be considered as a component of the "methylome," that is, the methylation information of nucleotides and proteins.

2 RNA Species for Epitranscriptomics

RNA modifications have been identified in messenger (m)RNA, ribosomal (r)RNA, transfer (t)RNA, small nuclear (sn)RNA, small nucleolar (sno)RNA, and other non-coding (nc)RNAs in various species, including human (Frye et al. 2018; Wang and He 2014). Whereas the majority of modifications occur in tRNA and rRNA, many modifications also occur in other RNA species, including mRNA, snRNA, and snoRNA. Indeed, the chemical diversity of RNA modifications play facilitates their multiple roles in biological processes, including various during developmental stages, through to stem cell differentiation (Ontiveros et al. 2019). RNA modifications of non-coding RNAs have been demonstrated, and their roles in the control of gene expression have been studied (Coker et al. 2019). Various chemical modifications and their reactions have been investigated in each base of the RNAs, though some enzymatic reactions are not fully understood.

3 Epitranscriptomics Enzymes

Analogous to histone modifications, RNA modifications can be classified into three steps, writing, erasing, and reading RNAs, by the recognition of common motifs, and the enzymes involved in each step have been investigated (Konno et al. 2019a). For example, in the majority of mammalian N^6-methyl adenosine (m^6A) sites, the consensus sequence is typically GGACU; more precisely, RRm^6ACH (R notes G or A; H includes A, C, or U), or Pu (G > A) m^6AC (A/C/U) (where Pu represents purine), though there are some exceptions (Liu et al. 2014). Demethylases that reverse methylation, known as erasing enzymes, have been reported and include fat mass and obesity-associated protein (FTO) (Jia et al. 2011)and α-ketoglutarate-dependent dioxygenase (AlkB) homolog 5 (ALKBH5) (Zheng et al. 2013) (Fig. 1). Moreover, reading enzymes recognize the methylation and exert the function of different binding proteins to affect the translation status and lifetime of RNA (Meyer and Jaffrey 2014, 2017; Wang et al. 2014); examples include a protein family of heterogeneous nuclear ribonucleoproteins (hnRNP) and YT521-B homology (YTH) m^6A RNA-binding protein 1 (YTHDF1). A recent study demonstrated that the oncogene c-myc promotes the expression of YTHDF1 in colorectal cancer (Nishizawa et al. 2017). Furthermore, YTHDF2, another YTH domain family member, is involved in the localization of bound mRNA from the translatable pool to mRNA decay sites via the carboxy-terminal domain of YTHDF2, selectively binding to m^6A-containing mRNA (Wang et al. 2014). The methyltransferases are writing enzymes and function to add the methyl group to RNAs.

3.1 METTL1

Methyltransferase Like (METTL) 1 is a methyltransferase that transforms guanine in RNA into N^7-methyl-guanine (m^7G) and has roles in tRNA, mRNA, and microRNA (miRNA). (Alexandrov et al. 2002; Zhang et al. 2019; Pandolfini et al. 2019). Specifically, it catalyzes the formation of m^7G at position 46 (M^7G46) in tRNA (Alexandrov et al. 2002) (Fig. 2). METTL1 also acts as a methyltransferase for guanine present in the coding sequence region of mRNA to m^7G. It is known that the m^7G is involved in the regulation of translation of its mRNA (Zhang et al. 2019). Furthermore, METTL1 is known to methylate miRNAs, such as let-7, and m^7G of let-7 miRNAs can be used to cleave pri-miRNAs and promote the formation of mature miRNA (Pandolfini et al. 2019).

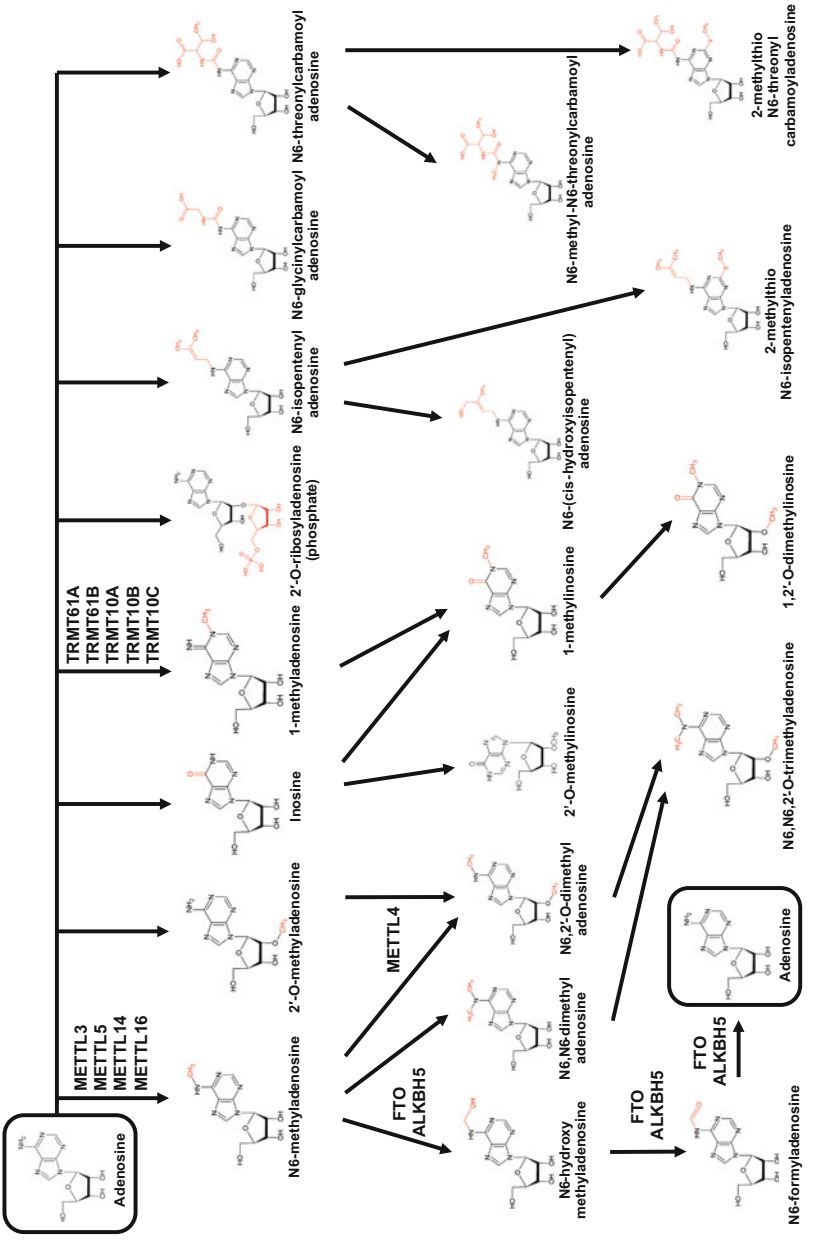


Fig. 1 Adenine modification in RNA. METTL3, 5, 14, and 16 methylate adenine to N6-methyl adenosine (m^6A). TRMT61A, 61B, 10A, 10B, and 10C involved in 1-methyl adenosine (1 mA). METTL4 involved in m^6Am. FTO and ALKBH5 were known about the demethylation enzymes

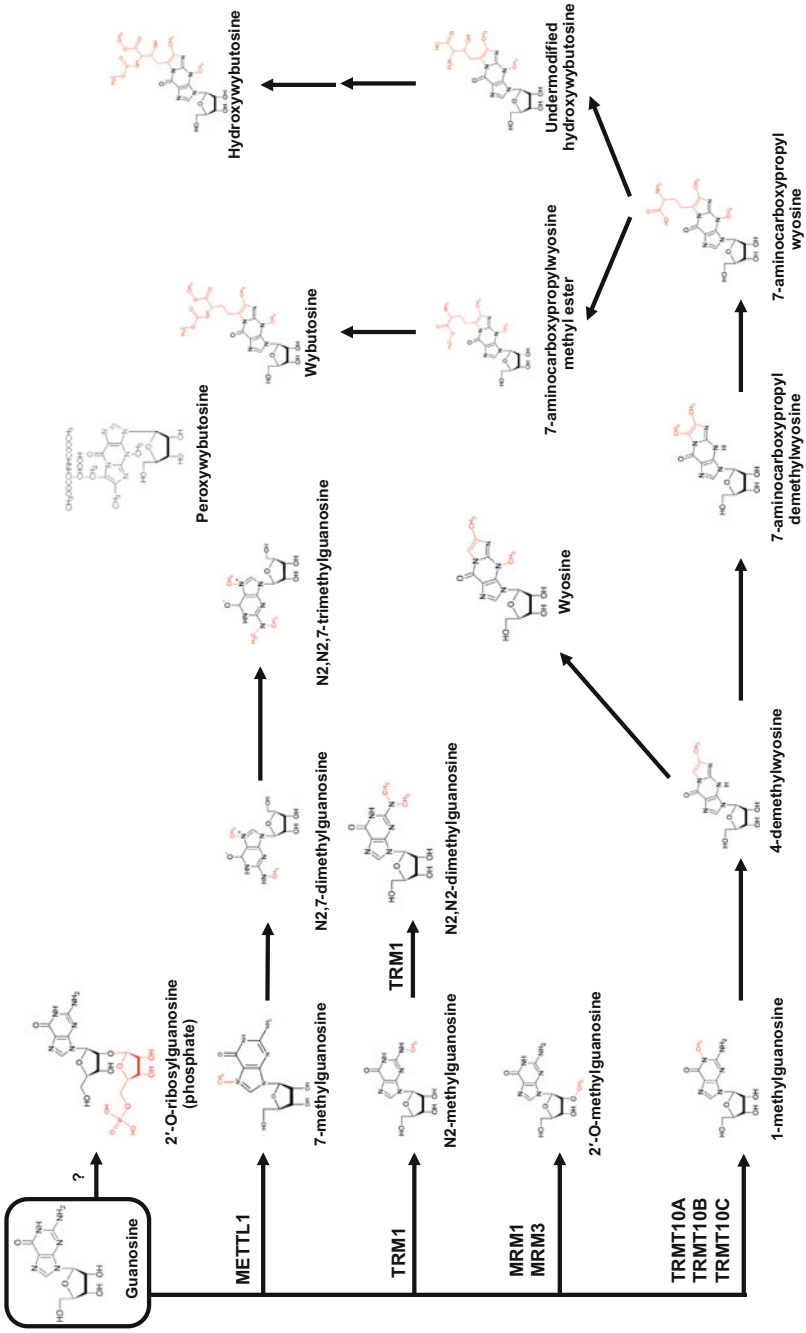


Fig. 2 Guanine modification in RNA. METTL7 methylate guanine to 7-methyl guanosine (7mG). TRM1 involved in N2, N2-dimethylguanine. MRM1, and 3 involved in 2′-*O*-methyl-guanine (Gm). TRMT10A, B, and C methylate guanine to 1-methyl guanosine (1mG)

3.2 METTL2A and 2B

METTL2A and METTL2B are S-adenosyl-L-methionine-dependent methyltransferase that mediate the 3-methylcytidine (3mC) modification of residue 32 of the tRNA anticodon loop of tRNA (Thr) (UGU) and tRNA (Arg) (CCU) (Fig. 3) (Xu et al. 2017a).

3.3 METTL3

METTL3 forms a heterodimer with METTL14 to mediate adenosine methylation at the N6 position of the RNA (Fig. 1). Furthermore, METTL3 constitutes the catalytic site in the heterodimer formed with METTL14 (Śledź and Jinek 2016; Wang et al. 2016a, b). m^6A in the 5′-[A/G]GAC-3′ consensus sequence of mRNA (m^6A) plays a role in mRNA stability, processing, translation efficiency, and editing (Dominissini et al. 2012; Wang et al. 2014; Liu et al. 2015; Alarcón et al. 2015a; Meyer et al. 2015). m^6A is a known regulator of mRNA stability, and m^6A of mRNAs is mainly mRNA is formed in the nucleus and promotes mRNA destabilization and degradation by recognizing the YTHDF family which m^6A reader proteins (Ke et al. 2017). Furthermore, m^6A modification of RNA has roles in the circadian clock, embryonic and hematopoietic stem cell differentiation, cortical neurogenesis, DNA damage response, T cell differentiation, and primary miRNA processing. In embryonic stem cells (ESCs), critical mRNAs for maintaining naive pluripotency have been found to be a key component of m^6A. m^6A is known to destabilization and degradation for transcripts and promote ESC differentiation (Batista et al. 2014; Geula et al. 2015; Bertero et al. 2018). Furthermore, m^6A has been shown to regulate the length of the circadian clock and has also been implicated in circadian regulation of the hepatic lipid metabolism (Zhong et al. 2018). m^6A also regulates sperm differentiation and meiosis and is known to be essential for male fertility and spermatogenesis (Chen et al. 2017; Xu et al. 2017b). Moreover, m^6A has been implicated in the response to DNA damage, and in response to UV irradiation, METTL3 is involved in the regulation of m^6A on the poly(A) of transcripts at the site of damaged DNA. METTL3 rapidly catalyzes the formation of m^6A and leads to the recruitment of DNA polymerase κ to sites of DNA damage (Xiang et al. 2017). m^6A is also required for T cell homeostasis and differentiation, and m^6A in suppressor of cytokine signaling (SOCS) family members (SOCS1, SOCS3, and CISH), which inhibit STAT signaling induce theses mRNA decay, and increase the levels of SOCS family proteins in naive T cells promotes mRNA destabilization and degradation and T cell differentiation (Li et al. 2017a, b, c). Moreover, increased SOCS family proteins inhibited IL-7-mediated STAT5 activation and T cell homeostatic proliferation and differentiation (Li et al. 2017a, b, c). m^6A of Xist leads to a target YTHDC1 leader on Xist, and Xist promotes the transcriptional repressive activity of m^6A (Patil et al. 2016). Xist m^6A methylation leads to the targeted YTHDC1 m^6A reader protein and

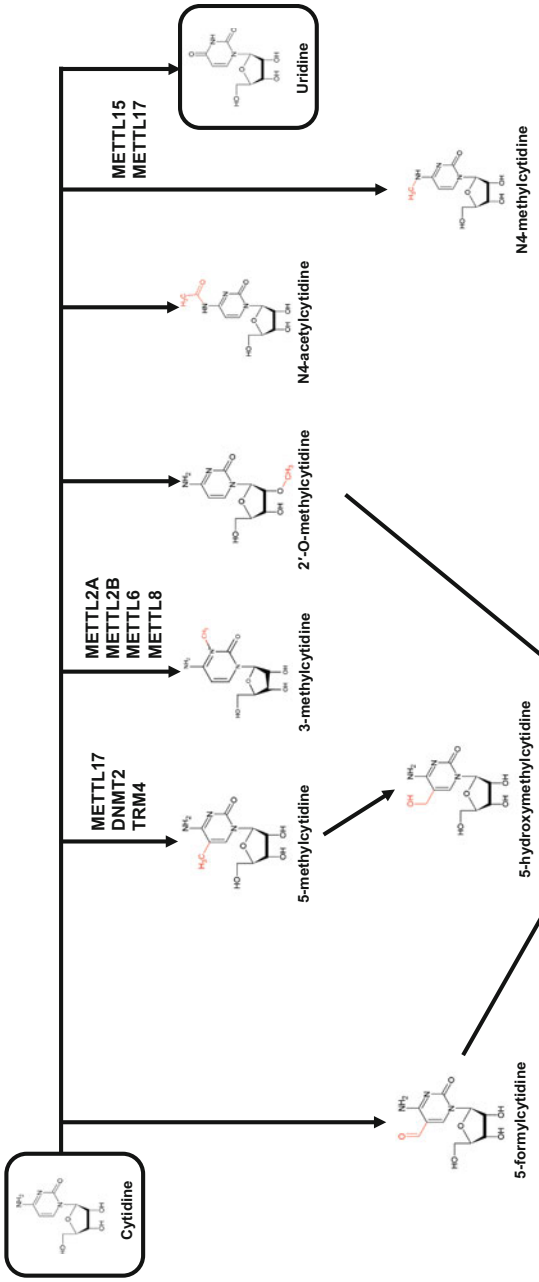


Fig. 3 Cytosine modification in RNA. METTL17, DNMT2, and TRM4 methylate cytosine to 5-methyl cytidine (5mC). METTL2A, 2B, 6, and 8 involved in 3-methyl cytidine (3mC). METTL15 and 17 involved in N^4-methyl cytidine (m^4C). DNMT2 also methylate cytosine in DNA to 5-methyl cytidine (5mC)

promotes resolved Xist mRNA (Patil et al. 2016). m^6A also regulates cortical neurogenesis, and during brain development, m^6A of transcription factors associated with neural stem cell maintenance, cell cycle, and neuronal differentiation promotes their destabilization and decay and promotes radial glial cell differentiation (Boles and Temple 2017; Wang et al. 2018). Furthermore, METTL3 promotes pri-miRNA methylation and promotes pri-miRNA recognition and processing by DiGeorge Syndrome Critical Region 8 (DGCR8) (Alarcón et al. 2015b). METTL3 is a positive regulator of mRNA translation independent of methyltransferase activity. Moreover, it is known to interact with translation initiation mechanisms in the cytoplasm to promote translation (Lin et al. 2016).

3.4 METTL4

METTL4 is an m^6A-specific DNA methyltransferase that mediates the methylation of DNA at position 6 of adenine the same in RNA and is known to be required to regulate polycomb silencing (Kweon et al. 2019). m^6A of DNA by METTL4 leads to the ubiquitination and degradation of additional sex combs like transcriptional regulator 1 (ASXL1) and MPN domain-containing protein (MPND), inactivation of the polycomb repressive-deubiquitinase (PR-DUB) complex, and subsequently leads to the preservation of its silencing (Kweon et al. 2019). Recently, METTL4 has also been reported to be involved in RNA methylation at adenine N6 site and $2'O$ in ribose, specifically in N6–$2'O$-methyladenosine (m^6Am) modification of snRNA, and is known to be important in the regulation of RNA splicing (Chen et al. 2020) (Fig. 1).

3.5 METTL5

METTL5 specifically methylates position 6 of adenine (m^6A) at position 1832 of the 18S rRNA. It is known to form a heterodimer with TRMT112 (van Tran et al. 2019) (Fig. 1).

3.6 METTL6

It is shown that METTL6 is an enzyme that modifies cytosine to 3-methylcytosine (3mC), which is the 32nd cytosine of the tRNA anticodon loop of tRNA (Ser) (Xu et al. 2017a) (Fig. 3).

3.7 METTL8

METTL8 is an *S*-adenosyl-L-methionine-dependent methyltransferase that mediates m^3C modification of mRNAs (Xu et al. 2017a) (Fig. 3).

3.8 METTL14

METTL14 is a non-catalytic subunit of m^6A methyltransferase that forms a heterodimer with METTL3 (Śledź and Jinek 2016; Wang et al. 2016a, b) (Fig. 1).

3.9 METTL15

METTL15 is an N^4-methylcytidine (m^4C) methyltransferase that is responsible for methylation at the C839 position of the mitochondrial 12S rRNA (Fig. 3). m^4C is involved in stabilizing the 12S rRNA folding of small mitochondrial ribosomal subunits (van Haute et al. 2019).

3.10 METTL16

METTL16 is an RNA N^6-methyltransferase that methylates adenosine residues at the N6 position of a subset of RNA and is involved in *S*-adenosyl-L-methionine homeostasis by regulating the expression of MAT2A transcripts (Fig. 1). METTL16 can N6-methylate mRNA and U6 snRNA (Pendleton et al. 2017). In contrast to the METTL3-METTL14 heterodimer, METTL16 requires both the 5′-UACAGAGAA-3′ sequence and a specific stem-loop RNA structure (Pendleton et al. 2017; Mendel et al. 2018; Doxtader et al. 2018). In the presence of *S*-adenosyl-L-methionine, METTL16 is known to bind to the 3′-untranslated region (UTR) region of mRNA of methionine adenosyl transferase 2A, or *S*-adenosylmethionine synthase isoform type-2 (MAT2A), which specifically methylate adenosine in the first hairpin structure of MAT2A mRNA, thus reducing MAT2A expression (Pendleton et al. 2017). In contrast, it is known that under *S*-adenosyl-L-methionine restriction conditions, the 3′-UTR of the MAT2A mRNA region is involved, but the lack of a methyl donor prevents its function and inhibits N^6-methylation, and promote the expression of MAT2A (Pendleton et al. 2017). In addition to mRNA, METTL16 can mediate N^6 methylation of U6 snRNA. Specifically, METTL16 N6-methylates the adenine at position 43 of U6 snRNA (Pendleton et al. 2017). METTL16 can also bind to a variety of lncRNAs, including 7SK RNA and 7SL RNA, and is known to specifically bind to the 3′ end of long non-coding RNAs (lncRNAs), such as Metastasis

Associated in Lung Adenocarcinoma Transcript 1 (MALAT1) (Warda et al. 2017; Brown et al. 2016).

3.11 METTL17

METTL17 is localized in the mitochondria, and METTL17 leads to m^4C840 and m^5C842 of 12S mt-rRNA (Shi et al. 2019) (Fig. 3). METTL17 leads to impaired translation of genes encoding mitochondrial proteins, resulting in mitochondrial oxidative phosphorylation of murine ESCs, which is important for pluripotency and major changes in the cellular metabolomics (Shi et al. 2019).

3.12 TRM1

tRNA methyltransferase (TRM) 1 is an enzyme that converts the guanine at position 26 of tRNA to N^2, N^2-dimethylguanine by using *S*-adenosyl-L-methionine as a methyl group donor (Ellis et al. 1986) (Fig. 2).

3.13 TRM4

TRM4 is an RNA methylating enzyme that methylates cytosines from various RNAs, including tRNAs, mRNAs, and some lncRNAs to 5-methylcytosines (5mC) (Flores et al. 2017) (Fig. 3). It is known that TRM4 conversion of RNA to 5mC is involved in the differentiation of stem cells and testis and the transfer from maternal to zygote by increasing protein synthesis (Flores et al. 2017; Llorens-Bobadilla et al. 2015; Signer et al. 2014; Baser et al. 2019). TRM4 methylates cytosine to 5-methylcytosine (5mC) at positions 34 and 48 of the tRNA (Leu) (CAA) precursor, and 48, 49, and 50 of the tRNA (Gly) (GCC) precursor (Auxilien et al. 2012; Gkatza et al. 2019). RNA methylation by TRM4 is required for the generation of tRNA (tRF)-derived RNA fragments (Auxilien et al. 2012; Gkatza et al. 2019). Moreover, TRM4 mediates 5mC conversion of mitochondrial tRNA (Haute et al. 2019), and catalyzes the cytosine methylation of mRNA, which stabilizes and prevents degradation (Zhang et al. 2012; Chen et al. 2019). Furthermore, it is known that 5mC modification of mRNA to cytosine is involved in the regulation of mRNA nuclear export by recognition of the 5mC modified mRNA by Aly/REF export factor (ALYREF, an mRNA transport adaptor, also named THOC4), which mediates nucleus-cytoplasmic shuttling (Yang et al. 2017).

3.14 TRMT61A and TRMT61B

TRMT 61A and B are enzymes that catalyze the formation of N^1-methyladenine (m^1A) for methionyl-tRNA at position 58 (m^1A58) (Ozanick et al. 2005) (Fig. 1).

TRMT61A and B are known to mediate the methylation of the adenosine at the N^1 position of mRNA in a region of mRNA that resembles the T-loop-like structure of tRNA (Li et al. 2017a, b, c; Safra et al. 2017).

3.15 TRMT10A and TRMT10B

TRMT10A and 10B are enzymes that catalyze the formation of guanine N^1-methylguanine (m^1G) in tRNA at position 9 (m^1G9) (Vilardo et al. 2012). However, it is believed that TRMT10A is not responsible for the formation of N^1-methyladenosine (m^1A) at position 9 (m^1A9) of the tRNA (Figs. 1 and 2) (Gillis et al. 2014).

3.16 TRMR10C

TRMT10C is an N^1-methyltransferase that is involved in the maturation of mitochondrial tRNA (Brzezniak et al. 2011). TRMT10C is a component of mitochondrial ribonuclease P, a complex that consists of TRMT10C/MRPP1, HSD17B10/MRPP2, and PRORP/MRPP3, and cleaves the 5′ end of the tRNA molecule (Holzmann et al. 2008). Together with HSD17B10/MRPP2, TRMT10C forms a mitochondrial ribonuclease P subcomplex named MRPP1-MRPP2, which has functions independent of ribonuclease P activity (Reinhard et al. 2017). The MRPP1-MRPP2 subcomplex catalyzes the formation of m^1G and m^1A at position 9 of the tRNA, while TRMT10C/MRPP1 acts as an N^1-methyltransferase subunit (Vilardo et al. 2012). In addition to its tRNA N^1-methyltransferase activity, TRMT10C/MRPP1 acts as an mRNA N^1-methyltransferase by mediating the methylation of adenosine residues at the N^1 position (m^1A) of mRNA of Mitochondrially Encoded NADH Dehydrogenase 5 (MT-ND5) (Safra et al. 2017) (Figs. 1 and 2).

3.17 MRM1

Mitochondrial RRNA Methyltransferase (MRM) 1 is an *S*-adenosyl-L-methionine-dependent 2′-*O*-ribose methyltransferase that catalyzes the formation of 2′-*O*-methylguanosine at position 1145 (Gm1145) of 16S mitochondrial large subunit ribosomal RNA (mtLSU rRNA) (Lee and Bogenhagen 2014). This represents a

universally conserved modification of the peptidyl transferase domain of 16S mtLSU rRNA (Lee and Bogenhagen 2014) (Fig. 2).

3.18 MRM3

MRM3 is a 2′-*O*-ribose methyltransferase that catalyzes the formation of 2′-*O*-methylguanosine at position 1370 (Gm1370) of 16S mtLSU rRNA (Lee and Bogenhagen 2014) (Fig. 2).

3.19 DNMT2

DNA Methyltransferase (DNMT) 2 specifically methylates DNA and cytosine 38 of the anticodon loop of tRNA (Asp) to form 5-methyl cytosine (5mC) (Goll et al. 2006) (Fig. 3).

4 Methods for Detecting on Modified RNAs

Given that RNA modifications are diverse with regards to the positions and splices of adducts, as well as in terms of length and structure of nucleotide chains, various challenging methods have to be applied to investigate and profile RNA modifications (Konno et al. 2019a; Li et al. 2020). Several methods for measuring RNA base modification have been reported. Indeed, dot blotting (Li et al. 2017a, b, c), Northern blotting with antibodies (Mishima et al. 2015), Maz RNA endonuclease (MazF is an *E. coli* toxin) (Imanishi et al. 2017), high-resolution melting (HRM) (Golovina et al. 2014), RNA photo crosslinking agents and quantitative proteomics (PCL-Proteomics) (Arguello et al. 2017), silver SiO_2-based electrochemical immunosensors (ECI) (Yin et al. 2017), and other methods have been reported for the purpose of measuring the amount of base modification. Moreover, a method to detect not only the amount of base modification but also the position of the modification has been reported (Molinie et al. 2016; Aschenbrenne et al. 2018; Linder et al. 2015; Meyer et al. 2012; Hafner et al. 2010; Liu et al. 2013; Konno et al. 2019b). A method for detecting the level of modification and isoform-characterization sequencing (LAIC-seq) (Molinie et al. 2016), direct sequencing (D-seq) (Aschenbrenne et al. 2018), base modification-specific immunoprecipitation (miCLIP) using the well-known antibodies (miCLIP) (Linder et al. 2015; Meyer et al. 2012). In addition, photoactivated ribonucleoside-enhanced cross-linked immunoprecipitation (PAR-CLIP) (Hafner et al. 2010), site-specific ligation-assisted extraction, and thin-layer chromatography after targeted cleavage and radiolabeling (SCARLET) (Liu et al. 2013), and capture mass spectrometry have also been

reported (Konno et al. 2019b). Using these methods, it is possible to identify RNA base modifications.

5 Diseases Involving the Epitranscriptomics

5.1 Mental Disorder

The brain contains the highest levels of m^6A expression of all major organs (Meyer et al. 2012). It is known that the function of m^6A contributes to the integration of embryonic brain development, movement, circadian clock regulation, dopaminergic midbrain circuits, and clue fear memory, etc. (Yoon et al. 2017; Wang et al. 2018; Lence et al. 2016; Fustin et al. 2013; Hess et al. 2013; Widagdo et al. 2016; Walters et al. 2017). The relationship between this RNA base modification and psychiatric disorders remains unclear, but it is currently under investigation.

5.2 Viral Infection

The epitranscriptome is related to several viral infections, such as loose sarcoma virus (Beemon and Keith 1977), flavivirus (Gokhale et al. 2016), Zika virus (Lichinchi et al. 2016a; Zhou et al. 2017), Kaposi's sarcoma-associated herpes virus (Ye et al. 2017; Hesser et al. 2018), human immunodeficiency virus (HIV) (Lichinchi et al. 2016b), influenza A virus (Courtney et al. 2017), tobacco mosaic virus (Li et al. 2018), SARS (Ma et al. 2015), and SARS-CoV-2 (Perveen et al. 2021). The epitranscriptomic conditions have also been reported to be involved in the replication of viral RNA in the nucleus of host cells (Beemon and Keith 1977; Gokhale et al. 2016; Lichinchi et al. 2016a, b; Zhou et al. 2017; Ye et al. 2017; Hesser et al. 2018; Courtney et al. 2017; Li et al. 2018). These viruses self-replicate in the host by utilizing the enhanced protein translation mechanism that is mediated by RNA base modification.

5.3 Cancer

1. Glioma
   In studies of glioma, knockdown of METTL3 or METTL14 has been shown to alter mRNA m^6A enrichment and alter mRNA expression of target genes, including ADAM19, which plays a key role in glioblastoma stem cells (Cui et al. 2017). In addition, it has been reported that inhibition of the m^6A-demethylase, FTO, suppresses tumor progression in glioblastoma stem cell-transplanted mice and prolongs the survival rate (Li et al. 2017a, b, c). This

study demonstrated that m^6A plays an important role in glioblastoma stem cell self-renewal and tumorigenesis and suggested that m^6A-modification is a promising therapeutic target for glioblastoma (Cui et al. 2017).

2. Leukemia

   FTO is up-regulated in acute myeloid leukemia (AML), such as mixed lineage leukemia (MLL) with t(11q23) rearrangements, promyelocytic leukemia (PML) with t(15;17), involving Retinoic Acid Receptor-a (RARA), and others with Internal Tandem Duplication (ITD) mutations in FLT3 (FMS-like tyrosine kinase 3) gene and nucleophosmin 1 (NPM1) mutations, suggesting that FTO plays an important carcinogenic role in AML (Li et al. 2017a, b, c). In a previous study, FTO reduced the expression of the ankyrin repeat and SOCS box containing 2 (ASB2) and RARA by the demethylation from m^6A modification of mRNA, such as ASB2 and RARA, and inhibited AML cell differentiation against therapeutic approach by all-trans-retinoic acid (ATRA), which subsequently led to leukemic transformation. It is generally considered that these findings led to the elucidation of the disease mechanism and development of therapeutic drugs (Li et al. 2017a, b, c).

3. Lung Cancer

   METTL3 increases the level of m^6A and interacts with the translation initiation mechanism to promote mRNA translation, and activates hippo pathway effector transcription with epidermal growth factor receptor and PDZ-binding motifs in human cancer cells. Furthermore, METTL3 has been shown to promote the translation of certain mRNAs, including the Hippo pathway effector, transcriptional co-activator with PDZ-binding motif (TAZ) (Lin et al. 2016). This study suggested that METTL3 can promote the proliferation, survival, and invasion of human lung cancer cells; thus, METTL3 may represent a good therapeutic target (Lin et al. 2016).

4. Breast Cancer

   In a study of human breast cancer cells, exposure to hypoxia was shown to stimulate the expression of the m^6A demethylase AlkBH5, which reduced NANOG mRNA methylation levels in the 3′-UTR sequence AAACU, increased NANOG protein levels, and increased the stem cell population (Zhang et al. 2016a). Furthermore, it was showed that the exposure to hypoxia stimulated NANOG expression in a hypoxia-inducible factor (HIF)-1α- and HIF-2-α-dependent manner (Zhang et al. 2016a). Another study demonstrated that hypoxic exposure induced zinc finger protein (ZNF) 217-dependent inhibition of m^6A modification of mRNAs encoding NANOG and Kruppel like factor (KLF) 4 in breast cancer (Zhang et al. 2016b). ZNF217 or ALKBH5 are thought to be involved in the regulation of pluripotency factor expression in breast cancer under hypoxic conditions (Zhang et al. 2016b).

5. Colorectal Cancer

   Recent studies have suggested that m^6A modification is involved in the malignant behavior of colorectal cancer (Nishizawa et al. 2017). Epigenetic data obtained by chromatin immunoprecipitation suggest that the oncogene

c-myc is involved in the transcription of the M6A leader YTHDF1 in colorectal cancer (Nishizawa et al. 2017). Immunohistochemical analysis of YTHDF1 showed that YTHDF1 expression was associated with the behavior of various malignancies, and YTHDF1 expression was identified as an independent prognostic factor in colorectal cancer patients (Nishizawa et al. 2017).

6. Pancreatic Cancer

   Previous case-control studies of genetic mutations have shown an association between FTO mutations and pancreatic cancer risk in Japan (Lin et al. 2013). Furthermore, a significant association between FTO rs9939609 mutation polymorphisms and endometrial cancer and pancreatic cancer has been demonstrated, especially in Asian populations, and may be a potential biomarker for early diagnosis (Huang et al. 2017). An association between the FTO rs9939609 variant and the risk of malignant pleural mesothelioma has also been reported (Khella et al. 2018). Another study of pancreatic cancer showed that METTL3 deficient cells were more sensitive to anticancer drugs, such as gemcitabine, 5-fluorouracil, cisplatin, and radiation. METTL3 can be used as a treatment for patients with pancreatic cancer, and it has been suggested to be a powerful target with increased efficacy than the other therapeutic agents (Taketo et al. 2018). Furthermore, recent reports have shown that the m^6A level of miRNA contained in serum can detect Stage I and II pancreatic cancer with very good sensitivity and specificity (Konno et al. 2019b); therefore, the m^6A level may represent an unprecedented but useful biomarker.

6 Conclusion

Studies on the epitranscriptomics are just beginning, and not all RNA base-modifying enzymes are fully understood. Furthermore, much remains to be known about the relationship between diseases and the epitranscriptomics. Further research into the epitranscriptomic information will lead to a greater understanding of the detailed mechanisms of its role in diseases, as well as the determination of potential therapeutic targets.

Acknowledgments We wish to thank the members of our laboratories. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology (17H04282, 17K19698, 18K16356, and 18K16355); and Japan Agency for Medical Research and Development (AMED) (16cm0106414h0001 and 17cm0106414h0002). Partial support was received from the Princess Takamatsu Cancer Research Fund. Partial institutional endowments were received from Hirotsu Bio Science Inc. (Tokyo, Japan); Kinshu-kai Medical Corporation (Osaka, Japan); Kyowa-kai Medical Corporation (Osaka, Japan); IDEA Consultants Inc. (Tokyo, Japan); and Unitech Co. Ltd. (Chiba, Japan).

References

Alarcón RC, Goodarzi H, Lee H et al (2015a) HNRNPA2B1 is a mediator of m(6)A-dependent nuclear RNA processing events. Cell 162:1299–1308

Alarcón RC, Lee H, Goodarzi H et al (2015b) N6-methyladenosine marks primary microRNAs for processing. Nature 519:482–485

Alexandrov A, Martzen RM, Phizicky ME (2002) Two proteins that form a complex are required for 7-methylguanosine modification of yeast tRNA. RNA 8:1253–1266

Arguello AE, DeLiberto AN, Kleiner RE (2017) RNA chemical proteomics reveals the N6-methyladenosine (m6A)-regulated protein-RNA interactome. J Am Chem Soc 139:17249–17252

Aschenbrenne J, Werner S, Marchand V et al (2018) Engineering of a DNA polymerase for direct m6 A sequencing. Angew Chem Int Ed Engl 57:417–421

Auxilien S, Guérineau V, Szweykowska-Kulińska Z et al (2012) The human tRNA m(5)C methyltransferase misu is multisite-specific. RNA Biol 9:1331–1338

Baser A, Skabkin M, Kleber S, Dang Y et al (2019) Onset of differentiation is post-transcriptionally controlled in adult neural stem cells. Nature 566:100–104

Batista PJ, Molinie B, Wang J et al (2014) m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell 15:707–719

Beemon K, Keith J (1977) Localization of N6-methyladenosine in the Rous sarcoma virus genome. J Mol Biol 113:165–179

Bertero A, Brown S, Madrigal P et al (2018) The SMAD2/3 interactome reveals that TGFβ controls m6A mRNA methylation in pluripotency. Nature 555:256–259

Boles NC, Temple S (2017) Epimetronomics: m6A marks the tempo of corticogenesis. Neuron 96:718–720

Brown AJ, Kinzig GC, DeGregorio JS et al (2016) Methyltransferase-like protein 16 binds the 3′-terminal triple helix of MALAT1 long noncoding RNA. Proc Natl Acad Sci USA 113:14013–14018

Brzezniak KL, Bijata M, Szczesny JR et al (2011) Involvement of human ELAC2 gene product in 3′ end processing of mitochondrial tRNAs. RNA Biol 8:616–626

Chen X, Li X, Guo J et al (2017) The roles of microRNAs in regulation of mammalian spermatogenesis. J Anim Sci Biotechnol 8:35

Chen X, Li A, Sun BF et al (2019) 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. Nat Cell Biol 21:978–990

Chen H, Gu L, Orellana EA et al (2020) METTL4 is an snRNA m6Am methyltransferase that regulates RNA splicing. Cell Res 30:544–547

Coker H, Wei G, Brockdorff N (2019) m6A modification of non-coding RNA and the control of mammalian gene expression. Biochim Biophys Acta Gene Regul Mech 1862:310–318

Courtney DG, Kennedy EM, Dumm RE et al (2017) Epitranscriptomic enhancement of influenza a virus gene expression and replication. Cell Host Microbe 22:377–386

Cui Q, Shi H, Ye P et al (2017) m6A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. Cell Rep 18:2622–2634

Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485:201–206

Doxtader AK, Wang P, Scarborough MA et al (2018) Structural basis for regulation of METTL16, an S-adenosylmethionine homeostasis factor. Mol Cell 71:1001–1011

Ellis SR, Morales MJ, Li JM et al (1986) Isolation and characterization of the TRM1 locus, a gene essential for the N2,N2-dimethylguanosine modification of both mitochondrial and cytoplasmic tRNA in *Saccharomyces cerevisiae*. J Biol Chem 261:9703–9709

Flores JV, Cordero-Espinoza L, Oeztuerk-Winder F et al (2017) Cytosine-5 RNA methylation regulates neural stem cell differentiation and motility. Stem Cell Rep 8:112–124

Frye M, Harada BT, Behm M et al (2018) RNA modifications modulate gene expression during development. Science 361:1346–1349

Fustin JM, Doi M, Yamaguchi Y et al (2013) RNA-methylation-dependent RNA processing controls the speed of the circadian clock. Cell 155:793–806
Geula S, Moshitch-Moshkovitz S, Dominissini D et al (2015) Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science 347:1002–1006
Gillis D, Krishnamohan A, Yaacov B et al (2014) TRMT10A dysfunction is associated with abnormalities in glucose homeostasis, short stature and microcephaly. J Med Genet 51:581–586
Gkatza AN, Castro C, Harvey FR et al (2019) Cytosine-5 RNA methylation links protein synthesis to cell metabolism. PLoS Biol 17:e3000297
Gokhale NS, McIntyre ABR, McFadden MJ et al (2016) N6-Methyladenosine in Flaviviridae viral RNA genomes regulates infection. Cell Host Microbe 20:654–665
Goll MG, Kirpekar F, Maggert KA et al (2006) Methylation of tRNA Asp by the DNA methyltransferase homolog Dnmt2. Science 311:395–398
Golovina AY, Dzama MM, Petriukov KS et al (2014) Method for site-specific detection of m6A nucleoside presence in RNA based on high-resolution melting (HRM) analysis. Nucleic Acids Res 42:e27
Hafner M, Landthaler M, Burger L et al (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141:129–141
Haute VL, Lee SY, McCann JB et al (2019) NSUN2 introduces 5-methylcytosines in mammalian mitochondrial tRNAs. Nucleic Acids Res 47:8720–8733
Hess ME, Hess S, Meyer KD et al (2013) The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. Nat Neurosci 16:1042–1048
Hesser CR, Karijolich J, Dominissini D et al (2018) N6-methyladenosine modification and the YTHDF2 reader protein play cell type specific roles in lytic viral gene expression during Kaposi's sarcoma-associated herpesvirus infection. PLoS Pathog 14:e1006995
Holzmann J, Frank P, Löffler E et al (2008) RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell 135:462–474
Huang X, Zhao J, Yang M et al (2017) Association between FTO gene polymorphism (rs9939609 T/A) and cancer risk: a meta-analysis. Eur J Cancer Care 26
Imanishi M, Tsuji S, Suda A et al (2017) Detection of N6-methyladenosine based on the methyl-sensitivity of MazF RNA endonuclease. Chem Commun 53:12930–12933
Jia G, Fu Y, Zhao X et al (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887
Ke S, Pandya-Jones A, Saito Y et al (2017) M 6 A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. Genes Dev 31:990–1006
Khella MS, Salem AM, Abdel-Rahman O et al (2018) The association between the FTO rs9939609 variant and malignant pleural mesothelioma risk: a case-control study. Genet Test Mol Biomarkers 22:79–84
Konno M, Taniguchi M, Ishii H (2019a) Significant epitranscriptomes in heterogeneous cancer. Cancer Sci 110:2318–2327
Konno M, Koseki J, Asai A et al (2019b) Distinct methylation levels of mature microRNAs in gastrointestinal cancers. Nat Commun 10:3888
Kweon SM, Chen Y, Moon E et al (2019) An adversarial DNA N(6)-methyladenine-sensor network preserves polycomb silencing. Mol Cell 74:1138–1147
Lee KW, Bogenhagen DF (2014) Assignment of 2′-O-methyltransferases to modification sites on the mammalian mitochondrial large subunit 16S rRNA. J Biol Chem 289:24936–24942
Lence T, Akhtar J, Bayer M et al (2016) m6A modulates neuronal functions and sex determination in Drosophila. Nature 540:242–247
Li HB, Tong J, Zhu S et al (2017a) M^6A mRNA methylation controls T cell homeostasis by targeting the IL-7/STAT5/SOCS pathways. Nature 548:338–342
Li Z, Weng H, Su R et al (2017b) FTO plays an oncogenic role in acute myeloid leukemia as a N6-methyladenosine RNA demethylase. Cancer Cell 31:127–141

Li X, Xiong X, Zhang M et al (2017c) Base-resolution mapping reveals distinct M^1A methylome in nuclear- and mitochondrial-encoded transcripts. Mol Cell 68:993–1005
Li Z, Shi J, Yu L et al (2018) N6-methyl-adenosine level in *Nicotiana tabacum* is associated with tobacco mosaic virus. Virol J 15:87
Li X, Liang QX, Lin JR et al (2020) Epitranscriptomic technologies and analyses. Sci China Life Sci 63:501–515
Lichinchi G, Zhao BS, Wu Y et al (2016a) Dynamics of human and viral RNA methylation during Zika virus infection. Cell Host Microbe 20:666–673
Lichinchi G, Gao S, Saletore Y et al (2016b) Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. Nat Microbiol 1:16011
Lin Y, Ueda J, Yagyu K et al (2013) Association between variations in the fat mass and obesity-associated gene and pancreatic cancer risk: a case-control study in Japan. BMC Cancer 13:337
Lin S, Choe J, Du P et al (2016) The m(6)A methyltransferase METTL3 promotes translation in human cancer cells. Mol Cell 62:335–345
Linder B, Grozhik AV, Olarerin-George AO et al (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods 12:767–772
Liu N, Parisien M, Dai Q et al (2013) Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA 19:1848–1856
Liu J, Yue Y, Han D et al (2014) A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol 10:93–95
Liu N, Dai Q, Zheng G et al (2015) N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature 518:560–564
Llorens-Bobadilla E, Zhao S, Baser A et al (2015) Single-cell transcriptomics reveals a population of dormant neural stem cells that become activated upon brain injury. Cell Stem Cell 17:329–340
Ma Y, Wu L, Shaw N et al (2015) Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc Natl Acad Sci USA 112:9436–9441
Mendel M, Chen KM, Homolka D et al (2018) Methylation of structured RNA by the m6A writer METTL16 is essential for mouse embryonic development. Mol Cell 71:986–1000
Meyer KD, Jaffrey SR (2014) The dynamic epitranscriptome: N6-methyladenosine and gene expression control. Nat Rev Mol Cell Biol 15:313–326
Meyer KD, Jaffrey SR (2017) Rethinking m6A readers, writers, and erasers. Annu Rev Cell Dev Biol 33:319–342
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149:1635–1646
Meyer KD, Patil PD, Zhou J et al (2015) 5' UTR m(6)A promotes cap-independent translation. Cell 163:999–1010
Mishima E, Jinno D, Akiyama Y et al (2015) Immuno-northern blotting: detection of RNA modifications by using antibodies against modified nucleosides. PLoS One 10:e0143756
Molinie B, Wang J, Lim KS et al (2016) m(6)A-LAIC-seq reveals the census and complexity of the m(6)A epitranscriptome. Nat Methods 13:692–698
Nishizawa Y, Konno M, Asai A et al (2017) Oncogene c-Myc promotes epitranscriptome m6A reader YTHDF1 expression in colorectal cancer. Oncotarget 9:7476–7486
Ontiveros RJ, Stoute J, Liu KF (2019) The chemical diversity of RNA modifications. Biochem J 476:1227–1245
Ozanick S, Krecic A, Andersland J et al (2005) The bipartite structure of the tRNA m1A58 methyltransferase from *S*. cerevisiae is conserved in humans. RNA 11:1281–1290
Pandolfini L, Barbieri I, Bannister JA et al (2019) METTL1 promotes let-7 microRNA processing via m7G methylation. Mol Cell 74:1278–1290
Patil PD, Chen CK, Pickering FB et al (2016) m(6)A RNA methylation promotes XIST-mediated transcriptional repression. Nature 537:369–373
Pendleton EK, Chen B, Liu K et al (2017) The U6 snRNA m6A methyltransferase METTL16 regulates SAM synthetase intron retention. Cell 169:824–835

Perveen S, Yazdi KA, Devkota K et al (2021) A high-throughput RNA displacement assay for screening SARS-CoV-2 nsp10-nsp16 complex toward developing therapeutics for COVID-19. SLAS Discov 10:2472555220985040
Reinhard L, Sridhara S, Hällberg BM (2017) The MRPP1/MRPP2 complex is a tRNA-maturation platform in human mitochondria. Nucleic Acids Res 45:12469–12480
Safra M, Sas-Chen A, Nir R et al (2017) The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature 551:251–255
Shi Z, Xu S, Xing S et al (2019) Mettl17, a regulator of mitochondrial ribosomal RNA modifications, is required for the translation of mitochondrial coding genes. FASEB J 33:13040–13050
Signer RA, Magee JA, Salic A et al (2014) Haematopoietic stem cells require a highly regulated protein synthesis rate. Nature 509:49–54
Śledź P, Jinek M (2016) Structural insights into the molecular mechanism of the m(6)a writer complex. elife 5:e18434
Taketo K, Konno M, Asai A et al (2018) The epitranscriptome m6A writer METTL3 promotes chemo- and radioresistance in pancreatic cancer cells. Int J Oncol 52:621–629
van Haute L, Hendrick AG, D'Souza AR et al (2019) METTL15 introduces N4-methylcytidine into human mitochondrial 12S rRNA and is required for mitoribosome biogenesis. Nucleic Acids Res 47:10267–10281
van Tran N, Ernst FGM, Hawley BR et al (2019) The human 18S rRNA m6A methyltransferase METTL5 is stabilized by TRMT112. Nucleic Acids Res 47:7719–7733
Vilardo E, Nachbagauer C, Buzet A et al (2012) A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase—extensive moonlighting in mitochondrial tRNA biogenesis. Nucleic Acids Res 40:11583–11593
Walters BJ, Mercaldo V, Gillon CJ et al (2017) The role of the RNA demethylase FTO (fat mass and obesity-associated) and mRNA methylation in hippocampal memory formation. Neuropsychopharmacology 42:1502–1510
Wang X, He C (2014) Dynamic RNA modifications in posttranscriptional regulation. Mol Cell 56:5–12
Wang X, Lu Z, Gomez A et al (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505:117–120
Wang P, Doxtader AK, Nam Y (2016a) Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. Mol Cell 63:306–317
Wang X, Feng J, Xue Y et al (2016b) Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. Nature 534:575–578
Wang Y, Li Y, Yue M et al (2018) N(6)-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications. Nat Neurosci 21:195–206
Warda SA, Kretschmer J, Hackert P et al (2017) Human METTL16 is a N6-methyladenosine (m6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep 18:2004–2014
Widagdo J, Zhao QY, Kempen MJ et al (2016) Experience-dependent accumulation of N6-methyladenosine in the prefrontal cortex is associated with memory processes in mice. J Neurosc 36:6771–6777
Xiang Y, Laurent B, Hsu CH et al (2017) RNA m6A methylation regulates the ultraviolet-induced DNA damage response. Nature 543:573–576
Xu L, Liu X, Sheng N et al (2017a) Three distinct 3-methylcytidine (m3C) methyltransferases modify tRNA and mRNA in mice and humans. J Biol Chem 292:14695–14703
Xu K, Yang Y, Feng GH et al (2017b) Mettl3-mediated m(6)A regulates spermatogonial differentiation and meiosis initiation. Cell Res 27:1100–1114
Yang X, Yang Y, Sun BF et al (2017) 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an M5C reader. Cell Res 27:606–625
Ye F, Chen ER, Nilsen TW (2017) Kaposi's sarcoma-associated herpesvirus utilizes and manipulates RNA N6-adenosine methylation to promote lytic replication. J Virol 91:e00466–e00417

Yin H, Wang H, Jiang W et al (2017) Electrochemical immunosensor for N6-methyladenosine detection in human cell lines based on biotin-streptavidin system and silver-SiO2 signal amplification. Biosens Bioelectron 90:494–500

Yoon KJ, Ringeling FR, Vissers C et al (2017) Temporal control of mammalian cortical neurogenesis by m6A methylation. Cell 171:877–889

Zhang X, Liu Z, Yi J et al (2012) The tRNA methyltransferase NSun2 stabilizes p16INK^4 mRNA by methylating the 3′-untranslated region of p16. Nat Commun 3:712

Zhang C, Samanta D, Lu H et al (2016a) Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m^{6}A-demethylation of NANOG mRNA. Proc Natl Acad Sci USA 113:E2047–E2056

Zhang C, Zhi WI, Lu H et al (2016b) Hypoxia-inducible factors regulate pluripotency factor expression by ZNF217- and ALKBH5-mediated modulation of RNA methylation in breast cancer cells. Oncotarget 7:64527–64542

Zhang L, Liu C, Ma H et al (2019) Transcriptome-wide mapping of internal N ^{7}-methylguanosine methylome in mammalian mRNA. Mol Cell 74:1304–1316

Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29

Zhong X, Yu J, Frazier K et al (2018) Circadian clock regulation of hepatic lipid metabolism by modulation of m6A mRNA methylation. Cell Rep 25:1816–1828

Zhou H, Wang F, Wang H et al (2017) The conformational changes of Zika virus methyltransferase upon converting SAM to SAH. Oncotarget 8:14830–14834

Epitranscriptomics Markers Regulate the Infection by RNA Viruses

Vassiliki Stamatopoulou and Apostolos Zaravinos

Contents

Abstract Epitranscriptomics is a relatively new field encompassing different kinds of post-transcriptional RNA modifications, such as nucleotide insertions, deletions, or substitutions within the nascent RNA. These post-transcriptional modifications play a key role in the regulation of gene expression and allow cells to respond to different environmental stimuli. In this review article, we discuss the major types of RNA modifications that take place in the genome of RNA viruses. We focus on Uridine-to-Pseudouridine editing, Adenosine-to-Inosine editing, Cytidine-to-Uridine editing, N^6-methyladenosine, N^6, 2-*O*-dimethyladenosine, 5-Methylcytosine, and 2′-*O*-methylation of ribose in viral RNAs. We summarize the most important and recent findings in the literature, and we discuss how these editing events regulate the infection of RNA viruses via altering gene expression. Taken together, this review highlights the importance of the different types of epitranscriptomic modifications in the regulation of infection by RNA viruses.

V. Stamatopoulou
Department of Biochemistry, School of Medicine, University of Patras, Patras, Greece

A. Zaravinos (✉)
Department of Life Sciences, School of Sciences, European University Cyprus, Nicosia, Cyprus
e-mail: a.zaravinos@euc.ac.cy

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_5

Keywords Epitranscriptomics · RNA editing · RNA viruses · Viral infection · Pseudouridine · Inosine · N^6-methyladenosine · 5-methylcytosine · 2′-*O*-methylation of ribose

1 Introduction

Epitranscriptomics is a relatively new field that includes the study of all the biochemical modifications of the RNA (the transcriptome) within a cell. In analogy to epigenetics that describes "functionally relevant changes to the genome that do not involve a change in the nucleotide sequence," epitranscriptomics involves all functionally relevant changes to the transcriptome that do not involve a change in the ribonucleotide sequence. Thus, the epitranscriptome can be defined as the ensemble of such functionally relevant changes or RNA edits. There are different types of RNA editing events that can impact the expression of a gene. These edits often happen to different RNA molecules, including ribosomal RNA (rRNA), transfer RNA (tRNA), messenger RNA (mRNA), and small nuclear RNA (snRNA), and can be grouped as follows: (1) Uridine-to-Pseudouridine editing, (2) Adenosine-to-Inosine (A-to-I) editing, (3) Cytidine-to-Uridine (C-to-U) editing, (4) methylation of the ribose 2′-hydroxyl (–OH) group and (5) other complex (hyper)modifications.

The main editors in A-to-I editing are adenosine deaminases acting on RNA (ADARs) and adenosine deaminases that act on tRNAs (ADATs), both of which regulate alternative splicing and transcriptional control. On the other hand, C-to-U editors of the AID/APOBEC family are mainly involved in innate and adaptive immunity (reviewed in Christofi and Zaravinos 2019).

2 RNA Editing in Viruses

RNA editing in (mainly mammalian) viruses, i.e., RNA modifications in viral transcripts, are the focus point of viral epitranscriptomics. In general, such modifications do not affect the viral sequence of the transcript, but they are functionally relevant, enhancing stability and generating new protein variants. The main RNA editing events in viruses include the isomerization of Uridine-to-5-ribosyl uracil (or pseudouridine, Ψ) and the deamination of A-to-I, among others.

2.1 Uridine-to-Pseudouridine (Ψ) Editing

Pseudouridine (Ψ) is the most abundant post-transcriptional edited nucleotide and thus, considered as the "fifth nucleotide." It consists of a uridine isomer, where the U

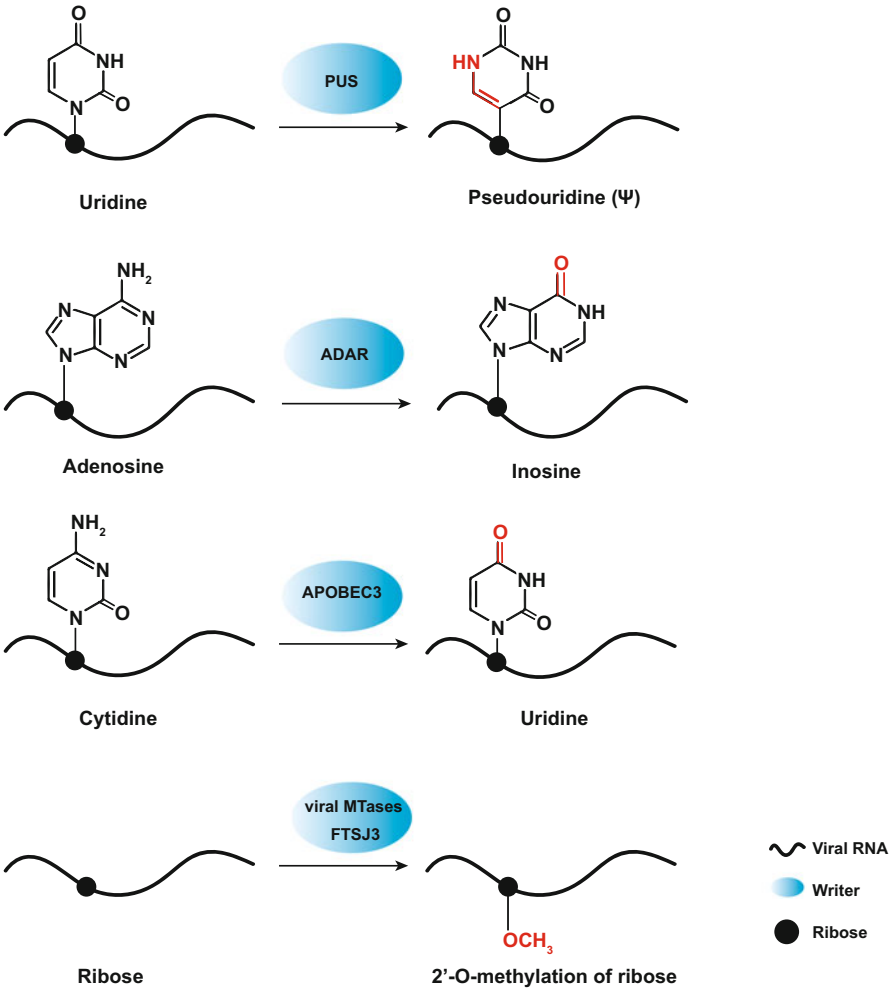


Fig. 1 Editing nucleotides within viral RNA. Pseudouridine, Inosine, Uridine, and 2′-*O* methylation of ribose are irreversible RNA modifications. Modification marks are indicated with red letters. The enzymes responsible for each modification are also indicated, where PUS correspond to Pseudouridine synthase, ADAR to adenosine deaminases acting on RNA, APOBEC3 to apolipoprotein-B mRNA editing enzyme, catalytic polypeptide-3, and MTase to Methyltransferase

base is attached via a C–C instead of an N–C glycosidic bond (Fig. 1). Pseudouridine synthases (PUS) are responsible for the isomerization of uridines to Ψ's (Netzband and Pager 2020). Although this isomerization leaves base pairing with adenosine intact, Ψ base pairs with A, G, C, and U. One technique to distinguish Ψ from uridine is by tagging it with the CMC cation *N*-cyclohexyl-*N*′-β-(4-methylmorpholinium) ethylcarbodiimide label (Li et al. 2015; Carlile et al. 2014; Lovejoy et al. 2014). Ψ-seq is a new single-nucleotide-resolution method that has allowed researchers to

identify pseudouridines genome-wide (Carlile et al. 2014; Schwartz et al. 2014). The method is based on the selective modification of Ψ's with *N*-cyclohexyl-*N*-′-(2-morpholinoethyl)carbodiimide metho-p-toluenesulfonate (CMC) to generate a block to reverse transcriptase (RT) one nucleotide 3′ to the pseudouridylated site. Mock CMC-treated RNA fragments are processed in parallel to identify Ψ-independent RT stops. This chemistry can be exploited to determine the location of Ψ's using next-generation seq (Carlile et al. 2014).

Pseudouridine is one among many post-transcriptional modifications taking place in viral RNAs (McIntyre et al. 2018). Having been poorly identified in viral RNAs, how it precisely regulates viral gene expression is widely unknown. Nevertheless, pseudouridylation has been detected in quite a few RNA viruses, including turnip yellow mosaic virus (TYMV) and brome mosaic virus (BMV) (Cuesta et al. 2019; Baumstark and Ahlquist 2001). TYMV belongs to the Tymoviridae family and mainly infects plants in the genus Brassica. The last 82 nucleotides in the virus's RNA genome form a "tRNA-like" domain, within which a stem-loop structure resembles the pseudouridine-containing T-arm in the tRNAs. The incubation of this virus's RNA with yeast extract or purified PUS1 and PUS4 synthases led to the pseudouridylation of this RNA fragment (Netzband and Pager 2020). BMV is another plant RNA virus of the Bromoviridae family, with an intergenic region in its RNA3 segment that also mimics the tRNA TΨC-stem loop in tRNA (Netzband and Pager 2020).

Interestingly, the amount of Ψ is regulated in response to environmental signals, including stress (Li et al. 2015; Carlile et al. 2014), implying that pseudouridylation is inducible. Indeed, viral infection can induce a stress response, as observed by an oscillatory pattern in the abundance of Ψ in RNA virus-infected cells (McIntyre et al. 2018).

Being broadly distributed within the mRNA, from the two untranslated regions (UTRs) to the coding part of the genome (Carlile et al. 2014), it is questionable how Ψ's affect viral gene expression. Pseudouridines use a water-ΨNH1 hydrogen bond to stabilize their conformation. Water-mediated hydrogen bonds were observed between the NH1 proton and phosphate oxygen atoms in crystal structures of tRNA molecules (Rodriguez-Hernandez et al. 2013) and molecular dynamics simulations. Such hydrogen bonds have been postulated to stabilize Ψ-containing helices in solution (Charette and Gray 2000; Hudson et al. 2013); however, no direct evidence for their existence has been obtained. This indicates that as in tRNAs, Ψ's within viral RNAs seem to preserve for RNA structure controlling the viral proteome via frameshifts, misreading, or suppression of stop and nonsense codons (Table 1 in chapter "RNA Modifications in Neurodegenerations"). The integration of Ψs into mRNA yields immune suppression, mRNA stabilization (Anderson et al. 2011) and enhances the translation of mRNAs containing pseudouridine, compared to those containing uridine, by diminishing the activation of protein kinase (PKR) (Anderson et al. 2010), among other functions, all of which impact on viral gene expression (Netzband and Pager 2020).

2.2 *Adenosine-to-Inosine (A-to-I) Editing*

A-to-I editing is the most common RNA modification (Pfaller et al. 2018; Danecek et al. 2012). It is catalyzed by ADARs, through deamination at the C6 position of A's producing I's within double-stranded (ds) RNA regions (George et al. 2014; Slotkin and Nishikura 2013; Samuel 2011) (Fig. 1). Aberrant ADAR activity has been linked to various diseases, such as different cancer types, neurological or metabolic diseases (Christofi and Zaravinos 2019), but also with viral infections and autoimmune disorders (Slotkin and Nishikura 2013). When an adenosine nucleotide is converted to inosine, it acts in a manner similar to a guanosine (G) nucleotide, with different consequences. When A-to-I editing occurs in the coding region of the mRNA, it results in an altered nucleotide codon and, therefore, can change the amino acid sequence of the coded protein in what is referred to as a re-coding editing event. A-to-I editing can also create or eliminate splice sites, thus, altering the portions of the RNA that remain in the final product. In addition, A-to-I edits alter base pairing, because I pairs preferentially with C, and this can potentially affect the secondary structure of the RNA. In the case of RNA molecules that bind target RNA segments, such as microRNAs (miRNAs), the altered base pairing can change binding specificities. It is also hypothesized that A-to-I edits could affect ribosomal coding or the function of viral RNA-dependent RNA polymerases, and hence, viral infection.

In mammals, there are three ADARs (ADAR1–3). ADAR1 acts as an interferon (IFN)-inducible protein, and ADAR2 is constitutively expressed. In contrast to the first two, ADAR3 does not exhibit any A-to-I activity. In addition, two ADAR-1 isoforms exist, p110 being constitutively expressed in the nucleus and p150 being expressed both in the cytoplasm and nucleus, and being induced by IFN signaling (Samuel 2011).

2.2.1 A-to-I Editing of Paramyxovirus RNA

Measles virus (MV) is a single-stranded, negative-sense [ssRNA(−)] RNA virus of the Paramyxoviridae family, and its infection to humans is typically acute. However, subacute sclerosing panencephalitis (SSPE) (or Dawson disease) is a rare and late form of chronic progressive brain inflammation, which is due to slow infection with some hypermutated forms of the virus (Griffin et al. 2018). These were found to harbor ADAR-induced inosine clusters (Pfaller et al. 2014, 2015; Netzband and Pager 2020) in the genes encoding the M protein.

ADAR editing has also been detected in other Paramyxoviridae viruses, like the human respiratory syncytial virus (hRSV), human metapneumovirus (HMPV), and parainfluenza virus 5 (PIV5) (van den Hoogen et al. 2014; Rima et al. 2014; Netzband and Pager 2020), where it mainly affects nucleocapsid proteins (Pfaller et al. 2014, 2015, 2018). Notably, ADAR1-mediated A-to-I editing of endogenous dsRNA inhibits antiviral inflammatory and interferon responses (Liddicoat et al. 2015; Mannion et al. 2014) and thereby introduces mutations that suppress the innate

immune response (Pfaller et al. 2018) (Table 1 in chapter "RNA Modifications in Neurodegenerations").

2.2.2 A-to-I Editing of HDV RNA

The hepatitis δ virus (HDV) of the Deltaviridae family has a very small single-stranded, circular RNA. The unique protein that is encoded by its genome is the hepatitis delta antigen (HDAg), and it directs the replication and assembly of the virus (Table 1 in chapter "RNA Modifications in Neurodegenerations"). In order to be able to assemble, the virus needs the envelope proteins of the hepatitis B virus (HBV). Therefore, HDV can only disseminate its RNA on liver cells that have also been infected with HBV, usually leading to severe damage in the liver (Lempp et al. 2016).

Two forms of HDAg exist, a small (HDAg-S), which controls replication, and a large (HDAg-L), which inhibits and promotes the assembly of the virus. During its replication, an RNA intermediate is formed known as the antigenome, and this was shown to carry U-to-I edits at nucleotide 1012. A-to-I conversion changes the UAG amber stop codon of HDAg-S to UIG (or UGG), which encodes for tryptophan, and allows the elongated form of the protein (HDAg-L) to be translated.

The stabilized or destabilized secondary structure of the HDV RNA can also affect the extent of A-to-I editing. In addition, increased base pairing within 15–25 nucleotides at the 3′ of the editing site could markedly increase the editing events; whereas, reduced base pairing in the same region did not affect editing (Netzband and Pager 2020).

2.2.3 A-to-I Editing of HIV-1 RNA

The human immunodeficiency virus (HIV) is a member of the Retroviridae family. Lentiviruses are transmitted as positive-strand single-stranded [ssRNA(+)] RNA viruses. Once they enter the host cell's cytoplasm, the viral RNA genome is reverse transcribed into dsDNA, which is then imported into the cell nucleus and integrated into the DNA (Hu and Hughes 2012). ADAR-mediated A-to-I editing was found to regulate the replication of HIV-1 RNA, and thus, viral infection (McIntyre et al. 2018; Netzband and Pager 2020). In line with this, ADAR overexpression increased the abundance of HIV-1 proteins, while ADAR silencing depleted them. Editing was shown to depend on the activator of HIV-1 gene expression, Tat protein (Netzband and Pager 2020). Although A-to-I editing seems to greatly affect HIV gene expression, its exact extent is still widely unknown (Table 1 in chapter "RNA Modifications in Neurodegenerations").

2.2.4 A-to-I Editing in +ss Viral RNA Genomes

A-to-I editing is also found in +ssRNA viruses of the Flaviviridae family, including Hepatitis C virus (HCV), Zika virus (ZIKV), Dengue virus (DENV), and poliovirus from the family of Picornaviridae (McIntyre et al. 2018). In HCV RNA, A-to-I editing was found to associate with antiviral activity. ADAR was found to be involved in a potent antiviral pathway and to eliminate HCV RNA through A-to-I editing. In specific, the inhibition of both the IFN-induced protein kinase R (PKR) and ADAR1 stimulated replicon expression and reduced the amount of I's recovered from HCV RNA in replicon cells. This indicates that ADAR1 can limit the replication of the viral RNA (Netzband and Pager 2020) (Table 1).

Adding to HCV, the Zika virus is another RNA virus which understanding of variations is of major importance as reflected by the medical health emergency attributed to it. ZIKV infection has been linked to microcephaly and fetal death in humans. Its RNA genome encodes for ten proteins (7 non-structural and 3 structural) (Cox et al. 2015). One of these, called flavivirus envelope glycoprotein, encapsulates the virus and binds the host cell's endosomal membrane, initiating endocytosis (Dai et al. 2016). The RNA genome forms a nucleocapsid along with copies of the 12-kDa capsid protein. In turn, the nucleocapsid envelops within a host-derived membrane modified with two viral glycoproteins. Viral genome replication depends on the making of dsRNA from the single-stranded, ssRNA(+) genome followed by transcription and replication to provide viral mRNAs and new ssRNA(+) genomes (Ferrero et al. 2018). The ZIKV genome was found to contain higher guanosine nucleotide levels in both the positive (+) and negative (−) strands, which could be due to ADAR-induced A-to-I editing (Khrustalev et al. 2017; Piontkivska et al. 2017). Such editing events can also affect the evolution of rhabdovirus sigma.

Similar to the ZIKV genome, the genomes of many RNA viruses were found to harbor A-to-I or A-to-G modifications (I's are decoded as G's by ribosomes during translation and by polymerases during RNA-dependent RNA replication) (Samuel 2011). Nevertheless, confirmatory mechanistic studies should tell us more about the consequences of such editing events in these viruses.

2.3 Cytidine-to-Uridine (C-to-U) Editing

In contrast to the human genome, C-to-U is a less frequent type of editing among RNA viruses. The APOBEC3 ("apolipoprotein-B mRNA editing enzyme, catalytic polypeptide-3") family contains seven members (APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H), all of which are involved in innate immune responses to retroviruses and endogenous retroelements (Orecchini et al. 2018; Desimmie et al. 2016; Simon et al. 2015; Harris and Dudley 2015; Milewska et al. 2018). Of these, APOBEC3G was the first member found to possess antiviral activity against HIV, in the absence of its viral

Table 1 The role of epitranscriptomic marks on RNA viruses

Epitranscriptomic modification	RNA virus type	Function	References
Inosine (I)	MV, hRSV, PIV5, HMPV	Introduces mutations that suppress the innate immune response	Pfaller et al. (2018); Griffin et al. (2018)
	HDV	Controls virus replication and virus assembly	Netzband and Pager (2020)
	HIV	Regulates the viral gene expression	McIntyre et al. (2018), Netzband and Pager (2020)
	HCV	Restricts virus replication	Netzband and Pager (2020)
	ZIKV, DENV, poliovirus	Unknown	McIntyre et al. (2018)
Pseudouridine (Ψ)	TYMV, BMV	Stabilizes RNA structure	Netzband and Pager (2020)
Cytidine to uridine	Retroviruses	Leads to G-to-A hypermutations in the viral genome reducing viral fitness	Lerner et al. (2018)
N6-methyladenosine (m^6A) *Nuclear-replicating viruses*	HIV-1	Affects viral RNA shuttle out of the nucleus Affects viral RNA stability Regulates virus gene expression during early & late infection Specific m^6A sites impact virus infectivity	Lichinchi et al. (2016a, b), Kennedy et al. (2016), Tirumuru et al. (2016), Ciuffi (2016)
	MLV	Affects viral gene expression	Courtney et al. (2019a, b)
	IAV	Promotes viral gene expression Promotes virus pathogenicity	Netzband and Pager (2020), Courtney et al. (2017)
	RSV	Promotes viral RNA splicing	Netzband and Pager (2020)
Cytoplasm-replicating viruses	HCV	Suppresses virion production	Gokhale et al. (2016)
	ZIKV, EV71	Regulates viral gene expression	Hao et al. (2019), McIntyre et al. (2018), Gokhale et al. (2016)
	DENV, WNV, YFV, poliovirus	Unknown	McIntyre et al. (2018), Gokhale et al. (2016)

(continued)

Table 1 (continued)

Epitranscriptomic modification	RNA virus type	Function	References
N6, 2′-*O*-Dimethyladenosine (m^6Am)	IAV	Unknown	Gonzales-van Horn and Sarnow (2017)
5-Methylcytosine (m^5C)	SINV	Regulates innate immune response	Bhattacharya et al. (2017)
	HIV-1	Probably affects the viral mRNA translation Promotes viral RNA splicing	Courtney et al. (2019a)
	MLV	Affects the viral gene expression Enhances the virus infectivity	Courtney et al. (2019b)
	DENV, ZIKV, HCV, poliovirus	Unknown	McIntyre et al. (2018)
2′-*O*-methylated ribose *Residues at 5′ end*	Members of Corona-, Flavi-, pox-viruses that replicate in the cytoplasm	Restricts innate immune response	Netzband and Pager (2020), Gonzales-van Horn and Sarnow (2017), Hyde and Diamond (2015)
Internal residues	DENV	Inhibits the elongation by the viral RNA-dependent RNA polymerase	Dong et al. (2012)
	HIV-1	Restricts innate immune response	Ringeard et al. (2019)
	ZIKV, HCV, poliovirus	Unknown	Lichinchi et al. (2016a, b), McIntyre et al. (2018)

BMV Brome mosaic virus, *DENV* Dengue virus, *EV71* Enterovirus 71, *HCV* Hepatitis C virus, *HDV* hepatitis δ virus, *HIV-1* Human immunodeficiency virus type 1, *HMPV* human metapneumovirus, *hRSV* human respiratory syncytial virus, *IAV* Influenza A virus, *MLV* Murine leukemia virus, *MV* Measles virus, *PIV5* parainfluenza virus 5, *RSV* Rous sarcoma virus, *SINV* Sindbis virus, *TYMV* Turnip yellow mosaic virus, *WNV* West Nile virus, *YFV* Yellow Fever virus, *ZIKV* Zika virus

neutralization partner, Vif (virion infectivity factor) (Lerner et al. 2018). Polymorphisms in human APOBEC3 members can affect their antiviral/retroviral activity and hence, a person's predisposition to infection by specific viruses and associated co-morbidities, such as HBC-associated hepatocellular carcinoma (He et al. 2017; Duggal et al. 2013). All APOBEC3 members can deaminate C-to-U in ssDNA formed during the viral/retroelement life cycle, leading to G-to-A transitions and reducing the virus's fitness (Lerner et al. 2018) (Table 1).

2.4 2′-O-*methylation of Ribose*

2.4.1 2′-*O*-methylation of Ribose at the 5′ End of Viral RNA

2′-*O*-methylation of the ribose represents a highly abundant modification detected in all types of eukaryotic RNAs and identified on all 4 ribonucleotides (A, U, G, C), as well as on edited nucleotides, like Ψ's and I's. Eukaryotic mRNAs harbor together with the 7-methylguanosine cap (m^7G), 2′-*O*-methylation on the first and sometimes on both the first and second transcription nucleotides. 2′-*O*-methylation protects mRNA from degradation, facilitates pre-mRNA splicing, nuclear export and translation, and functions as a signature that recognizes RNA as "self" to escape innate immune cascades (Zhao et al. 2017; Daffis et al. 2010). This strategy has also been adopted to knockdown genes successfully via RNA interference mechanism, where synthesized siRNAs have to be 2′-*O*-me-modified in order to avoid immune responses once transfected into the cell.

Mechanistically, 2′-*O*-methylated RNAs bind to the innate immune receptor Toll-like receptor 7 (TRL7) and inhibit the secretion of inflammatory cytokines, such as IFN-α and IFN-β. Once expressed, IFN-α and IFN-β bind to IFN receptor (IFNAR) and stimulate signaling in an autocrine and paracrine manner, inducing the expression of genes with specific antiviral activities, such as translation inhibition. More specifically, IFIT (interferon-induced protein with tetratricopeptides) proteins belong to the ISG group and interact with eIF3 (eukaryotic translation initiation factor), blocking the assembly of the 43S preinitiation complex and resulting in translation inhibition of the viral RNAs (Netzband and Pager 2020; Gonzales-van Horn and Sarnow 2017). Recent X-ray crystallography studies unveil that IFIT proteins are able to recognize and bind only m^7G-capped and non-2′-*O*-methylated RNAs, while they exclude all the m^7G-capped and 2′-*O*-methylated ones. Similarly, the cytoplasmic factor, RIG-I (retinoic acid-inducible gene I) activates the innate immune pathway by binding uncapped RNAs harboring a 5′-end triphosphate (Daffis et al. 2010). Altogether, IFIT and RIG-I proteins by binding non-2′-*O*-methylated nucleotides at the 5′-end of the RNA distinguish the host from the pathogenic RNAs (Netzband and Pager 2020; Gonzales-van Horn and Sarnow 2017; Hyde and Diamond 2015).

Many viruses that belong in Coronaviridae, Flaviviridae, Poxviridae, and all viruses that replicate in the cytoplasm, appear to use this mechanism by encoding their own 2′-*O* methyltransferase (2′-*O* MTase) to evade the innate immune sensing. Members of these families express dual-specific methyltransferases that act both by methylating the N^7G cap and by 2′-*O*-methylating the 2′-OH group of the ribose of the first nucleotide on the viral RNA (Decroly et al. 2012). Notably, WNV, a member of the Flaviviridae family, appears less virulent in the absence of 2′-*O* MTase activity, indicating that 2′-*O*-methylation modulates viral pathogenicity by escaping translation attenuation that ISG genes cause.

Furthermore, when human coronavirus, lacking the Nsp16 viral 2′-*O* MTase activity, infects primary human macrophages and stimulates the IFN-β production

in higher levels than wild-type coronavirus. This interferon expression is MDA-5 dependent (melanoma differentiation-associated protein 5), and *mda5* knockout murine macrophages failed to produce detectable levels of IFN-β upon infection with a 2′-*O* MTase deficient murine coronavirus (Züst et al. 2011). Collectively, viral RNAs have adopted the 2′-*O*-methylation strategy as a mask to delude the innate immune signaling by avoiding their recognition by the cytoplasmic RNA sensors Mda5, TLR7, RIG-I, IFIT1, and IFIT2 (Table 1).

2.4.2 2′-*O*-methylation at Internal Ribose Residues of Viral RNA

2′-*O*-methylations are not only found on the first adenosine of the viral genome, but also on adenosines within it. While 2′-*O*-methylation at the 5′-end of the RNA acts as a mask to evade innate immune responses, the role of internal 2′-*O*-methylated nucleotides in viral, as well as in cellular RNAs remains incomplete.

Recombinant 2′-*O* MTases of DENV can specifically and internally methylate adenosines of uncapped truncated viral RNAs and host ribosomal RNAs, without any sequence specificity requirement (Dong et al. 2012). Notably, mutagenesis studies revealed that the active site of the viral methyltransferase that catalyzes the internal 2′-*O*-methylation is composed of the K-D-K-E motif, similarly to the motif required for the 2′-*O* methylation of the 5′ cap. The 2′-*O*-methylation of adenosines in viral RNA inhibits its replication and elongation by the viral RNA-dependent RNA polymerase (Dong et al. 2012) (Table 1). Integration of 2′-*O*-methylated U, G, and A into short RNA strongly impairs the TLR7-mediated production of cytokines (Petes et al. 2017). Therefore, apart from affecting gene expression of the virus and its interaction with the host, the internal 2′-*O*-methylated A's could also restrict the innate immune response upon viral infection. However, such a function needs to be verified.

Yet, the internal 2′-*O*-methylation is not a global modification, since the human metapneumovirus (hMPV), a (−) RNA virus, lacks this activity. However, the Ebola virus, another negative-sense RNA virus, exhibits internal 2′-*O* adenosine MTase activity, but the role of the produced modifications needs further investigation (Martin et al. 2018).

On the contrary, HIV-1 and MLV genomes contain internal 2′-*O*-methylated nucleotides added by the cellular 2′-*O* MTase, FTSJ3 (Fig. 1). In particular, HIV-1 RNA transcription is accelerated by Tat protein that interacts with the TAR element at the 5′-end of the gRNA, which then recruits cellular RNA polymerase II. The cellular TAR RNA-binding protein (TRBP), which interacts with the HIV-1 RNA, binds FTSJ3 and MAT2A (a SAM metabolism protein), forming a complex. As such, it is shown that FTSJ3 can 2′-*O*-methylate the m^7G cap of HIV-1 RNA in vitro and an oligonucleotide containing 27-A's, U's, or G's. Thus, RiboMethSeq analyses detected 17 2′-*O*-methylated nucleotides within HIV-1 RNA isolated from viral particles. HIV-1 RNA lacking or harboring a reduced number of 2′-*O*-methylated nucleotides could significantly increase the expression of IFN-α and IFN-β by activating the Mda5 cytosolic RNA sensor (Ringeard et al. 2019). Thus, similar to

2′-*O*-methylation of the first adenosine, internal 2′-*O*-methylation may also enable viruses to escape immune sensing. It is also possible that these 2′-*O*-methylated nucleotides within the HIV-1 genome might limit the activity of reverse transcriptase, similar to the internal 2′-*O*-methyladenosines, which restrict the DENV NS5 RNA-dependent RNA polymerase (Dong et al. 2012). Nevertheless, this remains to be elucidated.

In spite of HIV-1, 2′-*O*-methyladenosines, 2′-*O*-methylguanosines, 2′-*O*-methylcytosines, and 2′-*O*-methyluridines have also been detected on the gRNA of DENV, ZIKV, HCV, and poliovirus (McIntyre et al. 2018; Lichinchi et al. 2016a, b). Cellular 2′-*O* MTase activities appear to be responsible for internal 2′-*O*-methylation of poliovirus and HCV RNA that has been isolated either from infected cells or released virions. Internal 2′-*O*-methyladenosines in the genome of DENV and ZIKV viruses most probably result from the activity of viral methyltransferases; whereas, the addition of 2′-*O*-methylation to other nucleotides might be the results of cellular 2′-*O* methyltransferases.

2.5 N6-Methyladenosine

N6-methyladenosine (m^6A) is a well-established type of co-transcriptional RNA editing. Recently, the emergence of high-throughput and deep RNA sequencing approaches, along with advanced classic analytical tools, like liquid chromatography coupled with mass spectrometry, and the availability of specific anti-m^6A antibodies allowed the detailed mapping of m^6A marks by performing multiple novel techniques—including PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation), PA-m^6A-seq (photo-crosslinking-assisted m^6A-seq), m^6A-seq and MeRIP-seq (m^6A-specific methylated RNA immunoprecipitation) (Tan and Gao 2018; Li et al. 2017). Therefore, we now recognize that m^6A is a reversible and abundant type of editing that is detected both in cellular and viral RNA genomes (Kennedy et al. 2017; Gonzales-van Horn and Sarnow 2017). m^6A are preferentially located near the 3′ untranslated region (UTR), near stop codon, and long exons of cellular mRNAs (Dang et al. 2019; Dominissini et al. 2012). Their identification and targeted mutagenesis have shed light into the actual role of this modification and revealed the proteins that write, erase or read m^6A marks. Hence, m^6A is added to nuclear pre-mRNA by a multimeric complex comprised of the catalytic subunit methyltransferase METTL3 (Methyltransferase Like 3), the RNA-binding protein METTL14 (Methyltransferase Like 3), and the cofactors WTAP (Wilms tumor 1-associated protein) and KIAA1429. WTAP is required for METTL3/METTL14 complex localization into nuclear speckles and improves the m^6A modification efficiency (Shi et al. 2019; Kobayashi et al. 2018; Scholler et al. 2018; Ping et al. 2014), while KIAA1429 possess a regulatory role and is essential for an efficient METTL3/METTL14 complex.

On the other hand, the removal of m^6A from the modified mRNAs is catalyzed by the demethylases FTO (fat mass and obesity-associated protein) and ALKBH5

(α-ketoglutarate-dependent dioxygenase AlkB homolog 5) (Zheng et al. 2013; Jia et al. 2011) (Fig. 2).

In addition, a combination of affinity chromatography and mass spectrometry identified several reader proteins for m^6A, among which the most well studied belong to the YTH-domain-containing proteins. Five of them are found in humans, termed YTHDF1–3 and YTHDC1–2. YTHDF1and YTHDC2 are cytoplasmic, whereas YTHDC1 is nuclear (Shi et al. 2017, 2019). Specifically, YTHDF1 enhances the translation efficiency of bound mRNAs, while YTHDF2 induces the destabilization and decay of m^6A-modified mRNAs by recruiting the CCR4-NOT deadenylase complex (Du et al. 2016). On the other hand, YTHDF3 promotes the function of either YTHDF1 or YTHDF2, while YTHDC2 regulates the mRNA levels during meiosis, and YTHDC1 is involved in nuclear gene splicing. These data have attracted the interest of the scientific community, and thus, several studies lately have focused on mapping such m^6A residues on viral RNA genomes in an effort to determine their regulatory role on viral infection and pathogenicity.

2.5.1 N6-methyladenosine at Genomic RNA of Nucleus-Replicating Viruses

Influenza A virus (IAV) was the first virus where internal m^6A residues were detected on its genomic RNA (gRNA). IAV as a member of the Orthomyxoviridae family contains a negative-sense, single-stranded gRNA (-ssRNA) segmented in eight subgenomic parts. m^6A modifications are randomly distributed among different IAV mRNAs, however, both the IAV gRNA and mRNA are more frequently modified at the 3′-end, similarly to the cellular mRNAs (Courtney et al. 2017; Kennedy et al. 2017; Netzband and Pager 2020). Recent studies clearly showed the m^6A impact on the virus gene expression, when mutation of m^6A sites or loss of the METTL3 activity decreased IAV protein and mRNA levels, leading to reduction of virus replication and load. On the other hand, ectopic overexpression of YTHDF2, but not of YTHDF1 or YTHDF3, enhances the replication and production of infectious IAV virions. Most importantly, m^6A-deficient IAV mutants exhibited decreased pathogenicity when introduced into mice, which remarkably indicates that the m^6A marker is a positive regulator of IAV production and infection (Netzband and Pager 2020; Courtney et al. 2017).

In addition to IAV, four retroviruses, Avian sarcoma virus (ASV), Rous sarcoma virus (RSV), HIV, and Murine Leukemia virus (MLV) contain m^6A-modified gRNAs, presenting, surprisingly, higher m^6A modification level than average cellular mRNA.

Using [methyl-^{3}H]methionine to label methyl-moieties on B77 ASV RNA, around 14 to 16 internal m^6A marks were detected in total, excluding those found proximal to the m^7G cap structure that is not randomly distributed on RNA transcripts. The consensus motif that is recognized by the m^6A writer complex was first revealed from studies on ASV and was identified to be DRm^6ACH (where D

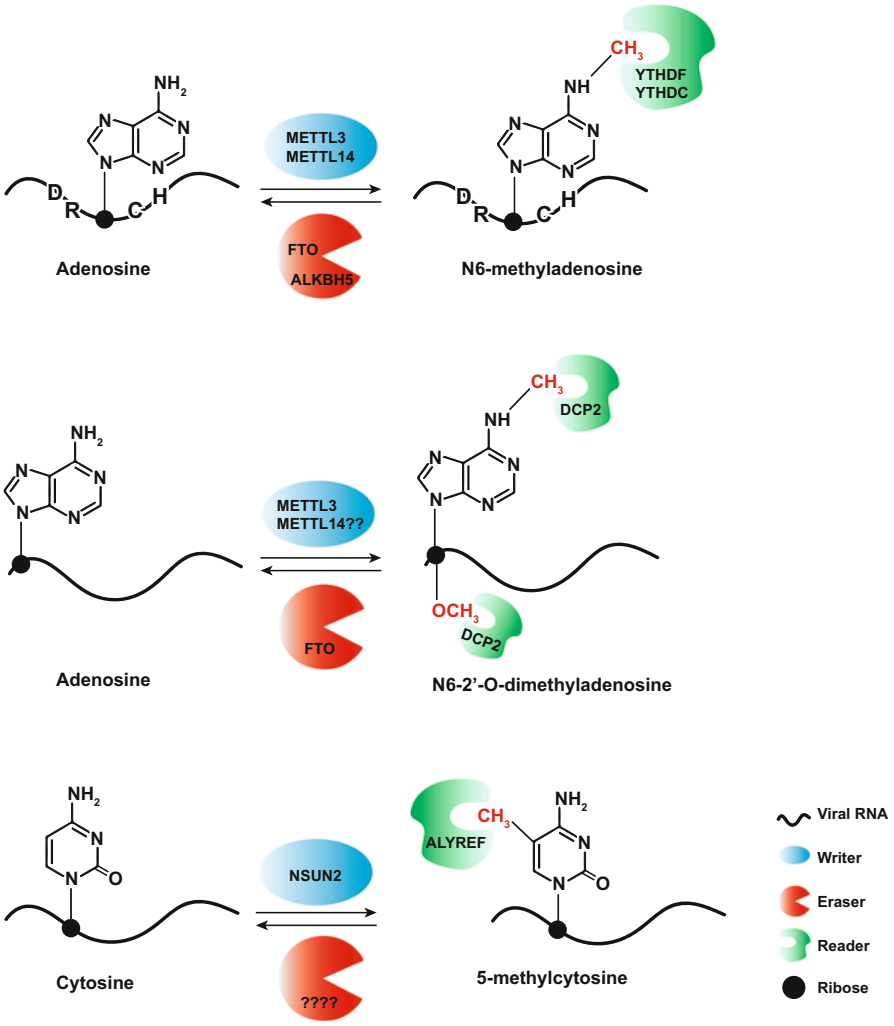


Fig. 2 Modification of viral RNA by cellular enzymes. N6-methyladenosine, N6–2′-*O*-dimethyladenosine, and 5-methylcytosine are reversible RNA modifications. The consensus motif DR(m^6A)CH recognized by the m^6A writer complex is illustrated, where D corresponds to bases G, A, or U; R corresponds to G and A purines; and H to U, C, or A. Modified bases are indicated with red letters. The enzymes responsible for each modification are also indicated, where METTL3/METTL14 corresponds to Methyltransferase Like 3/14 and NSUN2 to NOP2/Sun RNA Methyltransferase 2. The proteins that remove or read these modifications are demonstrated, as well, where FTO corresponds to fat mass and obesity-associated protein, ALKBH5 to α-ketoglutarate-dependent dioxygenase AlkB homolog 5, YTHDF/C to YTH domain-containing family protein F/C, DCP2 to mRNA-decapping enzyme 2 and ALYREF to Aly/REF Export Factor. Question marks (?) indicate that enzymes implicated in the respective modification reaction have not been yet identified

corresponds to G, A, or U; R corresponds to G and A purines; and H to U, C, or A) (Dominissini et al. 2012; Liu et al. 2014) (Fig. 2).

Additional experiments identified 10–15 m^6A modifications, the majority of which were accumulated within the 3′-end and upstream of the consensus splicing acceptor sequence on the RSV RNA. In the same context, cell treatment with cycloleucine, an inhibitor of internal methylation, resulted in the accumulation of unspliced viral RNAs, indicating the potential role of m^6A in regulating RSV gRNA splicing (Netzband and Pager 2020).

Recently, MeRIP-Seq analyses mapped numerous m^6A sites in the 5′ and 3′ long terminal regulatory repeats (LTR), *env* and *rev* genes, and the Rev response element (RRE) of the HIV-1 genome (Lu et al. 2018; Lichinchi et al. 2016a, b; Ciuffi 2016; Kennedy et al. 2016; Tirumuru et al. 2016). RRE is an RNA stem-loop structure present in full-length unspliced viral RNA and it is essential for the binding of the Rev protein to achieve an efficient export of the gRNA to the cytoplasm. Two conserved m^6A sites within the RRE region enhance Rev binding and stimulate the viral RNA export from the nucleus, and subsequently the translation, virions production, and release (Lichinchi et al. 2016a, b). Interestingly, the consensus motif of m^6A appears to be highly conserved among HIV isolates, an observation that contradicts the plasticity that the HIV genome exhibits. Altogether, these data clearly suggest that m^6A marks contribute to viral RNA stability and infectivity. Notably, it has been shown that infection of a human CD4+ T cell line with HIV results in a global increase of m^6A positions both on viral and host RNAs. In addition, some cellular transcripts exhibit an altered methylation profile before and after HIV infection, including mRNAs of TRAF2, PABPC3, and ETS2, all of which are known to have pro-viral functions (Lichinchi et al. 2016a, b).

Moreover, changes of the writer complex METTL3/METTL14, FTO, and ALKBH5 erasers, as well as YTHDF1-YTHDF3 readers abundance, affects the gene expression of HIV-1, both during early or late infection (Lu et al. 2018; Lichinchi et al. 2016a, b; Tirumuru et al. 2016). Specifically, silencing of the writer complex decreases gp120 viral transcript (glycoprotein of the HIV envelop), p24 protein level (component of the HIV virion capsid), and virus replication up to five-fold. On the contrary, silencing of the demethylase ALKBH5 increases viral expression and gRNA replication up to eight-fold. Regarding HIV assembly, m^6A marks seem to play a critical role, as well. The primer-binding site (PBS) is a critical region that promotes the initiation of the HIV gRNA reverse transcription, while the dimer initiation sequence (DIS) is required for the HIV genome dimerization and encapsidation. Disruption of two m^6A sites within PBS and in the proximity of the 5′ UTR of DIS largely impacts the subsequent HIV infections.

Additional studies on the retrovirus MLV identified 20 m^6A sites within the MLV genome, and mutation of three of them located within the *env* gene reduced both intracellular and viral proteins level and the viral gene expression in subsequent infections (Courtney et al. 2019a, b).

2.5.2 N6-methyladenosine at Genomic RNA of Cytoplasm-Replicating Viruses

Considering that the m^6A marks are post-transcriptionally added to the RNA by the nuclear writer complex METTL3/METTL14, along with the accessory factors WTAP and KIAA1429, m^6A moieties have unexpectedly been detected on mRNA transcripts encoded by several RNA viruses that exclusively replicate in the cytoplasm. However, some data show that METTL3 and METTL14 may shuttle between the nucleus and the cytoplasm raising the question of m^6A function in cytoplasm-replicating RNA viruses again. Thus, m^6A-modified residues have been detected at the gRNA of Flaviviridae (Zika, Dengue, Hepatitis C, West Nile, Yellow Fever viruses) (Gokhale et al. 2016) and Picornaviridae families (Poliovirus, Enterovirus 71) (McIntyre et al. 2018; Hao et al. 2019). All members of both families contain a +ssRNA that replicates and is translated exclusively in the cytoplasm.

In HCV, 14 m^6A-modified sites have been identified across the gRNA. Although neither overexpression of METTL3 and METTL14 nor knockdown of FTO eraser enzyme affect the gRNA replication, they do change the viral proteins level, resulting in an increase of the virion production and titers. Notably, during HCV infection, YTHDF readers colocalize with the HCV Core protein in the lipid droplets, sites where viral assembly and morphogenesis occur, suggesting the potential regulatory role of these proteins in virus assembly. HCV m^6A-modified gRNA is recognized by all three methyl YTHDF readers, which in contrast to cellular mRNAs, promote viral RNA retention in the cytoplasm and suppress virion production. Moreover, disruption of m^6A sites in the envelope *E1* gene does not affect HCV gRNA replication and translation but highly increases the virus titers and promotes the interaction between YTHDF2 and HCV Core protein (Gokhale et al. 2016). Altogether this data indicates that m^6A marks can negatively modulate the viral RNA package into virions without affecting the viral replication.

Finally, m^6A moieties regulate the gene expression of ZIKV and enterovirus 71 gRNA. In the case of ZIKV, METTL3 and METTL14 writers suppress the viral replication, and all three YTHDF1, YTHDF1 and YTHDF3 readers reduce viral replication by causing the gRNA decay (Lichinchi et al. 2016a, b).

Other members of the Flaviviridae family, including the West Nile virus (WNV), DENV and Yellow Fever virus (YFV), have been described to bear m^6A marks in conserved sites within the NS3, NS5 non-structural genes and 3′ UTRs, making m^6A a conserved regulatory moiety across this family (McIntyre et al. 2018; Gokhale et al. 2016; Hao et al. 2019).

2.6 N6, 2-O-dimethyladenosine

Apart from the common m^6A, another prevalent and reversible modification, *N*6,2-*O*-dimethyladenosine (m^6Am), has been identified on cellular mRNAs and viral

gRNAs. m^6Am is located at the 5′ end of most viral RNAs and all eukaryotic mRNAs, and in particular in the first adenosine next to the 7-methylguanosine cap (m^7GpppN) (Mauer et al. 2017; Gonzales-van Horn and Sarnow 2017). The proximity of this adenosine to the m^7G cap makes it a preferential substrate for the demethylase FTO, which in fact exhibits a 100-fold higher catalytic efficiency toward m^6Am than m^6A. Interestingly, the reader protein of m^6Am residues is DCP2, the enzyme that removes the cap from mRNAs. m^6Am-modified transcripts appear to be more stable compared to mRNAs that do not contain adenosine as the first transcription nucleotide, since m^6Am protects them from the decapping enzyme DCP2.

On the other hand, the writer methyltransferase that produces this cap-dependent moiety has not yet been identified, though the multicomponent METTL3/METTL14 complex is expected to produce this modification, as well.

Although m^6Am has been identified in IAV gRNA as early as 1976, its role in the viral life cycle remains essentially unexplored (Netzband and Pager 2020; Gonzales-van Horn and Sarnow 2017). Similar to its role in cellular mRNAs, it could also be implicated in the viral RNA stability, modulating the viral replication and translation rate within the host.

2.7 5-Methylcytosine

Although N6-methyladenosine has been considerably studied, less is known about the function of 5-methylcytosine. Methylcytosine is another post-transcriptional modification that results from the addition of a methyl group to the C5 position of cytosine (m^5C) (Li et al. 2017; Squires et al. 2012). m^5C modifications have been detected in highly abundant cellular RNAs, such as tRNAs, rRNAs and mRNAs; however, the first m^5C marks were found intracellularly on the Sindbis virus (SINV) RNA, a +ssRNA virus, member of the Togaviridae family (Netzband and Pager 2020). Though common, the regulatory role of this modification, either in cellular or viral RNA, is still under investigation. Despite the fact that none conserved RNA motif has been verified, these moieties are most frequently found in CG-rich regions, near the translation initiation sites and 3′-UTRs of cellular mRNAs (Yang et al. 2017).

Nowadays, the performance of UPLC-MS/MS (ultra-performance liquid chromatography-tandem mass spectrometry) and PA-m^5C-seq permitted m^5C mapping and the identification of the m^5C writer and reader proteins. Therefore, the human genome encodes seven m^5C RNA methyltransferases, which are all members of the NSUN family that use SAM (*S*-adenosyl methionine) as the methyl donor group (Squires et al. 2012). Extensive studies elucidated NSUN2 as the writer methyltransferase for the m^5C formation on viral RNAs, including gRNA of retroviruses, such as MLV and HIV (Courtney et al. 2019a, b). Moreover, ALYREF corresponds to the reader protein, which acts as an essential chaperone that

contributes to the shuttle of processed mRNAs from the nucleus to the cytoplasm (Yang et al. 2017). However, the demethylase that "erases" m^5C moieties is not yet known.

HIV gRNA, derived from either virus particles or infected cells, is highly m^5C-modified. In particular, each viral gRNA contains around 11–19 m^5C modifications, most commonly detected at the 3′ end of the viral RNA, corresponding to 14-fold higher modification level compared to the average cellular mRNAs. Of note, knockout of NSUN2 in HIV-infected cells decreases m^5C abundance, which in turn results in the decrease of viral proteins (Li et al. 2017; Xing et al. 2015). Moreover, lack of m^5C on the HIV-1 gag mRNA, though it is efficiently exported from the nucleus, it binds to the ribosomes with lower affinity. Hence, m^5C probably affects the viral mRNA translation via its deficient loading on the ribosome. Yet, it has to be determined whether this occurs through the cap-independent translation mechanism. Notably, m^5C marks are implicated in HIV-1 viral RNA splicing regulation. Specifically, splicing site A2, which contains an m^5C residue 150 nucleotides upstream, is not well recognized when NSUN2 is absent (Courtney et al. 2019a, b).

MLV gRNA is also highly modified similarly to HIV-1 gRNA. Loss of m^5C, by substituting m^5C with U, decreases the expression of MLV Gag p65 and p30 proteins by two-fold, releasing fewer virions. In the same line, disruption of the m^5C sites in the polymerase gene has no effect on the virions assembly. Similarly, knockdown of the writer NSUN2 modestly decreases the protein levels of MLV Gag, in contrast to NSUN2 overexpression that had no effect, suggesting that m^5C modification could be a clever strategy that retroviruses use to increase their replication efficiency. Consistent with this data, the newly produced virions that lack m^5C modifications reveal a reduced infectivity when used for a subsequent round of infection, implying a putative role of m^5C in camouflaging the viral RNA from innate immune sensing (Courtney et al. 2019a, b). Overall, viruses by hijacking the m^5C writer and reader proteins of the infected cells promote their genome replication, expression and pathogenicity. However, m^5C marks role remains unclear for Dengue virus, Zika virus, HCV and poliovirus (McIntyre et al. 2018).

In contrast to retroviruses, where m^5C is detected on the viral RNA that has been isolated from both cells and virions, in SINV gRNA, the m^5C marks are only detectable within the intracellular viral RNA required for the virions structural assembly, but not within the RNA isolated from the released virions. This points toward that this epitranscriptomic modification must be a critical regulator of the SINV gene expression and of the host's innate immune response (Bhattacharya et al. 2017).

3 Conclusions and Future Perspectives

Although in its infancy, viral epitranscriptomics is a rapidly progressed field, supported by the emergence and advance of sensitive high-throughput techniques and reagent availability. Viral epitranscriptomics intends to clarify the virus-host

interactions and pathways that regulate the viral life cycle and infection efficiency. In spite of being parasitic entities, viruses have exquisitely evolved to hijack the cellular epitranscriptomic modifying factors for their own genome and transcripts modification in order to skip the host's innate immune response and enhance their replication and virions release. Therefore, the detailed mapping of specific modification marks on the viral RNAs and the identification of their exact functional role, as well as the determination of the writers that establish, the erasers that remove and the readers of such modified nucleotides, is crucial and of major importance for antiviral drug development. Especially today that we are experiencing the COVID-19 pandemic, an issue of the highest priority worldwide, we necessitate an urgent search for new molecular targets and novel virus-specific and effective antiviral therapies. SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) is a positive-sense, single-stranded RNA virus responsible for the COVID-19 pandemic. For example, SARS-CoV-2 transcripts, similarly to all members within the Coronaviridae family, encodes the cofactor nsp10 that harbors 2′-*O*-RNA methyltransferase activity required for nsp16 proper function, a viral non-structural protein. It has been shown that ablation of *nsp16* reduces the viral RNA synthesis by ten-fold and that mice vaccination with nsp16-defective SARS-CoV-1 protects them from death, making nsp16 or the complex nsp16-nsp10 promising and SARS-CoV-specific drug targets (Viswanathan et al. 2020; Krafcikova et al. 2020). Moreover, several other drugs that target the viral epitranscriptomic pathways appear promising. As such, 3-deazaadenosine (DAA), an inhibitor of the intracellular methyl donor SAM, has been reported as a potent suppressor of many viruses in mice and rats (Kennedy et al. 2016; Courtney et al. 2017).

Therefore, major questions, such as how writers select specific nucleotides on RNA to modify and whether the observed changes on the viral gene expression is a direct or indirect result of the viral epitranscriptome alterations, remain largely undefined and awaits to be addressed. Moreover, a deeper understanding of the viral epitranscriptome and its dynamics will shed light not only on the viral pathogenesis, but it will also likely provide insight into a plethora of human RNA-dependent diseases.

References

Anderson BR, Muramatsu H, Nallagatla SR et al (2010) Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. Nucleic Acids Res 38:5884–5892

Anderson BR, Muramatsu H, Jha BK et al (2011) Nucleoside modifications in RNA limit activation of 2′,5′-oligoadenylate synthetase and increase resistance to cleavage by RNase L. Nucleic Acids Res 39:9329–9338

Baumstark T, Ahlquist P (2001) The brome mosaic virus RNA3 intergenic replication enhancer folds to mimic a tRNA TψC-stem loop and is modified in vivo. RNA 7:1652–1670

Bhattacharya T, Newton ILG, Hardy RW (2017) Wolbachia elevates host methyltransferase expression to block an RNA virus early during infection. PLoS Pathog 13:e1006427

Carlile TM, Rojas-Duran MF, Zinshteyn B et al (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515:143–146
Charette M, Gray MW (2000) Pseudouridine in RNA: what, where, how, and why. IUBMB Life 49:341–351
Christofi T, Zaravinos A (2019) RNA editing in the forefront of epitranscriptomics and human health. J Transl Med 17:319
Ciuffi A (2016) Viral cell biology: HIV RNA gets methylated. Nat Microbiol 1:16037
Courtney DG, Kennedy EM, Dumm RE et al (2017) Epitranscriptomic enhancement of influenza a virus gene expression and replication. Cell Host Microbe 22:377–386.e5
Courtney DG, Chalem A, Bogerd HP et al (2019a) Extensive epitranscriptomic methylation of A and C residues on murine leukemia virus transcripts enhances viral gene expression. MBio 10: e01209–e01219
Courtney DG, Tsai K, Bogerd HP et al (2019b) Epitranscriptomic addition of m^5C to HIV-1 transcripts regulates viral gene expression. Cell Host Microbe 26:217–227
Cox BD, Stanton RA, Schinazi RF (2015) Predicting Zika virus structural biology: challenges and opportunities for intervention. Antivir Chem Chemother 24:118–126
Cuesta R, Yuste-Calvo C, Gil-Cartón D et al (2019) Structure of turnip mosaic virus and its viral-like particles. Sci Rep 9:15396
Daffis S, Szretter KJ, Schriewer J et al (2010) 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature 468:452–456
Dai L, Song J, Lu X et al (2016) Structures of the Zika virus envelope protein and its complex with a flavivirus broadly protective antibody. Cell Host Microbe 19:696–704
Danecek P, Nellåker C, McIntyre RE et al (2012) High levels of RNA-editing site conservation amongst 15 laboratory mouse strains. Genome Biol 13:26
Dang W, Xie Y, Cao P et al (2019) N6-Methyladenosine and viral infection. Front Microbiol 10:417. (1-112)
Decroly E, Ferron F, Lescar J et al (2012) Conventional and unconventional mechanisms for capping viral mRNA. Nat Rev Microbiol 10:51–65
Desimmie BA, Burdick RC, Izumi T et al (2016) APOBEC3 proteins can copackage and comutate HIV-1 genomes. Nucleic Acids Res 44:7848–7865
Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m6A RNA methylomes revealed by m^6A-seq. Nature 485:201–206
Dong H, Chang DC, Hua MHC et al (2012) 2′-O methylation of internal adenosine by flavivirus NS5 methyltransferase. PLoS Pathog 8:e1002642
Du H, Zhao Y, He J et al (2016) YTHDF2 destabilizes m^6A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat Commun 7:12626
Duggal NK, Fu W, Akey JM et al (2013) Identification and antiviral activity of common polymorphisms in the APOBEC3 locus in human populations. Virology 443:329–337
Ferrero D, Ferrer-Orta C, Verdaguer N (2018) Viral RNA-dependent RNA polymerases: a structural overview. Subcell Biochem 88:39–71
George CX, John L, Samuel CE (2014) An RNA editor, adenosine deaminase acting on double-stranded RNA (ADAR1). J Interf Cytokine Res 34:437–446
Gokhale NS, McIntyre ABR, McFadden MJ et al (2016) N6-methyladenosine in Flaviviridae viral RNA genomes regulates infection. Cell Host Microbe 20:654–665
Gonzales-van Horn SR, Sarnow P (2017) Making the mark: the role of adenosine modifications in the life cycle of RNA viruses. Cell Host Microbe 21:661–669
Griffin DE, Lin WHW, Nelson AN (2018) Understanding the causes and consequences of measles virus persistence. F1000Res 7:237
Hao H, Hao S, Chen H et al (2019) N6-methyladenosine modification and METTL3 modulate enterovirus 71 replication. Nucleic Acids Res 47:362–374
Harris RS, Dudley JP (2015) APOBECs and virus restriction. Virology 479-480:131–145

He XT, Xu HQ, Wang XM et al (2017) Association between polymorphisms of the APOBEC3G gene and chronic hepatitis B viral infection and hepatitis B virus-related hepatocellular carcinoma. World J Gastroenterol 23:232–241

Hu WS, Hughes SH (2012) HIV-1 reverse transcription. Cold Spring Harb Perspect Med 2:a006882

Hudson GA, Bloomingdale RJ, Znosko BM (2013) Thermodynamic contribution and nearest-neighbor parameters of pseudouridine-adenosine base pairs in oligoribonucleotides. RNA 19:1474–1482

Hyde JL, Diamond MS (2015) Innate immune restriction and antagonism of viral RNA lacking 2′-O methylation. Virology 479:66–74

Jia G, Fu Y, Zhao X et al (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887

Kennedy EM, Bogerd HP, Kornepati AVR et al (2016) Posttranscriptional m^6A editing of HIV-1 mRNAs enhances viral gene expression. Cell Host Microbe 19:675–685

Kennedy EM, Courtney DG, Tsai K et al (2017) Viral epitranscriptomics. J Virol 91:e02263–e02216

Khrustalev VV, Khrustaleva TA, Sharma N et al (2017) Mutational pressure in Zika virus: local ADAR-editing areas associated with pauses in translation and replication. Front Cell Infect Microbiol 7:44

Kobayashi M, Ohsugi M, Sasako T et al (2018) The RNA methyltransferase complex of WTAP, METTL3, and METTL14 regulates mitotic clonal expansion in adipogenesis. Mol Cell Biol 38: e00116–e00118

Krafcikova P, Silhan J, Nencka R et al (2020) Structural analysis of the SARS-CoV-2 methyltransferase complex involved in RNA cap creation bound to sinefungin. Nat Commun 11:3717

Lempp FA, Ni Y, Urban S (2016) Hepatitis delta virus: insights into a peculiar pathogen and novel treatment options. Nat Rev Gastroenterol Hepatol 13:580–589

Lerner T, Papavasiliou FN, Pecori R (2018) RNA editors, cofactors, and mRNA targets: an overview of the C-to-U RNA editing machinery and its implication in human disease. Genes (Basel) 10:13

Li X, Zhu P, Ma S et al (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol 11:592–597

Li Q, Li X, Tang H et al (2017) NSUN2-mediated m^5C methylation and METTL3/METTL14-mediated m^6A methylation cooperatively enhance p21 translation. J Cell Biochem 118:2587–2598

Lichinchi G, Gao S, Saletore Y et al (2016a) Dynamics of the human and viral m^6A RNA methylomes during HIV-1 infection of T cells. Nat Microbiol 1:16011

Lichinchi G, Zhao BS, Wu Y et al (2016b) Dynamics of human and viral RNA methylation during Zika virus infection. Cell Host Microbe 20:666–673

Liddicoat BJ, Piskol R, Chalk AM et al (2015) RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. Science 349:1115–1120

Liu J, Yue Y, Han D et al (2014) A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol 10:93–95

Lovejoy AF, Riordan DP, Brown PO (2014) Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*. PLoS One 9:e110799

Lu W, Tirumuru N, St. Gelais C et al (2018) N6-Methyladenosine-binding proteins suppress HIV-1 infectivity and viral production. J Biol Chem 293:12992–13005

Mannion NM, Greenwood SM, Young R et al (2014) The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. Cell Rep 9:1482–1494

Martin B, Coutard B, Guez T et al (2018) The methyltransferase domain of the Sudan ebolavirus L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure. Nucleic Acids Res 46:7902–7912

Mauer J, Luo X, Blanjoie A et al (2017) Reversible methylation of m^6Am in the 5′ cap controls mRNA stability. Nature 541:371–375

McIntyre W, Netzband R, Bonenfant G et al (2018) Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection. Nucleic Acids Res 46:5776–5791

Milewska A, Kindler E, Vkovski P et al (2018) APOBEC3-mediated restriction of RNA virus replication. Sci Rep 8:5960

Netzband R, Pager CT (2020) Epitranscriptomic marks: emerging modulators of RNA virus gene expression. Wiley Interdiscip Rev RNA 11:e1576

Orecchini E, Frassinelli L, Galardi S et al (2018) Post-transcriptional regulation of LINE-1 retrotransposition by AID/APOBEC and ADAR deaminases. Chromosom Res 26:45–59

Petes C, Odoardi N, Gee K (2017) The toll for trafficking: toll-like receptor 7 delivery to the endosome. Front Immunol 8:1075

Pfaller CK, Radeke MJ, Cattaneo R et al (2014) Measles virus C protein impairs production of defective copyback double-stranded viral RNA and activation of protein kinase R. J Virol 88:456–468

Pfaller CK, Mastorakos GM, Matchett WE et al (2015) Measles virus defective interfering RNAs are generated frequently and early in the absence of C protein and can be destabilized by adenosine deaminase acting on RNA-1-like hypermutations. J Virol 89:7735–7747

Pfaller CK, Donohue RC, Nersisyan S et al (2018) Extensive editing of cellular and viral double-stranded RNA structures accounts for innate immunity suppression and the proviral activity of ADAR1p150. PLoS Biol 16:e2006577

Ping XL, Sun BF, Wang L et al (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res 24:177–189

Piontkivska H, Frederick M, Miyamoto MM et al (2017) RNA editing by the host ADAR system affects the molecular evolution of the Zika virus. Ecol Evol 7:4475–4485

Rima BK, Gatherer D, Young DF et al (2014) Stability of the parainfluenza virus 5 genome revealed by deep sequencing of strains isolated from different hosts and following passage in cell culture. J Virol 88:3826–3836

Ringeard M, Marchand V, Decroly E et al (2019) FTSJ3 is an RNA 2′-O-methyltransferase recruited by HIV to avoid innate immune sensing. Nature 565:500–504

Rodriguez-Hernandez A, Spears JL, Gaston KW et al (2013) Structural and mechanistic basis for enhanced translational efficiency by 2-thiouridine at the tRNA anticodon wobble position. J Mol Biol 425:3888–3906

Samuel CE (2011) Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. Virology 411:180–193

Scholler E, Weichmann F, Treiber T et al (2018) Interactions, localization, and phosphorylation of the m(6)A generating METTL3-METTL14-WTAP complex. RNA 24:499–512

Schwartz S, Bernstein DA, Mumbach MR et al (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell 159:148–162

Shi H, Wang X, Lu Z et al (2017) YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. Cell Res 27:315–328

Shi H, Wei J, He C (2019) Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers. Mol Cell 74:640–650

Simon V, Bloch N, Landau NR (2015) Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. Nat Immunol 16:546–553

Slotkin W, Nishikura K (2013) Adenosine-to-inosine RNA editing and human disease. Genome Med 5:105

Squires JE, Patel HR, Nousch M et al (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40:5023–5033

Tan B, Gao SJ (2018) RNA epitranscriptomics: regulation of infection of RNA and DNA viruses by N6-methyladenosine (m^6A). Rev Med Virol 28:e1983

Tirumuru N, Zhao BS, Lu W et al (2016) N6-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. elife 5:e15528

van den Hoogen BG, van Boheemen S, de Rijck J et al (2014) Excessive production and extreme editing of human metapneumovirus defective interfering RNA is associated with type I IFN induction. J Gen Virol 95:1625–1633

Viswanathan T, Arya S, Chan SH et al (2020) Structural basis of RNA cap modification by SARS-CoV-2. Nat Commun 11:3718

Xing J, Yi J, Cai X et al (2015) NSun2 promotes cell growth via elevating cyclin-dependent kinase 1 translation. Mol Cell Biol 35:4043–4052

Yang X, Yang Y, Sun BF et al (2017) 5-Methylcytosine promotes mRNA export – NSUN2 as the methyltransferase and ALYREF as an m^5C reader. Cell Res 27:606–625

Zhao BS, Roundtree IA, He C (2017) Post-transcriptional gene regulation by mRNA modifications. Nat Rev Mol Cell Biol 18:31–42

Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29

Züst R, Cervantes-Barragan L, Habjan M et al (2011) Ribose 2′-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. Nat Immunol 12:137–143

Epitranscriptomic Modifications and How to Find Them

Megan L. Van Horn and Anna M. Kietrys

Contents

Abstract Recently, there has been significant progress made in regard to chemical protocols for the detection and investigation of epitranscriptomic modifications, such as new breakthroughs in high-throughput sequencing methods, base-specific chemistries, and spectroscopy. Herein, we describe the development in methodology for probing epitranscriptomic modifications. We characterize the prevalent RNA modifications and the most important breakthroughs in epitranscriptomics. Further, a summary of the available approaches for detection is presented, with a strong focus on the newest methodology for each modification. We characterize analytical methods for the following modifications: N6-methyladenosine (m^6A), 1-methyladenosine (m^1A), 5-methylcytidine (m^5C), 5-hydroxymethylcytidine (hm^5C), 5-formylcytidine (f^5C), 5-carboxycytidine (ca^5C), inosine (I), pseudouridine (Ψ), and 2′-*O*-methylation (Nm). These are framed in the context of mRNA and other coding and non-coding RNAs. These epitranscriptomic modifications often determine the structure, life span, and function of RNAs, which are major regulatory molecules of cellular biology.

M. L. Van Horn · A. M. Kietrys (✉)
Department of Chemistry, Carnegie Mellon University, Pittsburgh, PA, USA
e-mail: mvanhorn@andrew.cmu.edu; amkietrys@cmu.edu

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_6

Keywords Epitranscriptomics · RNA modifications · N6-methyladenosine (m^6A) · 1-methyladenosine (m^1A) · 5-methylcytidine (m^5C) · Inosine (I) · Pseudouridine (Ψ) · 2′-*O*-methylation (Nm) · RNA sequencing · Methods for epitranscriptomics

1 Introduction

Over the past few years, the pool of knowledge available regarding modifications of RNA, called the epitranscriptome, has only continued to expand. These modifications, generally but not limited to the purine or pyrimidine base, may structurally alter the pairing of nucleobases, cause structural rearrangements of RNA, and therefore regulate the function of the molecule. Epitranscriptomic modifications determine the multifunctional nature of the RNAs and the vast biological processes they regulate, including RNA splicing (Rueter et al. 1999), translation (Meyer et al. 2015), cellular localization (Liu et al. 2020), and lifespan (Wang et al. 2014). Decoding information regarding modifications that effect changes in either sequencing or structure is increasingly important (Decatur and Fournier 2002; Motorin and Helm 2010). Modifications in RNA are not limited to messenger RNA (Roundtree et al. 2017), but also occur in transfer RNA (Motorin and Helm 2010), ribosomal RNA (Decatur and Fournier 2002), and non-coding RNAs (Hussain et al. 2013). At this time, there are over 150 distinct RNA modifications. This review focuses on a set of well-known, biologically-active modifications, consisting of N6-methyladenosine (m^6A), 1-methyladenosine (m^1A), 5-methylcytidine (m^5C), 5-hydroxymethylcytidine (hm^5C), 5-formylcytidine (f^5C), 5-carboxycytidine (ca^5C), inosine (I), pseudouridine (Ψ), and 2′-*O*-methylation (Nm).

Previously, RNA modifications were detected through protocols utilizing a combination of chemical or enzymatic digestion, radiolabeling, thin-layer chromatography, and mass spectrometry. To begin, enzymatic and chemical digestion methods were developed, with enzymatic digestion involving a combination of ribonuclease digestion and ionophoretic separation for mapping of modifications to specific RNAs. The digestion products of modified RNAs migrate differently when compared to unmodified digestion products. Subsequent sequencing of the determined modified RNAs was then used to pinpoint the specific location of the modification within the surrounding sequence. Radioactive labeling was often incorporated to improve the sensitivity of the analysis (Sanger et al. 1965). Chemical digestion and other chemical treatments became more prominent with the advent of reverse transcriptase, as chemical modifications installed on an epimodification cause steric hindrance which stops reverse transcriptase. The reverse transcriptase stops show a specific pattern of truncations or misincorporations within RNA sequencing results that correlates to modified nucleosides.

RNA modifications were also characterized by paper chromatography and thin-layer chromatography, which are based on the distance traveled by an analyte in a mobile phase drawn up a stationary phase via capillary action. The main difference

between these two methods lies in the choice of the stationary and mobile phase. Thin-layer chromatography generally uses a layer of an adsorbent such as silica gel or cellulose as a stationary phase, while paper chromatography relies on the cellulose molecules in the paper. These techniques both use the differences in net charge, polarity, and hydrophobicity between nucleotides to separate them from one another on the stationary phase of the chromatography plate. As a result, identification of nucleotides and nucleosides can be done through comparison of their mobility against a known standard. When differentiating modified RNA bases, thin-layer chromatography protocols utilizing two-dimensional separation on cellulose plates have provided the greatest separation capability. As with enzymatic and chemical digestion, radiolabeling can also be used in conjunction with thin-layer chromatography to further increase the sensitivity of detection (Randerath 1965).

Once mass spectrometry was developed, it was also used to elucidate the atomic structure of unknown nucleosides. Determination of the molecular structure allowed researchers to identify suitable methods for further interrogation, including thin-layer chromatography or high-performance liquid chromatography. Mass spectrometry relies on measuring characteristics of the fragmented nucleoside, with each fragment compared to a known standard to identify composition before being placed within the molecular structural context of the complete base. Although the structure of the nucleoside was able to be characterized, mass spectrometry alone was unable to place the modified RNA base within the context of a particular sequence or region (Thomas and Akoulitchev 2006).

While these historical methods are able to detect RNA modifications, specific level of the identification lacks the depth for precise mapping of the locations or determination of the presence of less abundant modifications. Additionally, many of these techniques required large quantities of RNA in order to be digested with various RNase enzymes, which may be difficult to acquire in certain studies. In comparison, the more recent techniques that will be discussed in this review improve upon the historical methods to provide a site-specific detection and quantitation. Overall, these newer strategies take one or more of the following approaches: specific antibody-based enrichment, modification-driven enzymes, specific chemical labeling of the modification, or the use of unique base-pairing features of the modification. In this review, we will summarize the most characterized modifications of RNA, their structural nuances, biological progenitors, and proposed functions, and describe their various detection methods focusing on high-throughput analyses.

Currently, the majority of the methods for the detection of RNA modifications can be grouped into three categories: high-throughput transcriptome-wide sequencing techniques (Table 1), mass spectrometry methods (Table 2), and bioinformatics tools and pipelines. While the advent of next-generation sequencing has provided a glut of sequencing information within a single run, it requires advanced algorithms to process and interpret this information. For increased sensitivity and more comprehensive data, next-generation sequencing has been coupled with immunoprecipitation (IP) and base-specific chemistries (Dominissini et al. 2012; Sakurai et al. 2014). Beyond various high-throughput sequencing techniques, mass spectrometry

(MS) methods have also been developed for the identification of nucleic acid modifications. MS techniques have the advantage of extreme sensitivity to base modifications, and many mass spectrometry protocols are able to detect multiple modifications within one sample of RNA (Jinno et al. 2017; Tardu et al. 2019). Lastly, while bioinformatics tools are not the main focus of this review, there is a wide variety of published pipelines for not only locating RNA modifications but also uncovering each modifications' distinct distribution patterns throughout the transcriptome. Bioinformatics methods, in particular, are currently an area of intense development, with many research groups focused on producing new pipelines and tools for RNA modification recognition and analysis (Chen et al. 2018; Sun et al. 2019; Werner et al. 2020).

2 Modifications

2.1 N6-methyladenosine (m^6A)

The m^6A modification is formed by methylation at the N6 position of adenine (Fig. 1). It occurs at an estimated frequency of 0.1–0.6% m^6A/A. The majority of m^6A modifications occur preferentially near stop codons and 3′ untranslated regions (UTRs), specifically in the consensus motif RRm^6ACH (R = A, G; H = A, C, U) (Dominissini et al. 2012; Meyer et al. 2012). Significant enrichment of m^6A is particularly seen just upstream of the stop codon. m^6A was discovered to be a reversible RNA modification, with groups of enzymes that function as "writers" and "erasers." A methyltransferase complex including METTL3, METTL14, and WTAP is responsible for catalyzing the modification, and the fat mass and obesity-associated protein (FTO) and AlkB Homolog 5, RNA Demethylase (ALKBH5) are known to actively perform m^6A demethylation (Fig. 1) (Jia et al. 2012; Zheng et al. 2013; Liu et al. 2014; Ping et al. 2014). Examination of the binding sites of the METTL complex has shown that, beyond the preference for coding sequences and 3′ UTRs, a significant portion of binding sites are within intronic sequences. This suggests that m^6A could be installed co-transcriptionally, potentially in nuclear speckles. Additionally, the METTL complex has shown to be rapidly recruited to DNA damaged by UV irradiation to mediate local RNA m^6A methylation, which facilitates recruitment of DNA damage repair polymerase κ. This local methylation can be reversed both rapidly and easily by FTO. Regarding the removal of m^6A, both

Fig. 1 Structure of adenine nucleobase and cellular machinery for the formation of the m^6A modification in RNA

FTO and ALKBH5 oxidatively demethylate m^6A, with FTO having also been seen removing m^6Am adjacent to the mRNA 5′ cap and internal m^6A. ALKBH5 demethylase activity has been linked to mRNA nuclear export and spermatogenesis (Roundtree et al. 2017). A third group of m^6A-specific proteins, termed "readers," have also been discovered to regulate splicing and cellular processing events upon recognition of m^6A-containing mRNAs (Alarcón et al. 2015; Xiao et al. 2016). Moreover, the structural change caused by the m^6A modification has been proposed to alter the accessibility of the RNA sequence to various RNA-binding proteins (Liu et al. 2015). This results from a shift in the conformation of the adenine base in its bound versus unbound state. When unpaired, the methyl group adopts the *syn* conformation, whereas the *anti* conformation is preferred for Watson–Crick binding to uracil. However, the *anti* conformation induces a steric clash between the m^6A methyl group and the N7 in the purine ring and results in destabilization of the m^6A-U pairing, further precipitating changes in local secondary structure and biological function (Roost et al. 2015).

m^6A-modified areas of unstructured transcripts can allow for the recognition of these regions by various proteins. Particularly, the YTH domain-containing family of proteins has a high binding affinity to m^6A in the RRm^6ACH motif, determining its recognition and modulating activity (Dominissini et al. 2012). YTHDF1 and YTHDF3 mediate translation through interacting with common ribosomal proteins, while the direct association with YTHDF2 works to decay mRNA targets. Another protein group commonly associated with m^6A, the heterogeneous nuclear ribonucleoprotein (HNRNP) family, functions to regulate the processing of m^6A-containing transcripts. The destabilization of m^6A-U and subsequent unstructured regions provides an opportunity for RNA transcripts to be recognized by HNRNPC and HNRNPG, which bind m^6A-dependent structural switches. As a result of these binding interactions, HNRNPC and HNRNPG are able to mediate canonical splicing outcomes on these methylated transcripts (Liu et al. 2015; Roundtree et al. 2017). Consequently, this modification can be shown to modify many biological functions, such as translation (Zhou et al. 2015; Meyer et al. 2015), splicing (Dominissini et al. 2012), and structure regulation (Roost et al. 2015), as well as the lifespan of the RNA (Wang et al. 2014).

The first high-throughput methods for transcriptome-wide detection of m^6A were developed independently by two different groups in 2012. Both methods, m^6A-seq and m^6A-containing RNA immunoprecipitation (MeRIP-seq), enrich m^6A-specific methylated RNA by immunoprecipitation with an m^6A-specific antibody following an initial mRNA fragmentation step (Fig. 2) (Dominissini et al. 2012; Meyer et al. 2012). Immunoprecipitation (IP) allows for the pulling of a targeted molecule out of solution using an antibody with a high affinity to that molecule. RIP, or RNA immunoprecipitation, is a variant of IP that targets RNA modification instead of proteins (Li et al. 2013). Once the target RNA has been enriched, the procedure continues along the traditional RNA sequencing (RNA-seq) methodology (Mortazavi et al. 2008). RNA-seq begins from the isolation of target RNA, followed by RNA selection or depletion, and then the RNA fragmentation step. The majority of the methods included in this review are based on RNA selection towards specific

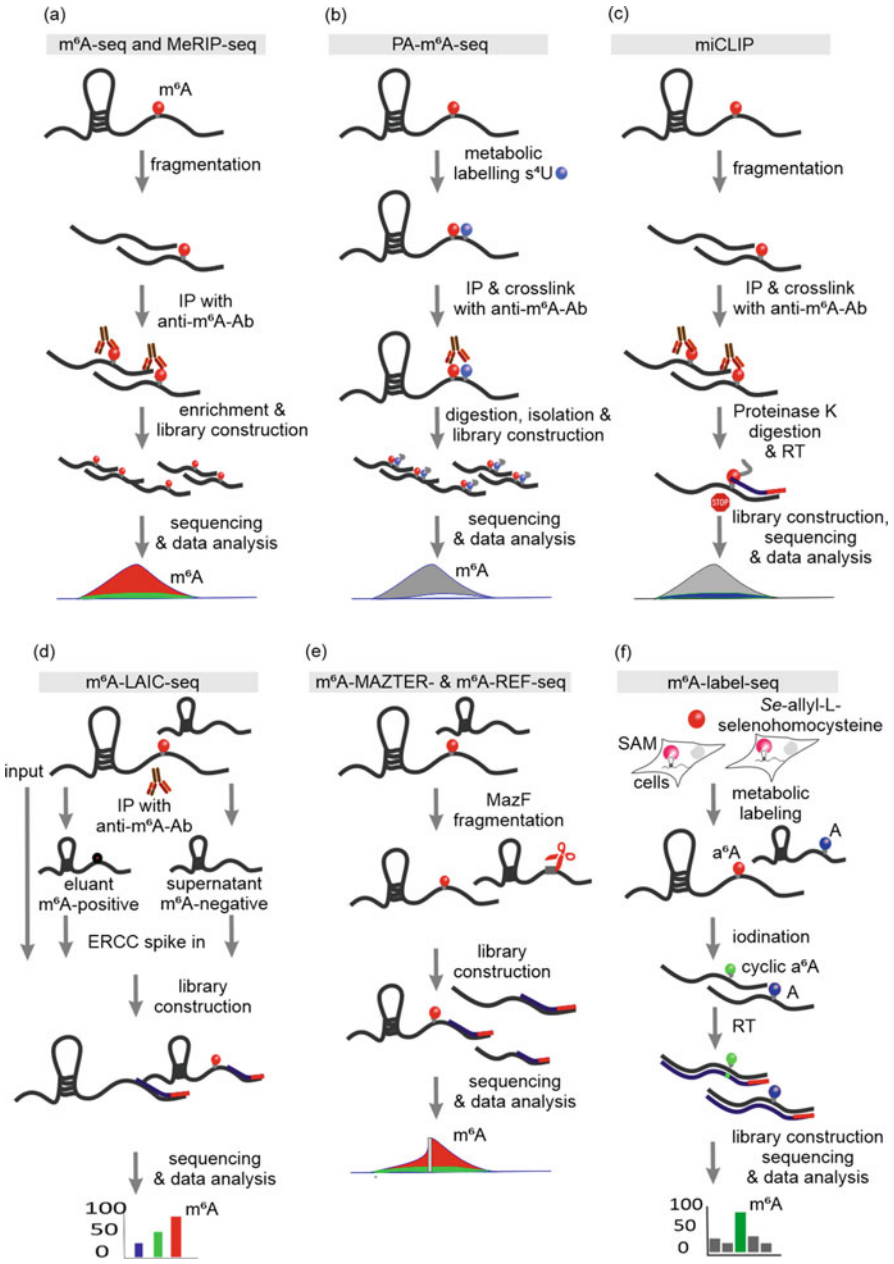


Fig. 2 High-throughput detection strategies for detection of m^6A modification in RNA. (**a**) m^6A-seq/MeRIP-seq identifies sites of m^6A modification using immunoprecipitation with an m^6A-specific antibody. (**b**) PA-m^6A-seq integrates s^4U metabolically into target RNA to induce a T-to-C shift. The m^6A site is identified with an m^6A-specific antibody and cross-linked. The location of the T-C shift indicates sites of m^6A methylation. (**c**) miCLIP combines

sequential motifs or epitranscriptomic modifications. For m^6A-seq/MeRIP-seq, enrichment of RNA fragments containing m^6A lowers the detection limit of modified RNAs within samples with potentially less abundant modification. The next step of fragmentation can be done with enzymes, divalent ions, sonication, or nebulizers and reduces bias by randomization of the fragmented sequence. These fragments are then ligated with adapter oligonucleotides. These adapters allow for primer binding and mark locations from which reverse transcriptase (RTase) can begin the process of reverse transcription (RT) to produce a library of complementary DNA (cDNA). Reverse transcription from RNA to cDNA takes advantage of the increased stability of DNA and allows for further amplification with DNA polymerases and the use of high-throughput sequencing technologies. The cDNA library is then amplified via polymerase chain reaction (PCR) to increase the amount of material available for sequencing. Once the library has been constructed, size selection is required to remove any sequences that are not the intended targets of sequencing, including incomplete and improper sequences, excess primers, or primer dimers. High-throughput sequencing of the libraries constructed from m^6A-RNAs enriched by m^6A-seq/MeRIP-seq confirmed the tendency for m^6A to occur in 3′ untranslated regions (UTRs) and near stop codons.

Consequently, the majority of m^6A recognition tools have their basis in m^6A-seq and MeRIP-seq. These include well-known strategies such as PA-m^6A-seq and miCLIP (Chen et al. 2015; Linder et al. 2015)/m^6A-CLIP (Hsu and He 2019), and m^6A-LAIC-seq (Molinie et al. 2016). PA-m^6A-seq and miCLIP/m^6A-CLIP methods rely on cross-linking immunoprecipitation (CLIP) variants, photoactivatable ribonucleoside-enhanced CLIP (PAR-CLIP) (Danan et al. 2016) and UV CLIP (Ke et al. 2015) (Fig. 2). Cross-linking immunoprecipitation begins with in vivo cross-linking of the RNA–protein or RNA-antibody complexes with UV irradiation to form covalent bonds. In photo-cross-linking-assisted m^6A sequencing (PA-m^6A-seq), 4-thiouridine (s^4U), incorporated via metabolic labeling, induces a T-to-C mutation at a site of cross-linking. Once m^6A immunoprecipitation is completed, the selected m^6A-RNAs are cross-linked to the specific antibody under UV light, digested with RNase T1, which cleaves after guanines, and sequenced. As a

Fig. 2 (continued) immunoprecipitation with an m^6A-specific antibody with cross-linking to determine m^6A sites. Proteinase K retrieval of cross-linked RNA produces truncations or misincorporations in RT indicative of m^6A methylation. (**d**) m^6A-LAIC-seq separates immunoprecipitation products into a control supernatant (m^6A-negative fraction) and eluate (m^6A-positive fraction) group of RNA. ERCC internal standards are added to these two pools and into the input RNA. The amount of m^6A per gene is quantified by the ratio of RNA abundance between pools. (**e**) m^6A-REF-seq/MAZTER-seq uses the MazF enzyme to cleave RNA at unmethylated ACA motifs. m^6A-containing motifs are not cut and remain complete in the sequencing readout. (**f**) m6A-label-seq utilizes Se-allyl-L-selenohomocysteine to substitute the methyl group on SAM with an allyl, which inserts a^6A in the place of m^6A. a^6A sites undergo iodination-induced cyclization that cause misincorporation in the cDNA library and the mutation patterns in RNA-seq data

result of the T-to-C shift, sites of m^{6}A methylation can be detected at a single-base resolution (Chen et al. 2015).

The other UV-based cross-linking method, cross-linking immunoprecipitation (UV CLIP) (Ke et al. 2015), is used in both m^{6}A-CLIP (Hsu and He 2019) and m^{6}A individual nucleotide resolution cross-linking and immunoprecipitation (miCLIP) (Linder et al. 2015). As before, RNA fragments are immunoprecipitated and then cross-linked at the m^{6}A methylation site to the corresponding antibody with UV light. However, where m^{6}A-CLIP and miCLIP diverge from m^{6}A-seq is that they use Proteinase K to retrieve the cross-linked RNA. The protein-modified RNA interaction causes either truncation or misincorporations in RT that can be read as m^{6}A patterns in the sequencing readouts. In the situation of co-existent m^{6}A and Nm modifications (m^{6}Am), m^{6}Am-Exo-seq can be applied (Sendinc et al. 2019).

While the methods listed above target single-base resolution methylation events, they are not able to fully quantify the level of m^{6}A modification in individual transcripts. A method, developed in 2013 and called site-specific cleavage and radioactive-labeling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET), can quantify m^{6}A stoichiometry at specific loci. SCARLET has the ability to identify both the precise location of the modification as well as the fraction of transcripts containing the modification at that specific site but is limited to specific loci (Liu et al. 2013).

For a transcriptome-wide m^{6}A stoichiometric profile, m^{6}A-level and isoform-characterization sequencing (m^{6}A-LAIC-seq) was developed (Fig. 2). This method uses full-length rather than fragmented RNAs in the m^{6}A immunoprecipitation. Additionally, it incorporates External RNA Controls Consortium (ERCC) internal standards into the input, control supernatant (m^{6}A-negative fraction), and eluate (m^{6}A-positive fraction) pools of RNA. Determination of m^{6}A levels per transcript is then quantified by the ratio of RNA abundance in the different pools. These abundances are collected from the RNA-seq fragment counts that have been normalized across the input, supernatant, and eluate using the ERCC control RNAs (Molinie et al. 2016).

Another technique depending on antibody recognition, m^{6}A-Cross-linking-Exonuclease-sequencing (m^{6}ACE-seq) is one of the most recent antibody-based methods for m^{6}A identification, developed for quantitative single-base resolution sequencing of methylation sites. The m^{6}A-specific antibodies are first photo-cross-linked with the target RNAs, which are then protected from subsequent 5′-to-3′ exoribonuclease (XRN1) digestion. Sequencing of the remaining fragments should therefore reveal the m^{6}A sites at the nucleotide located on the 5′ end (Koh et al. 2019).

All previously presented methods for m^{6}A methylation detection have relied on m^{6}A-specific antibodies for the enrichment of RNA transcripts with low abundance. However, m^{6}A-specific antibodies may also recognize modifications similar to m^{6}A, such as m^{6}Am or N6,N6-dimethyladenosine. Co-recognition with variants may cause many false positive assignments of m^{6}A. Within the past few years, a new group of antibody-independent m^{6}A sequencing techniques has been published: m^{6}A-REF-seq/MAZTER-seq, DART-seq, m^{6}A-label-seq, and m^{6}A-SEAL-seq.

Two independent groups reported m^{6}A-sensitive RNA-Endoribonuclease-Facilitated sequencing (m^{6}A-REF-seq) and RNA digestion via m^{6}A-sensitive RNase (MAZTER-seq) as antibody-independent m^{6}A sequencing methods, instead of using endoribonuclease-based strategies for modification detection (Fig. 2). The enzyme of choice is an m^{6}A-sensitive RNase, MazF, which cuts RNA only at unmethylated ACA motifs, not m^{6}ACA motifs. Sequencing then reveals m^{6}A methylation sites where there are complete, uncleaved ACA motifs in the readout (Zhang et al. 2019b; Garcia-Campos et al. 2019).

Deamination adjacent to RNA modification targets (DART-seq) is another recently developed antibody-free method for m^{6}A detection that takes advantage of the consensus sequence surrounding sites of methylation. The preferred consensus sequence RRm^{6}ACH invariably contains a cytidine residue immediately following the methylation. Therefore, researchers fused the cytidine deaminase APOBEC1, which induces cytosine-to-uracil conversion, with the m^{6}A-binding YTH domain and expressed this construct in cells of interest. APOBEC1-YTH recruitment to m^{6}A sites then causes the deamination of the proximal cytidine. After sequencing, instances of C-to-U mutations are then used to identify m^{6}A sites at the adenosine immediately prior to the mutation (Meyer 2019).

The most recently published antibody-free methods, m^{6}A-label-seq and m^{6}A-SEAL-seq, utilize a metabolic labeling and an FTO-assisted chemical labeling approach, respectively. The initial m^{6}A biogenesis involves both m^{6}A methylation enzymes and the cofactor SAM (S-adenosyl methionine). To co-opt this biogenesis process in m^{6}A-label-seq, cells are fed with a methionine analog *Se*-allyl-L-selenohomocysteine to substitute the methyl group on SAM with an allyl, which produces metabolically-modified RNAs containing N6-allyladenosine (a^{6}A) in the place of m^{6}A. These a^{6}A sites are then detected based on a^{6}A iodination-induced cyclization misincorporation at the opposite site in the cDNA library and the subsequent mutation pattern in RNA-seq data (Shu et al. 2020).

The chemical labeling technique of m^{6}A-SEAL-seq (an FTO-assisted m^{6}A selective chemical labeling method) also takes advantage of inherent biological machinery through the use of FTO, which enzymatically oxidates m^{6}A to N6-hydroxymethyladenosine (hm^{6}A). Then, treatment with dithiothreitol (DTT) mediates a thiol-addition reaction that converts the unstable hm^{6}A to the more stable N6-dithiolsitolmethyladenosine (dm^{6}A), which has a free sulfhydryl group. This group can be exploited for conjugation with a variety of tags, including biotin, via reaction with methanethiosulfonate (MTSEA), then allowing for affinity purification. Further steps follow regular RNA sequencing protocol (Wang et al. 2020).

Not based on RNA sequencing, SELECT, a single-base elongation- and ligation-based qPCR amplification method, can also be used for m^{6}A site identification. SELECT consists of two selection steps, with the first having m^{6}A hinder the ability of the DNA polymerase to continue elongation and add a thymine opposite the m^{6}A site. The second step has the m^{6}A sites in the RNA template selectively prohibit DNA-ligase-catalyzed nick ligation between the two PCR probes. After these two rounds of selection, the amount of ligation products formed from methylated RNA templates will be sharply reduced when compared to products from unmodified

RNA templates. Quantification of the modified versus unmodified templates can then be done using qPCR. SELECT can be used for novel identification of m^6A sites, taking advantage of m^6A's hindrance of both the single-base elongation activity of DNA polymerases and the nick efficiency of ligases. FTO-assisted SELECT was also shown to enable clear identification of m^6A target sites in both low-abundance lncRNA and mRNA (Xiao et al. 2018).

Due to the structure of m^6A, it is considered an "RT-silent" modification, meaning that it is chemically stable and, therefore, unable to introduce mutations or truncations during reverse transcription. Using a Se-modified deoxythymidine triphosphate (4SedTTP*) allows for efficient A-T* pairing but not m^6A-T* pairing during cDNA synthesis. Combining 4SedTTP* with FTO creates a strategy for RT-based detection of m^6A through the truncation signatures of cDNA transcripts upon encountering the treated m^6A sites (Hong et al. 2018).

Recently, a variety of electrochemical immunosensors have also been published for the detection of low levels of m^6A modification within an RNA sample, although they lack site-specificity. These biosensors are based on a variety of nanoparticles, including gold nanoparticles (AuNPs), SiO_2 nanospheres, and PtCo mesoporous nanospheres, where incidences of the detection result in electrochemical reduction signals (Yin et al. 2017; Dai et al. 2018; Ou et al. 2020).

2.2 *1-methyladenosine (m^1A)*

Methylation at the N1 position of the adenine (Fig. 3) is about ten times less abundant than m^6A at an estimated frequency of 0.015–0.054% m^1A/A (Dominissini et al. 2016). This modification occurs mainly in the GC-rich regions at 5′ UTRs and has been identified in tRNA, rRNA, and most recently in mRNA. It also has an associated set of "writers" and "erasers," suggesting that adenine modifications are highly dynamic with regulatory functionality (Dominissini et al. 2016; Li et al. 2016b, 2017; Safra et al. 2017). TRMT6 and the TRMT61A complex or TRMT61B are responsible for the installation of m^1A in cytoplasmic and mitochondrial tRNA, respectively. The TRMT6-TRMT61A complex recognizes the consensus sequence GUUCRA (R = A, G), as well as the characteristic T-loop structure of tRNA (Li et al. 2017; Safra et al. 2017). For human rRNA, methylation

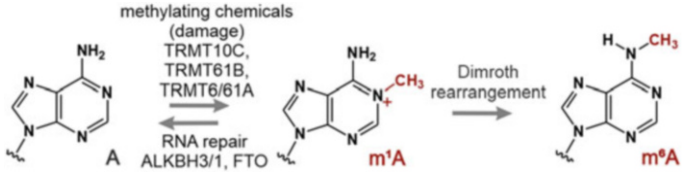


Fig. 3 Structure of adenine nucleobase and cellular machinery for the formation of the m^1A modification in RNA

occurs via RRP8 and has been determined necessary for proper rRNA biogenesis (Hauenschild et al. 2015; Waku et al. 2016). Removal of m^1A can be catalyzed by ALKBH1, ALKBH3, and FTO (Fig. 3) (Liu et al. 2016; Dominissini et al. 2016; Li et al. 2016b; Wei et al. 2018).

Upon methylation, the adenine acquires both a positive charge and a methyl group that blocks canonical Watson–Crick pairings. Both of these features also encourage strong electrostatic RNA-protein interactions and alternative RNA secondary structure motifs (Harcourt et al. 2017). As a result of the hindered Watson–Crick pairing, m^1A destabilizes the surrounding duplex as it remains unpaired instead of participating in Hoogsteen base-pairing (Yang and Lam 2009; Zhou et al. 2016). In the case of loop structures, the introduced positive charge could stabilize interactions with the RNA phosphate backbone. Additionally, the m^1A modification can be subjected to the Dimroth rearrangement, wherein under alkaline conditions, it is converted to an m^6A (Fig. 3). The Dimroth rearrangement facilitates the switching of the endocyclic and exocyclic nitrogen atoms, which shifts the methylation from the N1 to the N6 position. This then effectively converts the m^1A to m^6A. The presence of m^1A in RNA induces termination or misincorporation during reverse transcription as a result of the methyl group protruding from the Watson–Crick edge of adenosine, allowing for identification of modifications at single-base resolution (Kietrys and Kool 2016).

The positioning of the m^1A modification near the translation start site and first splice site in coding transcripts correlates with modulation of translation. Potentially resulting from its structure and subsequent poor pairing abilities, m^1A has been suggested to affect translation by causing a change in RNA folding to allow access to a previously paired region of RNA. It has also been correlated with increases in translation and changes in RNA cellular metabolism (Dominissini et al. 2016; Li et al. 2016b, 2017; Safra et al. 2017).

Identification of m^1A is similar to m^6A, as the most common method for detection is through the use of a commercial antibody. Various m^1A-seq methods have been developed for transcriptome-wide mapping, including m^1A-ID-seq, m^1A-seq, ARM-seq, DM-tRNA-seq, m^1A-MAP, m^1A-seq-TGIRT, and m^1A-miCLIP.

In 2016, two research groups independently published methods for the first transcriptome-wide m^1A landscape, m^1A-ID-seq (Dominissini et al. 2016) and m^1A-seq (Fig. 4) (Li et al. 2016b). Both rely on an m^1A-specific antibody for the enrichment of modified RNA, but differ slightly in the subsequent steps before final sequencing. The m^1A-ID-seq method utilizes an RNA/DNA demethylase for conversion of m^1A into canonical A following immunoprecipitation, while m^1A-seq chooses to use a Dimroth rearrangement for conversion of m^1A to the RT-silent m^6A. Sites of m^1A modification can be pinpointed by locating the truncated transcripts in the sequencing outputs when comparing treated versus untreated RNA templates (Dominissini et al. 2016; Li et al. 2016b).

ARM-seq and DM-tRNA-seq look to remove the m^1A base methylations that cause RT stops and subsequent truncations of the cDNA. AlkB-facilitated RNA methylation sequencing (ARM-seq) treats the target RNA with AlkB, a dealkylating enzyme that removes m^1A modification, thus removing RT hard stops and preventing truncation (Cozen et al. 2015). Similarly, demethylase-thermostable

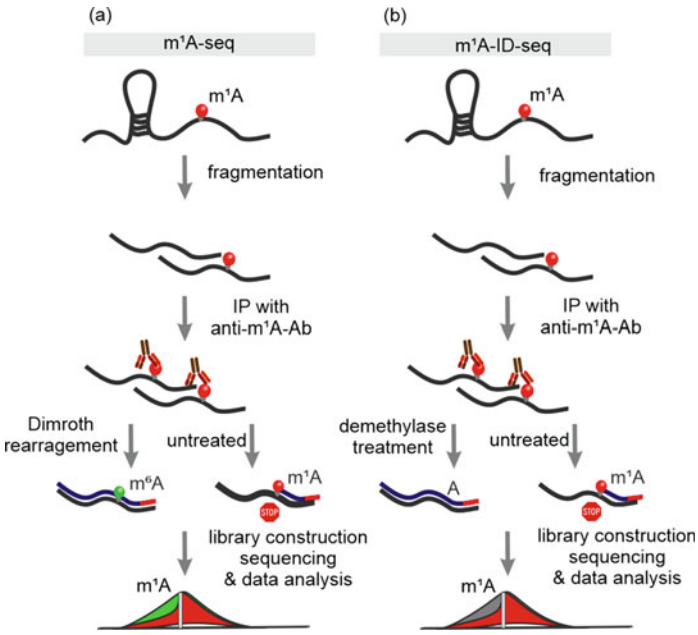


Fig. 4 High-throughput detection strategies for detection of m^1A modification in RNA. (**a**) m^1A-seq isolates m^1A-containing RNA fragments via immunoprecipitation with an m^1A-specific antibody. Conversion of m^1A to m^6A via Dimroth rearrangement allows complete read-through by RT. Comparison against the truncated m^1A-containing transcripts identifies m^1A methylation sites. (**b**) m^1A-ID-seq immunoprecipitates m^1A-containing RNA via m^1A-specific antibodies. A demethylase treatment removes m^1A sites and allows complete read-through by RT. Comparison against the truncated m^1A-containing transcripts identifies m^1A methylation sites

group II tRNA sequencing (DM-tRNA-seq), uses a demethylase mixture consisting of D135S AlkB and thermostable group II intron reverse transcriptase (TGIRT) to remove methylation sites and allow for full elongation. Again, locations of m^1A are determined through comparison of demethylase-treated and -untreated sequencing, which are now the locations of completed transcripts in the sequencing outputs (Zheng et al. 2015).

Continuing on the same trends for m^1A detection, misincorporation-assisted profiling of m^1A (m^1A-MAP) combines the use of an m^1A antibody for the enrichment of modified RNA fragments with a demethylase treatment to improve the likelihood of positive detection. However, m^1A-MAP differs from previous protocols in the choice of reverse transcriptase. Depending on the RTase used, both truncations and misincorporations during cDNA synthesis can be caused by an m^1A modification. Therefore, researchers chose TGIRT for precise induction of misincorporation at the m^1A modification site, as m^1A-induced truncations are less accurate (Li et al. 2017). The m^1A-seq-TGIRT method works similarly to m^1A-MAP, but, instead of demethylation, includes an optional Dimroth

rearrangement to transform m^1A to m^6A as a control. m^1A locations can then be identified by comparing the results from the sequencing of samples including and excluding the Dimroth rearrangement step (Safra et al. 2017).

One of the most recent developments in m^1A identification at single-base resolution is m^1A-modification individual nucleotide resolution cross-linking and immunoprecipitation, or m^1A-miCLIP. After an initial fragmentation of the RNA, an m^1A-specific antibody is cross-linked to the RNA with UV irradiation. After the RNA-antibody complexes are immunoprecipitated, Proteinase K induces the release of the RNA from the antibody, with the remaining peptide adducts on RNA. These adducts block full RTase read-through of the truncated RNA, thus producing cDNA with 3′ ends directly corresponding to the m^1A RNA site. Previously in other methods, any m^1A antibody-binding site located at a transcription start site would result in the reads terminating at the +1 position relative to the modification. The produced cDNA is then circularized to preserve the m^1A-induced truncations at the cDNA 3′ end. This final step is followed by amplification and sequencing, with locations of m^1A visible in the sequencing output as truncated transcripts (Grozhik et al. 2019).

In 2019, a CRISPR-Cas13a-based system for global m^1A detection within a sample predicated on the fact that m^1A cannot bind with its corresponding base pair has been published. As Cas13a has a single-base mismatch specificity, it does not cleave the reporter RNA on-target, leading to a low fluorescent signal which indicates the presence of m^1A within the target (Chen et al. 2019).

2.3 *5-methylcytosine (m^5C)*

Methylation of cytosine at the fifth position (Fig. 5) has been shown to be present in tRNA, rRNA, and mRNA. This modification is relatively common, with RNA sequencing showing more than 8000 m^5C sites in coding and non-coding mRNA regions and an estimated frequency of 0.025–0.095% m^5C/C (Huber et al. 2015). Subtle enrichment of m^5C is apparent in both the 5′ and 3′ UTRs (Squires et al. 2012). The distribution of the modified bases within this region further favors the

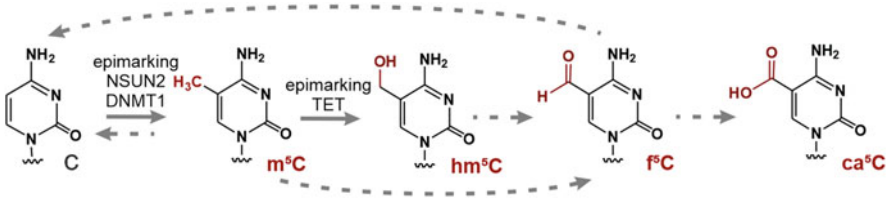


Fig. 5 Structure of cytosine nucleobase and cellular machinery for the formation of the m^5C, hm^5C, fm^5C, and ca^5C modifications in RNA. Both NSUN2 and DNMT1 catalyze methylation of C to m^5C. TET family enzymes have been proposed to further oxidize m^5C to hm^5C, f^5C, and ca^5C

binding sites for the Argonaute (AGO) proteins 1–4, which are involved in the RNA interference (RNAi) pathway that suppresses gene expression. A different set of methyltransferases, NSUN2 and DNMT2 (Fig. 5), have been designated as the main catalysts for m^5C methylation in mRNA and tRNA (Brzezicha et al. 2006; Jurkowski et al. 2008; Squires et al. 2012; Hussain et al. 2013; Yang et al. 2017).

While the m^5C modification does not interrupt base-pairing, it does increase the hydrophobicity of the major groove of RNA and potentially increases base stacking (Wang and Kool 1995). As a result, m^5C has stabilizing effects on the secondary structure in tRNA and also can affect translational fidelity in rRNA (Chow et al. 2007; Motorin and Helm 2010; Squires and Preiss 2010).

Unlike the previously presented modifications, there have been less conclusive suggestions on the role that m^5C plays within mRNA. Beyond the tendency of m^5C sites to congregate in the binding sites for AGO 1–4, there has also been a substantial association of m^5C sites with binding regions for central components of the miRNA/RNA-induced silencing complex (RISC). RISC functions to post-transcriptionally regulate levels of particular mRNA transcripts through translation inhibition or degradation of the target transcript. Degradation is carried out by members of the AGO family of proteins, which cleave the target mRNA. Consequently, since distribution patterns of m^5C have been implicated in these binding regions, it has been proposed that m^5C could play a role in post-transcriptional regulation of RNA metabolism (Squires et al. 2012). Variants of m^5C, such as hm^5C, f^5C, and ca^5C, have been proposed to affect the structure of the surrounding RNA through hydrogen bonding or polar interactions in the major groove. hm^5C, in particular, has been shown to be present in various transcripts involved in basic cellular processes and development (Delatte et al. 2016). The existence of hm^5C and additional variants dependent on further oxidation suggest that there is a potential avenue for m^5C dynamic reversibility, though evidence supporting such has not yet been presented. Additional profiling of this family of modifications is needed to more fully characterize their roles within the cell.

Stemming from its structure, m^5C is considered an RT-silent modification, as the enzyme may still incorporate the complementary deoxynucleoside triphosphate (dNTP) that would conform to Watson–Crick pairing as if it were an unmodified base. Therefore, detection of m^5C most commonly depends on chemical alteration of unmodified cytosines through sodium bisulfite treatment, which converts these cytosines into uracil (bisulfite-seq/BS-seq) (Fig. 6). However, m^5C is unaffected and will be read as cytosine in RNA sequencing. Therefore, a comparison of treated and untreated RNA transcripts can identify the presence of m^5C in the remaining cytosine sites (Schaefer et al. 2009; Squires et al. 2012). Variants of the bisulfite treatment have been reported, including m^5C-RIP, Aza-IP, and miCLIP (m^5C).

Similar to m^6A and m^1A detection methods, m^5C RNA immunoprecipitation (m^5C-RIP) utilizes an m^5C-specific antibody for the enrichment of fragmented RNA prior to cDNA library construction and sequencing (Fig. 6) (Edelheit et al. 2013). The other two approaches search for sites of m^5C modification by utilizing the ability of m^5C RNA methyltransferases to pinpoint such sites. The first, 5-aza-cytidine-mediated RNA immunoprecipitation (Aza-IP) (Fig. 6), incorporates this cytidine

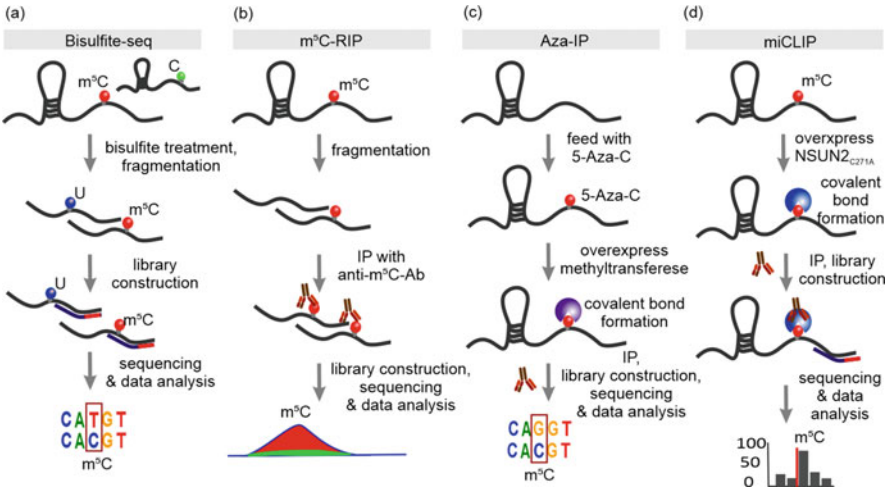


Fig. 6 High-throughput detection strategies for detection of m^5C modification in RNA. (**a**) Bisulfite-seq treats target RNA with sodium bisulfite, which converts unmodified C into U. Sites of m^5C are unaffected and remain to be read as C during sequencing. (**b**) m^5C-RIP utilizes an m^5C-specific antibody to immunoprecipitate m^5C-containing RNAs. (**c**) Aza-IP metabolically incorporates 5-Aza-C in RNAs also overexpressing an epitope-tagged m^5C RNA methyltransferase. A covalent bond is formed between the 5-Aza-C and methyltransferase, which allows the RNA to be immunoprecipitated. (**d**) miCLIP uses overexpression of $NSUN_{C271A}$ m^5C methyltransferase to form a covalent bond with m^5C sites. The protein–RNA complex is then immunoprecipitated

analog randomly into nascent RNAs that are overexpressing an epitope-tagged m^5C RNA methyltransferase. Substituting the analog interposes a nitrogen at the C5 position, which allows for a stable covalent bond to form between the RNA methyltransferase and the C6 position of the base. Once this bond coalesces, the targets are enriched by immunoprecipitation and sequenced. However, Aza-IP relies heavily on the complete replacement of m^5C sites with the cytosine analog, as any sites not replaced will be missed in the sequencing results (Khoddami and Cairns 2013).

The other method, methylation iCLIP (miCLIP) (Fig. 6), also employs the m^5C methyltransferase. The cysteine-to-alanine mutation (C271A) in human NSUN2 prevents the enzyme from releasing once it has formed a protein-RNA complex. The formation of a stable covalent bond allows for immunoprecipitation of the NSUN2 protein and its RNA targets, which presumptively contain m^5C methylation sites. These sites are determined through a pattern of reverse transcription stops, which terminates at the polypeptide-nucleotide crosslink site. Subsequent steps follow the regular RNA sequencing protocol (Hussain et al. 2013).

Further oxidation of m^5C by 10–11 translocation (TET) family enzymes creates 5-hydroxymethylcytidine (hm^5C) (Fu et al. 2014; Delatte et al. 2016). As opposed to the distribution of m^5C in the UTRs, hm^5C is predominately located in the coding sequences, which suggests differing functions for the hm^5C modifications (Delatte

et al. 2016). However, as methods for hm^5C discovery are still limited, these roles have not yet been elucidated. The hm^5C modification also protects cytosine from deamination to uracil when treated with the current bisulfite-seq technique that is also used for m^5C detection. BS-seq does not allow for differentiation between hm^5C and m^5C, only revealing that there is a cytosine modification present in the RNA transcript. In addition to hm^5C, another oxidation product of m^5C by TET family enzymes is 5-formylcytidine (f^5C) (Zhang et al. 2016), and it has been suggested that hm^5C can be further oxidized, forming 5-carboxycytidine (ca^5C) (Ito et al. 2011). Little is known about these m^5C derivatives, presumably due to the inherently transient nature of the modifications. There have been efforts to visualize these modifications through high-throughput sequencing and mass spectrometry.

For specific hm^5C detection, there are a few recently published protocols. An antibody-based method, hydroxymethylated RNA immunoprecipitation (hMeRIP-seq), uses an hm^5C-specific antibody to pull down fragments of RNA containing the modification to enrich the sequencing pool. Similar to previous RNA-seq methods, the pattern of the RT stops allows for the identification of modified bases (Delatte et al. 2016).

Recently, a peroxotungstate-based bisulfite-free analysis for both m^5C and hm^5C (TAWO-seq/WO-seq) has been proposed. Tet-assisted WO-seq (TAWO-seq) for m^5C identification originated from peroxotungstate oxidation sequencing (WO-seq), utilized for hm^5C identification. The WO-seq method does not rely on the traditional bisulfite-based enrichment scheme for m^5C detection. Rather, hm^5C-containing RNA is specifically oxidized by treatment with peroxotungstate, resulting in the conversion of hm^5C to trihydroxylated-thymine (^{th}T). This is then reverse transcribed into T by TGIRT during cDNA synthesis. This technique shows high specificity for hm^5C, as sequencing of peroxotungstate-treated canonical C-containing RNA and m^5C-containing RNA indicated no change in the number of sites versus a control. Additionally, treatment of RNA with peroxotungstate is a milder reaction than bisulfite, overall showing comparatively less damage on the treated RNA.

TAWO-seq is a variant of WO-seq, as it couples the WO-seq procedure with prior Naegleria Tet-like oxygenase (NgTET1) or mouse Tet1 (mTet1) for initial oxidation of m^5C to hm^5C (Yuan et al. 2019). Additionally, researchers showed that the original hm^5C sites within the RNA transcript could be protected from peroxotungstate oxidation by labeling the sites with glucose via β-glucosyltransferase (βGT). This then allows for specific m^5C identification without the potential confluence of hm^5C signals in the sequencing process. Canonical bisulfite sequencing has a tendency towards false positives, since the unmodified cytosine base is targeted. Through the development of a bisulfite-free system, the TAWO-seq/WO-seq methods could potentially decrease the number of false positives and present a more accurate picture of the m^5C/hm^5C landscape, as they directly detect modified cytosines (Yuan et al. 2019).

2.4 *Inosine (I)*

Characteristically referred to as "A-to-I editing," conversion of adenosine-to-inosine happens through hydrolytic deamination at the C6 position catalyzed by adenosine deaminases that act on RNA (ADARs) (Fig. 7) (Okada et al. 2019). Inosine modifications are common across metazoans, with an increased abundance in primates, including humans, over other animals (Paz-Yaacov et al. 2010). ADARs preferentially modify in dsRNA, frequently within *Alu* elements in untranslated regions and introns. *Alu* elements consist of short interspersed sequence fragments approximately 300 nucleotides in length that generally contain high amounts of CpG dinucleotides, making them targets for methylation. Additionally, while ADARs are expressed across a variety of mammalian tissues, the majority of ADAR targets are within brain tissue (Athanasiadis et al. 2004; Levanon et al. 2004). The conversion that ADARs catalyze shifts the base into a chemical structure similar to guanine, causing hypoxanthine (inosine base)-cytosine base-pairing (Nigita et al. 2015).

Changing base-pairing preference can cause ripples of change in other areas, specifically in amino acids produced from a modified coding sequence (Sommer et al. 1991). Alternately, switching pairing preference inherently changes the stability of the base-pairing, leading to potential effects on the local secondary structure along with coding and recognition (Serra 2004). This mainly appears in the form of a destabilizing I-U wobble pair, which has the tendency to cause the "unwinding" of double-stranded RNAs (Wagner et al. 1989; Serra 2004). A direct connection between these secondary structure changes and specific biological effects has yet to be made, although inosine has been broadly associated with codon alteration (Ohlson et al. 2007) and alternative splicing (Rueter et al. 1999; Sakurai et al. 2014).

Rather than using antibody-based detection methods, inosine in RNA requires taking advantage of these differing base-pairing properties between the modified and unmodified adenine base or through a chemical labeling strategy. For the former, once the inosine site has undergone reverse transcription, it can be read as guanine which then shows as an A-to-G mutation in cDNA. Comparison of the RNA sequencing results against the genomic sequencing results for a particular sample then reveals the locations of inosine modification (Athanasiadis et al. 2004; Levanon et al. 2004; Li et al. 2009). However, A-to-G mutations can also arise from single-nucleotide polymorphisms or database errors, making this method not wholly reliable for inosine site identification.

The main method for chemical labeling relies on treatment with acrylonitrile to generate inosine-specific cyanoethylation, producing strong reverse transcription stops (ICE-seq) (Fig. 8). The locations of A-to-I editing can then be pinpointed to

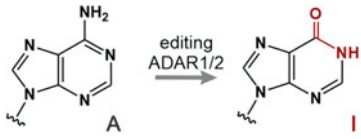


Fig. 7 Structure of adenine nucleobase and cellular machinery for the formation of inosine in RNA

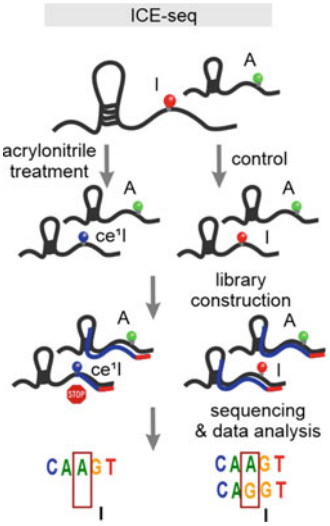


Fig. 8 High-throughput detection strategies for detection of inosine modification in RNA. ICE-seq treats the target RNA with acrylonitrile to produce cyanoethylated inosine. The altered inosine site produces truncated transcripts during RT. Comparison with untreated RNA transcripts identifies inosine sites

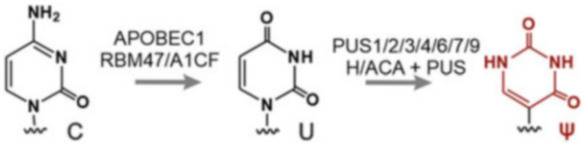


Fig. 9 Structure of uracil nucleobase and cellular machinery for the formation of the Ψ modification in RNA

these prematurely truncated transcripts and not confused with other sequencing errors (Helm and Motorin 2017). Alternate chemical labels have been recently developed, including acrylonitrile with an azidoethyl group as a clickable moiety and an acrylamidofluorescein reagent that allows for affinity capture with anti-fluorescein antibodies (Knutson et al. 2018; Li et al. 2019). Additionally, a 2018 study proposed the use of CRISPR-associated Argonaute proteins (MpAgo) for preferential enrichment of inosine-modified RNAs (Lapinaite et al. 2018).

2.5 Pseudouridine (Ψ)

Pseudouridine is an isomer of uridine, created by the rotation of uracil around the C-C glycosidic bond (Fig. 9), and the overall most abundant modification at a Ψ/U ratio of 0.2%–0.6% (Lovejoy et al. 2014; Schwartz et al. 2014). Since isomerization to pseudouridine causes a carbon-carbon bond to form between the base and sugar, pseudouridylation is not thought to be readily reversible. The rotation of uracil can be induced in an RNA-dependent manner with the box H/ACA ribonucleoproteins or in an RNA-independent manner with Ψ synthases (Fig. 9) (Hamma and

Ferré-D'Amaré 2006; Kiss et al. 2010; Li et al. 2016a). These Ψ synthases catalyze the isomerization in response to stress conditions, like heat shock. In humans, there have been 13 Ψ synthases discovered to date, with "stand-alone" synthases and an RNA-dependent Ψ synthase DKC1 that catalyzes a subset of Ψ in mRNA (Carlile et al. 2019). This modification is highly abundant and widespread in both mRNAs and ncRNAs, including rRNAs, tRNAs, and snRNAs, mainly located in coding sequences and 3′ UTR (Carlile et al. 2014; Schwartz et al. 2014; Li et al. 2015). The nature of the pseudouridine modification does not interrupt the canonical Watson–Crick preference of uridine to pair with adenosine. However, studies have shown that pseudouridine constrains flexibility in single-stranded RNA, which can subsequently modulate the function of that RNA. This constraint has been potentially attributed to the pseudouridine's coordination of an additional water molecule, causing hydrogen bonding to occur with the adjacent phosphate backbone. The additional water molecule also increases base stacking by favoring a 3′-endo conformation of the ribose. This has the effect of restricting the base to an axial *anti* conformation, which also restricts the flexibility of the residue (Charette and Gray 2000). Further, it has been proposed that the additional hydrogen bonding group on the non-pairing edge of pseudouridine could allow for selective protein recognition through polar interactions in the major groove (Harcourt et al. 2017).

While the exact biological role of pseudouridine has still yet to be fully elucidated, certain studies have indicated that pseudouridine-containing mRNAs had both increased and decreased translation levels compared to an unmodified transcript. When placed in nonsense codons, the presence of pseudouridine sites allows for complete read-through, which then produces a protein product with a C-terminal extension (Roundtree et al. 2017). These effects could suggest that pseudouridylation has a function in the modulation of translation, but is dependent on the transcript in question, the sequence context, and the expression system (Schwartz et al. 2014).

Due to its structural nature, pseudouridine is also considered an "RT-silent" modification like m^6A or m^5C and cannot be discriminated from uridine through direct sequencing methods. Therefore, chemical treatment with *N*-cyclohexyl-*N*′-β-(4-methylmorpholinium) ethylcarbodiimide (CMC) specifically labels pseudouridine's base. As a result, the CMC-Ψ construct causes reverse transcription stops on the 3′ side of the labeled Ψ site. Many methods also use the CMC derivative 1-cyclohexyl-(2-morpholinoethyl)carbodiimide metho-*p*-toluene sulfonate (CMCT). Three methods, published independently, take advantage of the CMC labeling: Ψ-seq (Schwartz et al. 2014), Pseudo-seq (Carlile et al. 2014), and pseudouridine site identification (PSI-seq) (Lovejoy et al. 2014) (Fig. 10). A newer method, N3-CMC-enriched pseudouridine sequencing (CeU-seq) (Fig. 10), uses a clickable CMC derivative, azido-CMC (N3-CMC), to conjugate the modified site with a biotin molecule. The biotin pull-down helps to enrich a pool of Ψ-rich RNAs and increases the sensitivity of the mapping of the Ψ sites (Li et al. 2015).

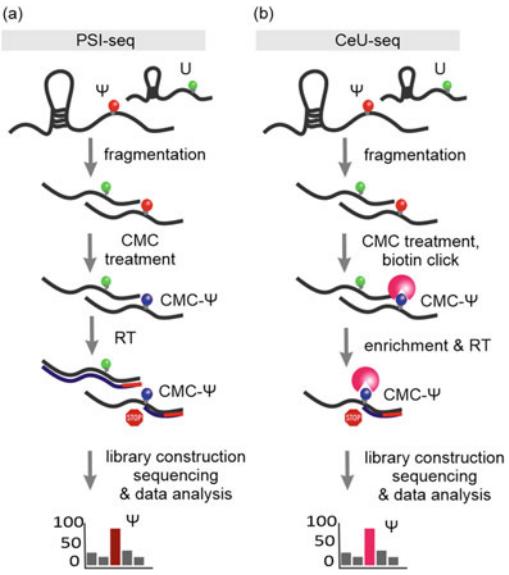


Fig. 10 High-throughput detection strategies for detection of Ψ modification in RNA. (**a**) PSI-seq treats the target RNA with CMC to specifically label the Ψ base. The CMC-modified Ψ base induces RT stops and produces truncated transcripts compared to untreated RNA transcripts. (**b**) CeU-seq modifies Ψ sites with the clickable azido-CMC. Biotin conjugated to the azido-CMC moiety can then be used to pull-down Ψ-containing RNA

Fig. 11 Structure of ribose and cellular machinery for the formation of methylation at the 2′-*O* position in ribose of RNA

2.6 2′-O-methylation (Nm)

Beyond the methylations mentioned above that occur on the RNA base, there can be methylations that occur on the ribose (Fig. 11). Methylation at the 2′- hydroxyl group of the ribose forms Nm (N = A, U, C, G), which is present at a frequency of about two modifications per transcript (Elliott et al. 2019). Nm has been seen in mRNA, tRNA, rRNA, and snRNA, preferentially placed in the first two nucleotides adjacent to the 5′ cap (Schibler and Perry 1977). A variety of enzymes have been associated with the formation of Nm, including stand-alone methyltransferases CMTR1 and CMTR2 for cap-proximal Nm installation (Fig. 11) (Bélanger et al. 2010; Werner et al. 2011). A ribonucleoprotein complex containing fibrillarin (FBL) and C/D-box guide snoRNA is another option for Nm modifications (Somme et al. 2014; Shubina et al. 2016).

On a structural basis, the presence of Nm masks the hydrophilic character of the 2′-hydroxyl group of RNA, and subsequently, the presence of Nm increases RNA-RNA duplex stability. The methylation of the hydroxyl group also allows for interactions promoting more complex RNA secondary structures and influencing

RNA-protein binding events, potentially due to a change in the availability of the minor groove (Roundtree et al. 2017).

As compared to the previous modifications, the location of Nm on the ribose means that it does not directly impact the Watson–Crick base-pairing. However, Nm instead stops the reverse transcription and therefore may modulate the biological activity of RNAs. This acylation reaction at the 2′ position of the ribose is commonly used to map single-stranded regions of RNA and predict RNA secondary structure (Spitale et al. 2013). However, more recently, there have been studies proposing that Nm modification of the codon inhibits ribosomal protein translation through sterically hindering the interaction between the ribosomal RNA (rRNA) and the ribose backbone of the mRNA-tRNA complex. Therefore, Nm also has the potential for post-transcriptional gene expression regulation (Elliott et al. 2019).

Multiple high-throughput sequencing techniques have been developed for recognition of Nm modifications, the first being RiboMeth-seq. This chemical-based method uses the tendency of the 2′-hydroxyl to be hydrolyzed upon alkaline treatment. Once RNAs are subjected to limited alkaline conditions, they fragment, whereas modified RNAs with the 2′-methyl group, hindering alkaline hydrolysis, are left intact. Therefore, the sites of 2′-methylation are seen in sequencing data as underrepresented read ends (Limbach and Paulines 2017).

Other detection techniques include Nm-seq, RibOxi-seq, 2MeO-seq, and MeTH-seq. Nm-seq (Fig. 12) relies on an iterative oxidation-elimination-dephosphorylation (OED) cycling process that removes one 2′ unmodified nucleotide per round in the 3′-to-5′ direction. This removal works by oxidizing the vicinal diols in these nucleotides by sodium periodate to generate a dialdehyde intermediate which is spontaneously β-eliminated under mild basic conditions. The OED process repeats until an Nm is reached, which prevents any further elimination due to the lack of vicinal diols able to be oxidized. This is followed by a single final round of oxidation-elimination (OE) reaction, which generates an unligatable 3-′-monophosphate on unmodified ends. Transcripts containing Nm-modified ends, however, are resistant to oxidation and keep their 3′-hydroxyl group. This means that only Nm-modified ends have the ability to be ligated with the adapter and then subsequently sequenced (Dai et al. 2017).

Ribose oxidation sequencing (RibOxi-seq) functions on a similar principle to Nm-seq, also using a periodate oxidation treatment to prevent non-2′-*O*-methylated riboses from being ligated to adapters during construction of the library. Prior to the periodate treatment, however, the RNA is randomly digested with Benzonase nuclease to ensure that all possible 3′-ends are represented for the analysis (Zhu et al. 2017).

Differing from the previous sequencing methods, 2′-*O*-methylation sequencing (2MeO-seq) instead manipulates dNTP concentrations for Nm modification discovery, using random hexamers hybridized to the 3′ sequencing adapter. During reverse transcription, low dNTP concentration causes RTase to stall one nucleotide downstream from the site of methylation, which is then visible when sequenced as compared to sequencing results using a high dNTP concentration (Incarnato et al. 2017).

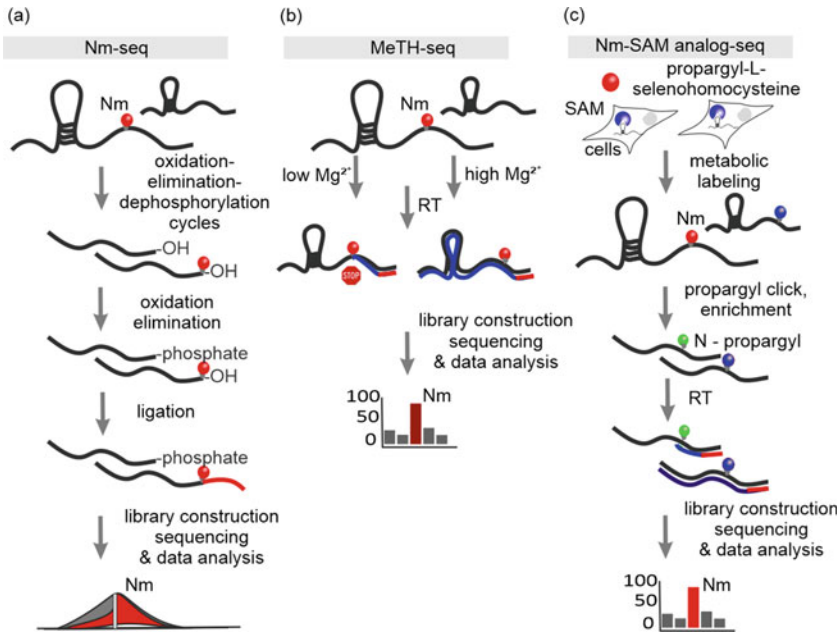


Fig. 12 High-throughput detection strategies for detection of Nm modification in RNA. (**a**) Nm-seq uses a sequence of oxidation-elimination-dephosphorylation cycling to iteratively remove a 2′ unmodified nucleotide in the 3′-to-5′ direction. Nm is unable to be removed through oxidation which terminates the cycle. A final oxidation-elimination round converts the 3′-hydroxyl to an unligatable phosphate. Transcripts with a terminal Nm retain the 3′-hydroxyl group and can be ligated with adapters for sequencing. (**b**) MeTH-seq reduces Mg^{2+} concentration to cause truncations one nucleotide 3′ to the Nm site. Comparison against transcripts reverse transcribed under high Mg^{2+} concentration reveals locations of Nm. (**c**) Nm-SAM analog-seq employs propargyl-L-selenohomocysteine, an *S*-adenosyl-methionine analog, which adds a clickable propargyl group in place of a methyl group at the 2′ *O*-position on the sugar. The propargyl allows for biotin to be clicked on and enrich the modified RNAs. RT stops caused by the modification allow for the identification of the Nm sites

Another alternate Nm sequencing method, methylation at two prime hydroxyls sequencing (MeTH-seq) (Fig. 12) forgoes an initial enrichment step of Nm-containing RNA. Instead, it utilizes a limited concentration of Mg^{2+} to cause RT to stop one nucleotide 3′ to the Nm sites. These characteristic stops will then be seen in the cDNA library and in subsequent sequencing to determine specific locations of Nm (Bartoli et al. 2018).

Sites of Nm can also be enriched through treatment with an *S*-adenosyl-methionine analog similar to m^6A. In this case, the analog contains propargyl-L-selenohomocysteine, which adds a clickable propargyl group rather than a methyl group to the nucleotides. Biotin can then be conjugated to the propargyl group via click chemistry for the enrichment of the RNA targets of any RNA methyltransferase incorporating this methionine analog. The samples were processed using the regular

Table 1 Characterization of methods for detection and validation of epitranscriptomic modifications

Method	Approach	Analytes	Pros	Cons
m^6A-seq MeRIP-seq	Ab-based enrichment tandem with high-throughput RNA-seq	m^6A, m^1A, and m^5C	Enrichment step increases sequencing depth	Not single-nucleotide resolution and variability in the specificity of Ab
MeRIP-CLIP PA-seq	Ab-based enrichment and UV-crosslink tandem with high-throughput RNA sequencing	m^6A	Enrichment step increases sequencing depth and single-nucleotide resolution	Low crosslink efficiency and variability in the specificity of Ab
SCARLET and similar historical and present methods	Site-specific cleavage and radioactive-labeling followed by ligation-assisted extraction and thin-layer chromatography	m^6A and Ψ	Specialized instruments not required, quantitative and qualitative data	Only for single-site analysis and time and labor consuming
m^1A-seq-TGIRT m^1A-MAP	Pattern from RT stops, or miscoding caused by m^1A tandem with high-throughput RNA sequencing	m^1A	Single-nucleotide resolution and quantitative data	Semi-quantitative and may requires additional treatment with AlkB enzyme
BS-seq	Bisulfite RNA treatment tandem with high-throughput RNA-seq	m^5C	Single-nucleotide resolution and quantitative data	Substantial input material and its degeneration upon BS treatment, mapping requires increased sequencing depth, and bias for incomplete deamination in highly structured regions
m^6A-REF-seq MASTER-seq DART-seq ICE-seq Aza-IP m^6Am-Exo-seq	Chemical reagent or enzymatic modification causes RT stop or miscoding, tandem with high-throughput RNA-seq	m^6A, I, Ψ, Nm	Single-nucleotide resolution	False positives from structural or sequential motifs and reagents available only for a few types of modifications
2MeO-seq	RT stop at low dNTP concentrations	Nm	Easy to conduct and provides information on sequence	Only for single-site analysis and indirect validation

(continued)

Table 1 (continued)

Method	Approach	Analytes	Pros	Cons
RiboMeth-seq	Alkaline treatment tandem with high-throughput RNA-seq	Nm	Single-nucleotide resolution, quantitative data, and small amount of input material	False positives from structural motifs, mapping requires increased sequencing depth and good for abundant RNA species
HPLC-MS/MS	Mass spectrometry of RNase digested RNA; mainly nucleosides	Multiple modifications	Includes sequence information, single-nucleotide resolution, and quantitative data	Substantial amount of input material and requires mass spectrometers

RNA protocol, with the RT stops allowing for identification of the Nm sites (Hartstock et al. 2018).

3 Methods for Multiple Modification Identification

The methods summarized in the previous sections have all been designed for targeting a single specific modification within a sample. However, there are certain cases in which transcriptome-wide profiling of multiple modifications is desirable. At this time, there exist a broad range of methods giving the possibility to screen and analyze multiple RNA modifications. These include mass spectrometry methods, high-throughput sequencing techniques, and bioinformatics tools that allow for the determination of multiple distinct modifications within a given sample.

To begin with mass spectrometry, RNA analysis has been adapted from proteomics methods. The sample is subjected to enzymatic cleavage to break down the oligonucleotides into nucleosides. These nucleosides are run through the mass spectrometer, and the presence of any modifications is detected through both the fragment patterns and comparison with the estimated mass of an unmodified fragment. Specific development of protocols is essential, given that RNA is a poly-anion and therefore does not transfer into the vacuum, or "fly," as well as proteins or DNA (Tromp and Schürch 2005).

Mass spectrometry methods have been adapted and geared towards the discovery of m^5C and are consequently also able to detect the related hm^5C, f^5C, and ca^5C modifications. These run the gamut from osmium-tagged derivatization-MS, MnO_2 oxidation/dansylhydrazine (DNSH) derivatization, 2-bromo-1-(4-diethylaminophenyl)-ethanone (BDEPE) derivatization, Girard reagent P (GirP) derivatization with in-tube solid-phase microextraction (SPME), *p*-dimethylaminophenacyl (DMPA) derivatization-LC/MS, 8-(diazomethyl)quinoline derivatization-LC/MS, and GCB-solid-phase extraction-liquid chromatography-electrospray ionization-MS/MS protocols for use with LC-MS (Debnath and

Okamoto 2018; Lan et al. 2019; Zhang et al. 2019a). There is also malic acid-enhanced hydrophilic interaction liquid chromatography-tandem mass spectrometry (HILIC-MS/MS) and radical transfer dissociation mass spectrometry (RTD-MS) for m^5C and related modification identification (Guo et al. 2018; Calderisi et al. 2020). A summary of described mass spectrometry methods can be seen in Table 2.

Table 2 Characterization of mass spectrometry methods available for detection of multiple epitranscriptomic modifications

Methods	Analytes	Pros	Cons
CMC-derivatization-MALDI/MS	Ψ	Able to determine Ψ sites in RNA	Non-specific CMC modification of other uridine nucleosides
DSPE-acetone derivatization-LC/MS	m^6A, m^1A, m^5C, f^5C, I, Ψ	7–30-fold increase in detection sensitivity	Cannot detect Nm
MnO_2 oxidation-DNSH derivatization-LC/MS	hm^5C, f^5C	363–380-fold increase in detection sensitivity	Tedious sample processing
BDEPE derivatization-LC/MS	m^5C, hm^5C, f^5C, ca^5C	70–313-fold increase in detection sensitivity	Requires SPE for salt removal
GirP derivatization combined with in-tube SPME-LC/MS	f^5C	310–880-fold increase in detection sensitivity	Requires monolith preparation
DMPA derivatization-LC/MS	m^5C	88–372-fold increase in detection sensitivity	Requires SPE for excess derivatization reagent removal
8-(Diazomethyl)-quinoline derivatization-LC/MS	m^6A, m^1A, m^5C, hm^5C, f^5C, Nm	56–137-fold increase in detection sensitivity	Derivatization reagent unstable
Malic acid-enhanced HILIC-MS/MS	m^5C, hm^5C	20–40-fold increase in detection sensitivity and specific detection of low-abundance modifications: hm^5C	Malic acid addition needed
RTD-MS	hm^5C, f^5C	Minimal sample preparation and good for highly labile modifications	FT-ICR tandem with ESI/MS required
Multi-column LC/MS-MS	m^1A, m^5C, Ψ, Nm	Highly rapid analysis	Specialized experimental set-up required
SPME-LC-MS/MS with ZrO_2/SiO_2 composite SPME column	m^6A, m^1A, m^5C, hm^5C, f^5C, I, Ψ	Increased separation and extraction on-column	ZrO_2/SiO_2-modified column required
Nano-LC/UPLC-ESI-MS/MS	m^6A, m^1A, f^5C, Nm	Minimal amounts of material required with high resolution of separation	Specialized experimental set-up required

Beyond the m^5C family of modifications, there are also other LC-MS/MS and ultra-performance liquid chromatography (UPLC)-MS/MS methods that can be used for a more comprehensive profiling of ribonucleoside modifications, including m^6A, m^5C, m^1A, inosine, Ψ, and Nm (Jiang et al. 2016; Jinno et al. 2017; Tardu et al. 2019). A liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS/MS)-based method has been developed for m^6A analysis (Yuan 2017). In 2019, another group presented a method for mapping m^1A sites in tRNA using two-dimensional polyacrylamide gel electrophoresis (PAGE), which provides higher levels of separation, followed by nano-LC-MS/MS (Antoine et al. 2019). Strides have also been made towards developing mass spectrometry methods to identify pseudouridylation. These include CMC derivatization-matrix-assisted laser desorption/ionization mass spectrometry (MALDI/MS), which provides each pseudouridine with a characterizable mass increase, and a direct MS-based sequencing using a Ψ-specific signature ion (Yamauchi et al. 2016; Lan et al. 2019). A modified high-performance liquid chromatography (HPLC)-UV assay has also been presented as a more economical choice for rapid screening of shifts in pseudouridine level within a sample (Xu et al. 2017).

High-throughput sequencing methods have typically been limited to the detection of one specific modification at a time. This can be a drawback when wanting to profile multiple RNA modifications within a singular sample. Recently, RNA bisulfite sequencing (RBS-seq) was published as a variant of bisulfite sequencing that allowed for the simultaneous detection of m^5C, m^1A, and Ψ at transcriptome-wide single-base resolution. This technique takes advantage of the canonical behavior of m^5C to resist deamination when treated with bisulfite, whereas C will convert to U. As discussed above, m^1A often causes misincorporation/mismatches when left as is during reverse transcription. However, bisulfite treatment converts m^1A to m^6A, which has no issue with reverse transcription. Lastly, Ψ will form a monobisulfite adduct upon bisulfite treatment, leading to a deletion signature at the sites of modification. Therefore, when comparing a bisulfite-treated (BS) versus non-bisulfite-treated (NBS) sequencing read, m^5C will appear as non-converted C in BS, m^1A will appear as mutated A sites only in NBS, and Ψ will appear as T deletion sites in BS (Khoddami et al. 2019).

Bioinformatics pipelines are also generally tuned for the discovery of one specific modification, but recent studies suggest a different tactic for detection, proposing an examination of the patterns embedded in deep sequencing data. The hypothesis offers a scenario in which each modification has its own fingerprint in patterns of mutations, deletions, insertions, and truncations, and, due to the nature of each modification, these may or may not be detectable above normal background levels seen for unmodified RNA. Researchers found distinguishable differences for m^1A and inosine as a result of reverse transcriptase-induced biases in the ultra-deep sequencing experiments, with the expectation that more unique modification fingerprints will be unearthed with further research (Kietrys et al. 2017; Sas-Chen and Schwartz 2019). Beyond RNA fingerprinting, there is a vast quantity of bioinformatics methods and pipelines for modification discovery and analysis, with more being added by the month as major efforts are concentrated in this area of research.

4 Conclusions/Outlook

There is a multitude of transcriptome-wide techniques available for the detection of modifications in a variety of RNAs. These techniques show impressive levels of specificity, sensitivity, and coverage, which resulted from important innovations in detection methods that have occurred over the past few years. Even so, many of the current methods still only paint part of the picture, providing a way to identify at a single-base resolution but not quantify modification levels at the same time, or vice versa. An ideal quantitative detection method would include single-base resolution and high levels of accuracy and precision. Inherent biases also remain in many sequencing technologies as a result of RNA structure and abundance. Therefore, to get a true picture of the RNA epitranscriptome, both single-cell RNA modification sequencing and more simultaneous multiple modification sequencing technologies need to be developed, as there is the potential for different RNA modifications to interact with one another within the cellular system. Further on in the future of epitranscriptome recognition, development of tools to allow for quantitative recognition of how dynamic a particular modification is—regarding the percentage of transcripts with specific single-site modification—and co-identification of two modifications on the same transcript within a sample would fill another gap in current technologies. Continued advances in epitranscriptomic identification technologies will allow for further probing into how epitranscriptomic modifications impact the RNA's structure and interactions with other biomolecules, which affects splicing, maturation, stability, expression, and degradation.

References

Alarcón CR, Goodarzi H, Lee H et al (2015) HNRNPA2B1 is a mediator of m^6A-dependent nuclear RNA processing events. Cell 162:1299–1308

Antoine L, Wolff P, Westhof E et al (2019) Mapping post-transcriptional modifications in *Staphylococcus aureus* tRNAs by nanoLC/MSMS. Biochimie 164:60–69

Athanasiadis A, Rich A, Maas S (2004) Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLoS Biol 2:2144–2158

Bartoli KM, Schaening C, Carlile TM et al (2018) Conserved methyltransferase Spb1 targets mRNAs for regulated modification with 2′-O-methyl ribose. BioRxiv. https://doi.org/10.1101/271916

Bélanger F, Stepinski J, Darzynkiewicz E et al (2010) Characterization of hMTr1, a human Cap1 2′-O-ribose methyltransferase. J Biol Chem 285:33037–33044

Brzezicha B, Schmidt M, Makałowska I et al (2006) Identification of human tRNA:m^5C methyltransferase catalysing intron-dependent m^5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). Nucleic Acids Res 34:6034–6043

Calderisi G, Glasner H, Breuker K (2020) Radical transfer dissociation for de novo characterization of modified ribonucleic acids by mass spectrometry. Angew Chemie Int Ed 59:4309–4313

Carlile TM, Rojas-Duran MF, Zinshteyn B et al (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515:143–146

Carlile TM, Martinez NM, Schaening C et al (2019) mRNA structure determines modification by pseudouridine synthase 1. Nat Chem Biol 15:966–974
Charette M, Gray MW (2000) Pseudouridine in RNA: what, where, how, and why. IUBMB Life 49:341–351
Chen K, Lu Z, Wang X et al (2015) High-resolution N6-methyladenosine (m^6A) map using photo-crosslinking-assisted m^6A sequencing. Angew Chemie Int Ed 54:1587–1590
Chen W, Feng P, Yang H et al (2018) iRNA-3typeA: identifying three types of modification at RNA's adenosine sites. Mol Ther Nucleic Acids 11:468–474
Chen Y, Yang S, Peng S et al (2019) N1-Methyladenosine detection with CRISPR-Cas13a/C2c2. Chem Sci 10:2975–2979
Chow CS, Lamichhane TN, Mahto SK (2007) Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. ACS Chem Biol 2:610–619
Cozen AE, Quartley E, Holmes AD et al (2015) ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. Nat Methods 12:879–884
Dai Q, Moshitch-Moshkovitz S, Han D et al (2017) Nm-seq maps 2′-O-methylation sites in human mRNA with base precision. Nat Methods 14:695–698
Dai T, Pu Q, Guo Y et al (2018) Analogous modified DNA probe and immune competition method-based electrochemical biosensor for RNA modification. Biosens Bioelectron 114:72–77
Danan C, Manickavel S, Hafner M (2016) PAR-CLIP: a method for transcriptome-wide identification of RNA binding protein interaction sites. Methods Mol Biol 1358:153–173
Debnath TK, Okamoto A (2018) Osmium tag for post-transcriptionally modified RNA. Chembiochem 19:1653–1656
Decatur WA, Fournier MJ (2002) rRNA modifications and ribosome function. Trends Biochem Sci 27:344–351
Delatte B, Wang F, Vo Ngoc L et al (2016) Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. Science 351:282–285
Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. Nature 485:201–206
Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S et al (2016) The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. Nature 530:441–446
Edelheit S, Schwartz S, Mumbach MR et al (2013) Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m^5C within archaeal mRNAs. PLoS Genet 9:1–14
Elliott BA, Ho H-T, Ranganathan SV et al (2019) Modification of messenger RNA by 2′-O-methylation regulates gene expression in vivo. Nat Commun 10:1–9
Fu L, Guerrero CR, Zhong N et al (2014) Tet-mediated formation of 5-hydroxymethylcytosine in RNA. J Am Chem Soc 136:11582–11585
Garcia-Campos MA, Edelheit S, Toth U et al (2019) Deciphering the "m^6A code" via antibody-independent quantitative profiling. Cell 178:731–747
Grozhik AV, Olarerin-George AO, Sindelar M et al (2019) Antibody cross-reactivity accounts for widespread appearance of m^1A in 5′ UTRs. Nat Commun 10:1–13
Guo C, Xie C, Chen Q et al (2018) A novel malic acid-enhanced method for the analysis of 5-methyl-2′-deoxycytidine, 5-hydroxymethyl-2′-deoxycytidine, 5-methylcytidine and 5-hydroxymethylcytidine in human urine using hydrophilic interaction liquid chromatography-tandem mass spectrometry. Anal Chim Acta 1034:110–118
Hamma T, Ferré-D'Amaré AR (2006) Pseudouridine synthases. Chem Biol 13:1125–1135
Harcourt EM, Kietrys AM, Kool ET (2017) Chemical and structural effects of base modifications in messenger RNA. Nature 541:339–346
Hartstock K, Nilges BS, Ovcharenko A et al (2018) Enzymatic or in vivo installation of propargyl groups in combination with click chemistry for the enrichment and detection of methyltransferase target sites in RNA. Angew Chemie Int Ed 57:6342–6346
Hauenschild R, Tserovski L, Schmid K et al (2015) The reverse transcription signature of N-1-methyladenosine in RNA-Seq is sequence dependent. Nucleic Acids Res 43:9950–9964

Helm M, Motorin Y (2017) Detecting RNA modifications in the epitranscriptome: predict and validate. Nat Rev Genet 18:275–291

Hong T, Yuan Y, Chen Z et al (2018) Precise antibody-independent m6A identification via 4SedTTP-involved and FTO-assisted strategy at single-nucleotide resolution. J Am Chem Soc 140:5886–5889

Hsu PJ, He C (2019) High-resolution mapping of N6-methyladenosine using m^6A crosslinking immunoprecipitation sequencing (m^6A-CLIP-Seq). Methods Mol Biol 1870:69–79

Huber SM, Van Delft P, Mendil L et al (2015) Formation and abundance of 5-hydroxymethylcytosine in RNA. Chembiochem 16:752–755

Hussain S, Sajini AA, Blanco S et al (2013) NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep 4:255–261

Incarnato D, Anselmi F, Morandi E et al (2017) High-throughput single-base resolution mapping of RNA 2′-O-methylated residues. Nucleic Acids Res 45:1433–1441

Ito S, Shen L, Dai Q et al (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science 333:1300–1303

Jia G, Fu Y, Zhao X et al (2012) N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887

Jiang H-P, Chu J-M, Lan M-D et al (2016) Comprehensive profiling of ribonucleosides modification by affinity zirconium oxide-silica composite monolithic column online solid-phase microextraction – mass spectrometry analysis. J Chromatogr A 1462:90–99

Jinno D, Kanemitsu Y, Saitoh K et al (2017) Rapid and selective simultaneous quantitative analysis of modified nucleosides using multi-column liquid chromatography-tandem mass spectrometry. J Anal Sci Technol 8:1–9

Jurkowski TP, Meusburger M, Phalke S et al (2008) Human DNMT2 methylates tRNA Asp molecules using a DNA methyltransferase-like catalytic mechanism. RNA 14:1663–1670

Ke S, Alemu EA, Mertens C et al (2015) A majority of m^6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes Dev 29:2037–2053

Khoddami V, Cairns BR (2013) Identification of direct targets and modified bases of RNA cytosine methyltransferases. Nat Biotechnol 31:458–464

Khoddami V, Yerra A, Mosbruger TL et al (2019) Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. Proc Natl Acad Sci USA 116:6784–6789

Kietrys AM, Kool ET (2016) A new methyl mark on messengers. Nature 530:423–424

Kietrys AM, Velema WA, Kool ET (2017) Fingerprints of modified RNA bases from deep sequencing profiles. J Am Chem Soc 139:17074–17081

Kiss T, Fayet-Lebaron E, Jády BE (2010) Box H/ACA small ribonucleoproteins. Mol Cell 37:597–606

Knutson SD, Ayele TM, Heemstra JM (2018) Chemical labeling and affinity capture of inosine-containing RNAs using acrylamidofluorescein. Bioconjug Chem 29:2899–2903

Koh CWQ, Goh YT, Goh WSS (2019) Atlas of quantitative single-base-resolution N6-methyladenine methylomes. Nat Commun 10:1–15

Lan M-D, Yuan B-F, Feng Y-Q (2019) Deciphering nucleic acid modifications by chemical derivatization-mass spectrometry analysis. Chin Chem Lett 30:1–6

Lapinaite A, Doudna JA, Cate JHD (2018) Programmable RNA recognition using a CRISPR-associated Argonaute. Proc Natl Acad Sci 115:3368–3373

Levanon EY, Eisenberg E, Yelin R et al (2004) Systematic identification of abundant A-to-I editing sites in the human transcriptome. Nat Biotechnol 22:1001–1005

Li JB, Levanon EY, Yoon JK et al (2009) Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. Science 324:1210–1213

Li Y, Song S, Li C et al (2013) MeRIP-PF: an easy-to-use pipeline for high-resolution peak-finding in MeRIP-Seq data. Genomics Proteomics Bioinforma 11:72–75

Li X, Zhu P, Ma S et al (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol 11:592–597

Li X, Ma S, Yi C (2016a) Pseudouridine: the fifth RNA nucleotide with renewed interests. Curr Opin Chem Biol 33:108–116
Li X, Xiong X, Wang K et al (2016b) Transcriptome-wide mapping reveals reversible and dynamic N1-methyladenosine methylome. Nat Chem Biol 12:311–316
Li X, Xiong X, Zhang M et al (2017) Base-resolution mapping reveals distinct m^1A methylome in nuclear-and mitochondrial-encoded transcripts. Mol Cell 68:993–1005
Li Y, Göhl M, Ke K et al (2019) Identification of adenosine-to-inosine RNA editing with acrylonitrile reagents. Org Lett 21:7948–7951
Limbach PA, Paulines MJ (2017) Going global: the new era of mapping modifications in RNA. Wiley Interdiscip Rev RNA 8:1–27
Linder B, Grozhik AV, Olarerin-George AO et al (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods 12:767–772
Liu N, Parisien M, Dai Q et al (2013) Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA 19:1848–1856
Liu J, Yue Y, Han D et al (2014) A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol 10:93–95
Liu N, Dai Q, Zheng G et al (2015) N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature 518:560–564
Liu F, Clark W, Klungland A (2016) ALKBH1-mediated tRNA demethylation regulates translation. Cell 167:816–828
Liu S, Feng Y, Wu J-J et al (2020) m^6A facilitates YTHDF-independent phase separation. J Cell Mol Med 24:2070–2072
Lovejoy AF, Riordan DP, Brown PO (2014) Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*. PLoS One 9:1–15
Meyer KD (2019) DART-seq: an antibody-free method for global m^6A detection. Nat Methods 16:1275–1280
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149:1635–1646
Meyer KD, Patil DP, Zhou J et al (2015) 5′ UTR m^6A promotes cap-independent translation. Cell 163:999–1010
Molinie B, Wang J, Lim KS et al (2016) m^6A-LAIC-seq reveals the census and complexity of the m^6A epitranscriptome. Nat Methods 13:692–698
Mortazavi A, Williams BA, McCue K et al (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621–628
Motorin Y, Helm M (2010) tRNA stabilization by modified nucleotides. Biochemistry 49:4934–4944
Nigita G, Veneziano D, Ferro A (2015) A-to-I RNA editing: current knowledge sources and computational approaches with special emphasis on non-coding RNA molecules. Front Bioeng Biotechnol 3:37
Ohlson J, Pedersen JS, Haussler D et al (2007) Editing modifies the GABAA receptor subunit α3. RNA 13:698–703
Okada S, Ueda H, Noda Y et al (2019) Transcriptome-wide identification of A-to-I RNA editing sites using ICE-seq. Methods 156:66–78
Ou X, Pu Q, Sheng S et al (2020) Electrochemical competitive immunodetection of messenger RNA modified with N6-methyladenosine by using DNA-modified mesoporous PtCo nanospheres. Microchim Acta 187:1–9
Paz-Yaacov N, Levanon Y, Nevo E et al (2010) Adenosine-to-inosine RNA editing shapes transcriptome diversity in primates. Proc Natl Acad Sci USA 107:12174–12179
Ping X-L, Sun B-F, Wang L et al (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res 24:177–189
Randerath K (1965) Two-dimensional separation of nucleic acid bases on cellulose. Nature 205:908
Roost C, Lynch SR, Batista PJ et al (2015) Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification. J Am Chem Soc 137:2107–2115

Roundtree IA, Evans ME, Pan T et al (2017) Dynamic RNA modifications in gene expression regulation. Cell 169:1187–1200
Rueter SM, Dawson TR, Emeson RB (1999) Regulation of alternative splicing by RNA editing. Nature 399:75–80
Safra M, Sas-Chen A, Nir R et al (2017) The m^1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature 551:251–255
Sakurai M, Ueda H, Yano T et al (2014) A biochemical landscape of A-to-I RNA editing in the human brain transcriptome. Genome Res 24:522–534
Sanger F, Brownlee GG, Barrell BG (1965) A two-dimensional fractionation procedure for radioactive nucleotides. J Mol Biol 13:373–398
Sas-Chen A, Schwartz S (2019) Misincorporation signatures for detecting modifications in mRNA: not as simple as it sounds. Methods 156:53–59
Schaefer M, Pollex T, Hanna K et al (2009) RNA cytosine methylation analysis by bisulfite sequencing. Nucleic Acids Res 37:12
Schibler U, Perry RP (1977) The 5′-termini of heterogeneous nuclear RNA: a comparison among molecules of different sizes and ages. Nucleic Acids Res 4:4133–4150
Schwartz S, Bernstein DA, Mumbach MR et al (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell 159:148–162
Sendinc E, Valle-Garcia D, Dhall A et al (2019) PCIF1 catalyzes m6Am mRNA methylation to regulate gene expression. Mol Cell 75:620–630.e9
Serra MJ (2004) Pronounced instability of tandem IU base pairs in RNA. Nucleic Acids Res 32:1824–1828
Shu X, Cao J, Cheng M et al (2020) A metabolic labeling method detects m^6A transcriptome-wide at single base resolution. Nat Chem Biol 16:887–895
Shubina MY, Musinova YR, Sheval EV (2016) Nucleolar methyltransferase fibrillarin: evolution of structure and functions. Biochemist 81:941–950
Somme J, Van Laer B, Roovers M et al (2014) Characterization of two homologous 2′-O-methyltransferases showing different specificities for their tRNA substrates. RNA 20:1257–1271
Sommer B, Köhler M, Sprengel R et al (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. Cell 67:11–19
Spitale RC, Crisalli P, Flynn RA et al (2013) RNA SHAPE analysis in living cells. Nat Chem Biol 9:18–20
Squires JE, Preiss T (2010) Function and detection of 5-methylcytosine in eukaryotic RNA. Epigenomics 2:709–715
Squires JE, Patel HR, Nousch M et al (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40:5023–5033
Sun P, Chen Y, Liu B et al (2019) DeepMRMP: a new predictor for multiple types of RNA modification sites using deep learning. Math Biosci Eng 16:6231–6241
Tardu M, Jones JD, Kennedy RT et al (2019) Identification and quantification of modified nucleosides in *Saccharomyces cerevisiae* mRNAs. ACS Chem Biol 14:1403–1409
Thomas B, Akoulitchev AV (2006) Mass spectrometry of RNA. Trends Biochem Sci 31:173–181
Tromp JM, Schürch S (2005) Gas-phase dissociation of oligoribonucleotides and their analogs studied by electrospray ionization tandem mass spectrometry. J Am Chem Soc Mass Spectrom 16:1262–1268
Wagner RW, Smith JE, Cooperman BS et al (1989) A double-stranded RNA unwinding activity introduces structural alterations by means of adenosine to inosine conversions in mammalian cells and Xenopus eggs. Proc Natl Acad Sci USA 86:2647–2651
Waku T, Nakajima Y, Yokoyama W et al (2016) NML-mediated rRNA base methylation links ribosomal subunit formation to cell proliferation in a p53-dependent manner. J Cell Sci 129:2382–2393
Wang S, Kool ET (1995) Origins of the large differences in stability of DNA and RNA helixes: C-5 methyl and 2′-hydroxyl effects. Biochemistry 34:4125–4132

Wang X, Lu Z, Gomez A et al (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505:117–120

Wang Y, Xiao Y, Dong S et al (2020) Antibody-free enzyme-assisted chemical approach for detection of N6-methyladenosine. Nat Chem Biol 16:896–903

Wei J, Liu F, Lu Z et al (2018) Differential m^6A, m^6Am, and m^1A demethylation mediated by FTO in the cell nucleus and cytoplasm. Mol Cell 71:973–985

Werner M, Purta E, Kaminska KH et al (2011) 2′-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. Nucleic Acids Res 39:4756–4768

Werner S, Schmidt L, Marchand V et al (2020) Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. Nucleic Acids Res 48:3734–3746

Xiao W, Adhikari S, Dahal U et al (2016) Nuclear m^6A reader YTHDC1 regulates mRNA splicing. Mol Cell 61:507–519

Xiao Y, Wang Y, Tang Q et al (2018) An elongation- and ligation-based qPCR amplification method for the radiolabeling-free detection of locus-specific N6-methyladenosine modification. Angew Chemie Int Ed 57:15995–16000

Xu J, Gu AY, Thumati NR, Wong JMY (2017) Quantification of pseudouridine levels in cellular RNA pools with a modified HPLC-UV assay. Genes 8:1–11

Yamauchi Y, Nobe Y, Izumikawa K et al (2016) A mass spectrometry-based method for direct determination of pseudouridine in RNA. Nucleic Acids Res 44:1–11

Yang H, Lam SL (2009) Effect of 1-methyladenine on thermodynamic stabilities of double-helical DNA structures. FEBS Lett 583:1548–1553

Yang X, Yang Y, Sun BF et al (2017) 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m^5C reader. Cell Res 27:606–625

Yin H, Wang H, Jiang W et al (2017) Electrochemical immunosensor for N6-methyladenosine detection in human cell lines based on biotin-streptavidin system and silver-SiO_2 signal amplification. Biosens Bioelectron 90:494–500

Yuan B-F (2017) Liquid chromatography-mass spectrometry for analysis of RNA adenosine methylation. In: Lusser A (ed) Methods in molecular biology. Springer Science+Business Media LLC, New York

Yuan F, Bi Y, Siejka-Zielinska P et al (2019) Bisulfite-free and base-resolution analysis of 5-methylcytidine and 5-hydroxymethylcytidine in RNA with peroxotungstate. Chem Commun 55:2328–2331

Zhang HY, Xiong J, Qi BL et al (2016) The existence of 5-hydroxymethylcytosine and 5-formylcytosine in both DNA and RNA in mammals. Chem Commun 52:737–740

Zhang Y-F, Qi C-B, Yuan B-F et al (2019a) Determination of cytidine modifications in human urine by liquid chromatography - mass spectrometry analysis. Anal Chim Acta 1081:103–111

Zhang Z, Chen L-Q, Zhao Y-L et al (2019b) Single-base mapping of m^6A by an antibody-independent method. Sci Adv 5:1–11

Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29

Zheng G, Qin Y, Clark WC et al (2015) Efficient and quantitative high-throughput tRNA sequencing. Nat Methods 12:835–837

Zhou J, Wan J, Gao X et al (2015) Dynamic m^6A mRNA methylation directs translational control of heat shock response. Nature 526:591–594

Zhou H, Kimsey IJ, Nikolova EN et al (2016) m^1A and m^1G disrupt A-RNA structure through the intrinsic instability of Hoogsteen base pairs. Nat Struct Mol Biol 23:803–810

Zhu Y, Pirnie SP, Carmichael GG (2017) High-throughput and site-specific identification of 2′-O-methylation sites using ribose oxidation sequencing (RibOxi-seq). RNA 23:1303–1314

Experimental Approaches and Computational Workflows for Systematic Mapping and Functional Interpretation of RNA Modifications

Quoseena Mir and Sarath Chandra Janga

Contents

Abstract In addition to four nucleotides in mRNA modified nucleotides are an essential addition to the standard genetic code in eukaryotes and prokaryotes. Epitranscriptomics has emerged as a new field to study nucleotide modifications in

Q. Mir
Department of BioHealth Informatics, School of Informatics and Computing, Indiana University Purdue University, Indianapolis, IN, USA

S. C. Janga (✉)
Department of BioHealth Informatics, School of Informatics and Computing, Indiana University Purdue University, Indianapolis, IN, USA

Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, 5021 Health Information and Translational Sciences (HITS), Indianapolis, IN, USA

Department of Medical and Molecular Genetics, Indiana University School of Medicine, Medical Research and Library Building, Indianapolis, IN, USA
e-mail: scjanga@iupui.edu

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_7

mRNA and examine their impact on gene expression. Systematic analysis of their occurrence and functions of RNA modifications is still a challenge. Most of the RNA modifications are low in abundance and technical limitations to study multiple modifications at a time, however, add more challenges in studying RNA modifications. Antibody-based selective methods played a major role in identifying global maps of modified nucleotides in mRNA and noncoding RNA, but limit to modifications with known antibodies. Long read sequencing introduced by Oxford Nanopore and PacBio now promise to overcome such limitations. We summarize here recent progress in experimental and computational methods to detect RNA modifications at single site and transcriptomic level and posit that emerging technologies would enable high-resolution tissue-specific spatial maps of epitranscriptomes in years to come.

Keywords Epitranscriptomic · RNA modifications · Computational methods · Nanopore sequencing · Experimental methods · Antibody detection · NGS · Long read sequencing

1 Introduction

The field of RNA modifications is relatively new but growing more and more every day, leading to the onset of a new research field known as epitranscriptomics (Delaunay and Frye 2019; Dimitrova et al. 2019; Jonkhout et al. 2017; Kadumuri and Janga 2018). The general purpose and exact functions of epitranscriptomic modification remain largely unknown, but many studies show their role in regulating gene expressions at transcriptomic level (Gilbert et al. 2016). RNA modifications are believed to perform their functions through two main approaches: structural changes in RNA that may either block or enhance interactions and direct recognition by different or more RNA binding proteins to induce various reactions (Liu et al. 2020). The outcome of modified transcripts is determined by coordinated action of the three groups of effector class of proteins (Kadumuri and Janga 2018): (1) RNA-modifying enzymes also known as writer proteins add a chemical modification to a target position on an RNA molecule; (2) RNA-binding proteins (RBPs) or reader proteins specifically recognize the modified nucleotides; and (3) eraser proteins recognize modified nucleotides and convert them back into unmodified nucleotides. RNA modifications can directly or indirectly shape the cellular transcriptome and proteome as they can affect a variety of molecular processes. Molecular processes that are affected by RNA modifications are transcription, splicing, RNA export, mRNA translation, mRNA stability, and RNA degradation.

Structure plays an important role in determining RNA function or role in cell. RNA modifications add chemical moieties to nucleotides and change the structural properties of RNAs post-transcriptionally on a large scale. Changes in structural organization of proteins directly or indirectly impact process depending on RNA

structure: (1) tRNAs, rRNAs, and other noncoding RNAs function depends on their structural arrangement, (2) arrangement of motifs or functional elements on lncRNAs and mRNA determines their structural stability for interactions, (3) regulatory RNAs (siRNAs or miRNAs) functionality is altered depending on accessibility of target sites, and (4) chemical changes in nucleotides affect base-pairing properties and impact RNA–RNA interactions (Tanzer et al. 2019).

Here, we describe certain experimental and computational methods used to study RNA modifications in detail (Tables 1 and 2). Many of these methods focus on modifications at single sites, but we mostly focus on methods that are applicable to study a spectrum of RNA modifications on a global scale and sensitive enough to quantify these modifications. Several methods can be used to analyze the entire transcriptome and reveal overall function of modification in an organism (Fig. 1). Global-scale approaches for detecting mRNA modifications are dependent on one of the following: (1) modification-based specific antibody to enrich modification-containing RNA, (2) RNA modification sensitive to ribonuclease, (3) specific chemical labeling of modification, and (4) base-pairing property of RNA modification (Limbach and Paulines 2017; Uddin et al. 2020; Zhu et al. 2019; Zhang et al. 2019a). The first epitranscriptome sequencing methods of m6A were developed using an m6A-specific antibody (m6A/MeRIP-seq), which identified thousands of m6A peaks in mammalian mRNA at a resolution of 100–200 nt (Dominissini et al. 2013; Schwartz et al. 2013). In the following sections, we highlight the major experimental approaches employed for detecting RNA modifications.

2 Experimental Methods for Detection of RNA Modifications at Single Sites

2.1 Reverse-Transcription-Based Methods

Reverse-transcription-based methods work on the principle that reverse transcriptase falls off while converting RNA into DNA strand if it encounters RNA modification in RNA strand. Reverse transcription-based methods are easy and cost-effective but require working on a known RNA molecule (Khoddami et al. 2019). RNA fragmentation during RNA isolation methods and RNA secondary structures are a major limitation to analyze results from reverse-transcription-based methods (Strobel et al. 2018). In both cases, reverse transcriptase will stop transcription and it is not possible to differentiate from enzyme fall-off due to nucleotide modification. 2′-*O* methylation is one of the modifications that cause enzyme to pause and continue reverse transcription instead of stopping it under special reaction conditions like low dNTPs. In certain conditions, a chemical treatment of RNA molecule helps control transcriptase pause or stop (Kellner et al. 2010; Kirpekar et al. 2005; Motorin et al. 2007; Ofengand et al. 2001).

Table 1 Experimental methods to detect RNA modifications

Detection technique	Modification	References
MeRIP-Seq	m^6A	Zeng et al. (2018)
m^6A-Seq	m^6A, m^6Am	Dominissini et al. (2012)
miCLIP	m^6A, m^6Am, m^5C	Grozhik et al. (2017), Hussain et al. (2013), Linder et al. (2015)
PA-m^6A-Seq	m^6A	Chen et al. (2015b)
m^6A-CLIP	m^6A	Grozhik et al. (2017), Linder et al. (2015)
SCARLET	m^6A	Liu and Pan (2016)
m^6A-LAIC-Seq	m^6A	Molinie et al. (2016)
RNA-BisSeq	m^5C	Chen et al. (2019a)
Aza-IP	m^5C	Khoddami and Cairns (2013)
m5C-RIP	m^5C	Edelheit et al. (2013)
hMeRIP-Seq	hm^5C	Delatte et al. (2016)
Pseudo-Seq	Ψ	Carlile et al. (2014)
PSI-Seq	Ψ	Schwartz et al. (2014)
CeU-Seq	Ψ	Li et al. (2015)
m^1A-Seq	m^1A	Dominissini et al. (2016)
m^1A-ID-Seq	m^1A	Li et al. (2017a)
m1A-MAP	m^1A	Li et al. (2017b)
m 6A-label-seq	m^6A	Shu et al. (2020)
mRNA editing	A to I and C to U mRNA editing	Tan et al. (2017)
Bs-seq	m^5C	Chen et al. (2018c)
NsunC271A-CLIP	m^5C	Hussain et al. (2013)
AlkAniline-Seq	m^7G, m^3C	Marchand et al. (2018)
m ^{7}G-MaP-seq	m^7G	Enroth et al. (2019)
m 6A-SEAL-seq	m^6A	Wang et al. (2020)
Nm-seq	Nm	Dai et al. (2017)
RibOxi-seq	Nm	Zhu et al. (2017)
RiboMethSeq	Nm	Marchand et al. (2016)
2′-OMe-Seq	Nm	Incarnato et al. (2017)
MAZTER-seq	m^6A	Garcia-Campos et al. (2019)
m 6A-REF-seq	m^6A	Zhang et al. (2019a)
DART-seq	m^6A	Meyer (2019)
TRIBE	m^6A	Worpenberg et al. (2019)
RBS-Seq	m^5C, Ψ, and m^1A	Khoddami et al. (2019)

Table 2 Computational methods to detect RNA modifications from short and long read sequencing data

Detection technique	Modification	References
iRNAm5C-PseDNC	m^5C	Qiu et al. (2017)
iRNA-PseColl	m^5C	Feng et al. (2017)
RNAm5Cfinder	m^5C	Li et al. (2018)
PEA-m5C	m^5C	Song et al. (2018)
m5C-HPCR	m^5C	Zhang et al. (2018)
RAMPred	m^1A	Chen et al. (2016b)
iRNA-3typeA	m^1A, m^6A	Chen et al. (2018b)
iRNA-Methyl	m^6A	Chen et al. (2015a)
m6Apred	m^6A	Chen et al. (2015c)
M6ATH	m^6A	Chen et al. (2016a)
RNA-MethylPred	m^6A	Jia et al. (2016)
TargetM6A	m^6A	Li et al. (2016)
pRNAm-PC	m^6A	Liu et al. (2016)
RNAMeth Pre	m^6A	Xiang et al. (2016a)
AthMeth Pre	m^6A	Xiang et al. (2016b)
M6A-HPCS	m^6A	Zhang et al. (2016)
SRAMP	m^6A	Zhou et al. (2016)
MethyRNA	m^6A	Chen et al. (2017a)
RAM-ESVM	m^6A	Chen et al. (2017b)
iRNA-PseColl	m^6A	Feng et al. (2017)
RAM-NPPS	m^6A	Xing et al. (2017)
iMethyl-STTNC	m^6A	Akbar and Hayat (2018)
iRNA(m6A)-PseDNC	m^6A	Chen et al. (2018a)
BERMP	m^6A	Huang et al. (2018)
M6AMRFS	m^6A	Qiang et al. (2018)
RFAthM6A	m^6A	Wang and Yan (2018)
M6APred-EL	m^6A	Wei et al. (2018)
HMpre	m^6A	Zhao et al. (2018)
WHISTLE	m^6A	Chen et al. (2019b)
iMethyl-Deep	m^6A	Mahmoudi et al. (2020)
Gene2Vec	m^6A	Zou et al. (2019)
DeepPromise	m^6A, m^1A	Chen et al. (2020)
EpiNano	m^6A	Liu et al. (2019)
ELIGOS	m^6A	Jenjaroenpun et al. (2021)
MINES	m^6A	Lorenz et al. (2020)

2.1.1 Detection of RNA Modifications Without Chemical Treatment of the RNA

Chemical modifications in purine (atoms 1, 2, and 6) and pyrimidine (atoms 2, 3, and 4) nucleobases prevent base-pairing interactions between complementary RNA bases. These chemical modifications at specific atoms in purines like m1G, m2A,

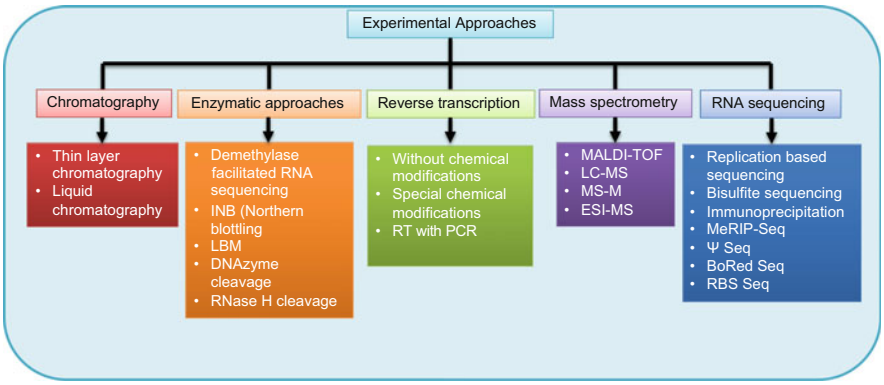


Fig. 1 Experimental approaches commonly employed to study RNA modifications. Some of these methods are specific to a single position or molecule while the newer variants coupled with high-throughput techniques such as sequencing are used to study RNA modifications at a transcriptomic level

m1A, m2A, and m3U make transcriptase stop converting RNA strand into DNA. But in the case of modifications on pyrimidine atoms like m6A, Nm, and m4C, reverse transcriptase enzymes pause on one nucleotide after modified base toward 3′ side (Behm-Ansmant et al. 2011; Gustafsson and Persson 1998). However, Hoogsteen modifications are small for RT enzyme to stop or pause like Ψ, m5C, m5U, A to I, m7G (Ofengand et al. 2001), so it is not possible to detect such modifications using this method, but in special conditions like in the case of a bulky modification if reactions are carried out at low dNTPs concentration.

2.1.2 Chemical Treatment of the RNA to Detect RNA Modifications by Reverse Transcription Method

Many of RNA modifications are difficult to detect using reverse-transcription-based methods; such modifications can be detected after treating RNA molecule with specific chemical reagents. Some of the common modifications that are detected using this approach are m5C, m3C, m7G, Nm, pseudouridine (Ψ), and dihydrouridine (D). Some of these methods are used for studying modifications at transcriptome level using high-throughput sequencing approach (Basturea et al. 2012; Toh et al. 2008). One such example is the study of Ψ modifications in yeast by Carlile et al. For sequencing Ψ modifications, RNA was chemically treated with CMC (N-cyclohexyl-N′-β-(4-methylmorpholinium)-ethylcarbodiimide) and converted into cDNA using reverse transcriptase method (Carlile et al. 2014; Basturea et al. 2006).

2.1.3 Detection of RNA Modifications by Reverse Transcription and PCR

Most of the RNA modifications are difficult to detect because of their low abundance at cellular level. For example, a small number of RNAs are 2′-*O*-methylated at their 3′-termini and are in low abundance, limiting its detection by common reverse transcription method. However, a version of reverse transcription RTL-P method developed by Dong et al. can be used to identify 2′-*O*-methylated residues (Dong et al. 2012). The method works at low concentrations of dNTPs for converting RNA into cDNA and is referred as reverse transcription at low (RTL-P) deoxyribonucleoside triphosphate (dNTP) concentrations and polymerase chain reaction (PCR) (Dong et al. 2012). In this method, RNA is purified from DNA contamination using DNAse I and reverse transcription is carried out at two concentrations of dNTPs, low and high (Dong et al. 2012). In normal or high dNTP concentrations, reverse transcriptase cannot differentiate between modified and unmodified bases. But in the case of low dNTP concentrations enzyme falls off at a base prior to modified base resulting in two amplified products (Dong et al. 2012; Anonymous 2018; Elliott et al. 2019; Motorin and Marchand 2018).

2.2 Chromatography-Based Methods

2.2.1 Thin-Layer Chromatography (TLC)

Thin-layer chromatography (TLC) method of studying RNA modifications is easy, fast, and cost-effective. It can be used in any lab as it does not require special expertise and equipment. But this method can be used to study modifications in small RNA fragments like noncoding RNA and tRNA. Small RNA fragments of ~100–150 nt can be studied using TLC method. Thus, larger RNA molecules need to be digested using multiple ribonuclease enzymes. Once RNA molecules are digested into smaller fragments and purified from ribosomal contamination, 2-D TLC method can be used to analyze modifications (Grosjean et al. 2004, 2007).

2.2.2 Liquid Chromatography

Over the years several methods have been developed to study RNA modifications and quantify them using liquid chromatography (LC) (Burtis 1970; Lakings and Gehrke 1971; Uziel et al. 1968). Usually, these methods require purifying RNA molecules and labeling them with radioactive labels. Before loading onto separation column RNA molecules are digested completely or partially using multiple nuclease enzymes followed by phosphatase treatment. Oligonucleotide bases are detected based on their radioactivity or UV absorption. The modified and unmodified bases are differentiated based on their mobility differences on chromatographic columns.

2.2.3 Mass Spectrometry-Based Methods

Given the sensitivity of mass spectrometry for mass detection of molecules, it has been most widely used to study post-translational modifications in proteins (Larsen and Roepstorff 2000); in the last couple of years, mass spectrometry methods were used to study modifications in RNA and DNA molecules as well (Giessing and Kirpekar 2012; Pandolfini et al. 2019). MS works on the simple principle of separating molecules based on mass to charge ratio. One of the main limitations is that mass spectrometry cannot quantity or locate modifications. This method can be used to identify modifications; however, locating them to a specific location remains a challenge; one adaptation was using RNA molecules obtained by specific tRNAs or using nuclease enzyme digested products (Jora et al. 2019; Jiang et al. 2019; Yu et al. 2019). But nuclease digestion-based methods lead to errors in mapping modifications to the genome. One major example of such errors is m7G in Let-7 miRNA; it is a known Gm site of rRNA fragments (Enroth et al. 2019; Pandolfini et al. 2019). Even though tRNA-based complementary approach has been used to overcome this issue but that also requires larger input concentrations and poses significant challenges at computational level (Jiang et al. 2019; Jora et al. 2019).

2.3 Enzymatic Methods to Detect Modification

2.3.1 RNase H Cleavage Method of Detecting 2′-*O*-methylation

This enzymatic method is mostly used to study enzymes causing RNA modifications as opposed to identify new modifications. Few RNA modifications have been identified using RNase H cleavage method like nucleotides methylated at 2′-*O* (Zhao and Yu 2004), methylated nucleotides, and pseudouridines (Lapham and Crothers 2000). One more enzyme that has been used to study nucleotide modifications is DNAzyme (Hengesbach et al. 2008; Santoro and Joyce 1997; Buchhaupt et al. 2007). m6A modification is one of the most abundant modification and is studied widely. One technique to detect m6A modification and quantify it is by using enzyme-based cleavage method and radioactive labeling (SCARLET) (Liu and Pan 2016). Most of these methods developed could not be applied to study modifications at transcriptome level like ligation-based methods (Saikia et al. 2006; Dai et al. 2007) and immuno-northern blotting (INB) (Cozen et al. 2015).

3 RNA Sequencing: Detection of RNA Modifications at Multiple Sites

3.1 Amplification-Based Sequencing Approaches

Most of the sequencing-based approaches need RNA to be converted into cDNA and amplified before sequencing. During the process of converting RNA into cDNA, most of the modifications are lost, inosines (A to I) being the only exceptional modification that can be detected even after RNA is converted to cDNA. One advantage is that reverse transcriptase reads I as G and adds mutations to modified locations. These mutations are compared to reference genome and are called as modified bases (Eisenberg et al. 2010; Wulff et al. 2017).

Bisulfite sequence is one of the most widely used amplification methods to study m5C modifications. When m5C-modified RNA molecules are treated with bisulfite, m5C remains unchanged while unmodified C base is converted into uridine (Motorin et al. 2010; Pollex et al. 2010; Schaefer et al. 2009). Sequencing after bisulfite treatment helps analyze m5c modifications at transcriptomic level.

3.2 Immunoprecipitation Method

Immunoprecipitation methods are used to target specific modifications and combined with sequencing approaches help to identify and quantify modifications at transcriptome level. One method developed by Khoddami and Cairns helped identify modifications in RNA species that are present in low abundance in the cell. But this method specifically targets m5C modifications (Khoddami and Cairns 2013). The main advantage of this method is it does not require gene modification of enzymes responsible for methylation. Once antibodies are crosslinked with modified nucleotides it results in either the addition of mutations or early fall off of the reverse transcriptase enzyme. To study mutations at transcriptome level, amplified cDNA molecules are sequenced using next-generation sequencing approaches (RIP-seq or meRIP-seq). Sequencing results in reads of ~100–200 nt in length; in order to track back modifications to single-nucleotide level mutations and read fragmentation is taken into consideration for data analysis (e.g., PA-m6A-seq (Chen et al. 2015b), miCLIP-seq (Boulias et al. 2019; Linder et al. 2015), m1A-MAP (Li et al. 2017b), and m7G-MeRIP-Seq (Zhang et al. 2019b)). In order to study differential modification level of various isoforms of same transcripts a method is used called LAIC-seq. This method studies m6A modifications in transcript isoforms but does not help in locating modification to individual base level (Molinie et al. 2016). A more specific antibody-based detection method of m6A modification was developed known as N (6)-methyladenosine-sequencing (m6A-seq) or MeRIP-seq (Dominissini et al. 2013).

3.3 RBS Sequencing

A major current limitation in detecting RNA modifications is the lack of methods that permit the mapping of multiple modification types simultaneously, at transcriptomic level and nucleotide-specific resolution. Khoddami et al. (2019) developed RBS-seq, a variant of the RNA bisulfite sequencing method to detect simultaneously m5C, Ψ, and m1A at single-base resolution transcriptome-wide. For m5C, bisulfite treatment deaminates unmodified Cs converting them to uridine (U)s (Ts upon cDNA sequencing), while methylated cytosines do not respond to the bisulfite treatment and remain Cs. For m1A, because m1A disturbs the canonical A:T base pairing, it pauses the reverse transcriptase leading to nucleotide misincorporation. Such mismatches act as m1A signatures in the synthesized cDNA. In contrast, under the alkaline conditions of the bisulfite treatment m1A converts to m6A (methyl passes from N1 to N6) (Macon and Wolfenden 1968), which reads as A and does not induce nucleotide misincorporation (see m6A-seq method above). Comparison of bisulfite-treated and untreated samples leads to identification of m1A sites. Bisulfite treatment of pseudouridine results in enzyme reverse transcriptase skipping a base at modification site.

4 Computational Methods for Predicting RNA Modifications

Over the years experimental techniques have advanced in the field of biology and genomics, helping molecular biologists identify different types biomolecular modifications. Even though techniques have developed, but we are still far from being able to study modifications on a larger scale. To help solve this problem many computational approaches are developed in parallel to experimental techniques to study modifications on a global scale. In the following sections, we highlight major computational frameworks that are commonly employed for detecting RNA modifications from short and long read sequencing datasets. Figure 2 presents a schematic listing various publicly available tools and the underlying biological rationale motivating these approaches.

4.1 Computational Methods Used to Predict Modifications from Short Read Sequencing Data

Advances in computational resources have led to an increase in the number of computational tools developed to predict RNA modifications in the last couple of years (Table 2). Since m6A is the most highly studied modification, there is abundant high-quality single-nucleotide resolution data facilitating computational

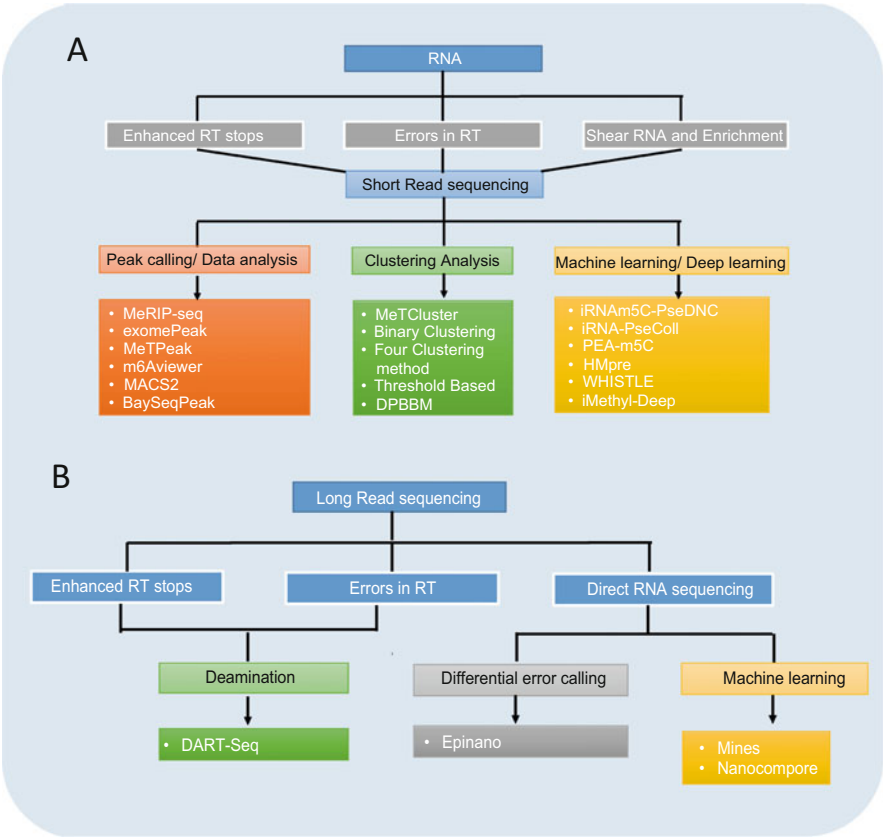


Fig. 2 Schematic showing the various computational workflows used to identify RNA modifications from transcriptome-wide datasets. (**a**) Short read-based approaches to study RNA modifications. (**b**) Long read-based workflows/steps to study single or multiple RNA modifications

tool development. So, many computational methods have been developed to predict m6A sites with SRAMP (Zhou et al. 2016), Gene2Vec (Zou et al. 2019), BERMP (Huang et al. 2018), and WHISTLE (Chen et al. 2019b). These four machine learning tools have become the basis for developing better modification prediction tools. In order to predict m6A site, SRAMP exploits sequence-derived features by combining three random forest classifiers (Zhou et al. 2016), while Gene2Vec (Zou et al. 2019) and BERMP (Huang et al. 2018) use convolutional neural networks (CNNs) and recurrent neural networks to predict m6A, respectively. WHISTLE (Chen et al. 2019b) is a tool based on support vector machines (SVMs) to predict the m6A sites; it not only uses sequence features but integrated 35 additional genomic features. RAMPred a m1A site predictor uses SVM-based model using features like nucleotide-chemical properties and nucleotide composition (Chen et al.

2016a). DeepPromise a CNN model uses direct RNA sequencing data to predict two major modifications m6A and m1A (Chen et al. 2020).

4.2 *Computational Methods Used to Predict Modifications from Long Read Sequencing Data*

Long read sequencing is a newly developed technology. Two major groups leading the field currently are Pacific Biosciences' (PacBio) single-molecule real-time (SMRT) sequencing (Roberts et al. 2017) and Oxford Nanopore Technologies' (ONT) nanopore sequencing. ONT can generate long sequencing reads of up to several kilobases from single mRNA/DNA molecules in real time (Banerjee et al. 2010; Stoddart et al. 2010). One of the main advantages to ONT is amplification-free direct RNA sequencing (DRS). As opposed to PacBio, Nanopore is nonsynthesis-based sequencing technique; thus, it does not require RNA to be converted into cDNA or amplified prior to sequencing. DRS is being explored to study RNA modifications; however, challenge remains in data analysis part of nanopore sequencing, base calling errors, and calling modifications from raw data signals (Anreiter et al. 2021). Computational methods to identify RNA modifications from single-molecule DRS data from ONT platforms can be grouped into two categories: (1) Algorithms based on the principle that modifications cause a significant alteration in raw signal as compared to unmodified bases that result in base calling error. (2) Machine learning models to interpret current raw signal from Nanopore machine. Lorenz et al. developed a software MINES, an Rf classifier trained using experimentally detected m6A modifications within DRACH sites in HEK293T cells. Its reported MINES can predict m6A modifications at an accuracy rate of 80% (Lorenz et al. 2020). The authors validated MINES using modification sites from CLIP data in WT and METTL3 knockdown samples. One more tool to study single-molecule-based RNA modifications was developed by Wongsurawat et al., ELIGOS (epitranscriptional landscape inferring from glitches of ONT signals). ELIGOS uses a rather different approach using cDNA sequencing data and comparing the error rates with DRS data calculating % error of specific bases (%ESB). So far it has been used to predict modification in rRNA data from *E. coli*, yeast, and human cells (Jenjaroenpun et al. 2021). EpiNano an algorithm to predict m6A predictions has been trained on synthetic methylated and unmethylated data signals (Liu et al. 2019). So far machine learning-based prediction models developed completely rely on training dataset quality.

5 Future Directions

Advances in sequencing technologies to generate whole genome sequencing and longer read lengths in addition to development of mass spectrometry technologies have enabled experimentalists to study RNA modifications at transcriptomic level. However, a lot about RNA modification role and function in cellular process is still unknown and low abundance of RNA modification along with low abundance of various species still leaves a lot of room to uncover and study (Shi et al. 2020). In recent years, many of the direct or indirect roles of RNA modifications in various processes have been demonstrated by researchers like process related to RNA structures, e.g., initiating translation, mRNA stability, splicing, nuclear transport, regulating RNA protein and RNA–RNA interactions (Alarcon et al. 2015a, b; Liu et al. 2015; Wang et al. 2014; Zheng et al. 2013). An open question concerning most of the publicly available datasets is how much the variability observed in mRNA modifications at a global level can be explained by various factors ranging from cell-based variability or limitations posed by various methods. One of the common limitations to both experimental and computational methods is studying all modifications separately instead of the fact that multiple modifications can be present on the same transcript governing various processes. Thus, it is important to develop a method that requires smaller inputs and helps identify multiple modifications at a cell-specific level or transcript level (Kadumuri and Janga 2018). While not many methods are available to study differential modification levels in isoforms of a transcript and whether they govern isoform role to a specific cell or tissue is not known. Modifications on isoforms can dictate its expression in a cell/tissue-specific manner, localization within the cell for translational purposes, stability, and interactions, thereby regulating its expression and translation in a spatiotemporal context (Fig. 3) (Chang et al. 2017).

It is possible to suggest based on several recent studies that RNA modifications can determine the function of the RNA message at three distinct levels: (1) Cell type—depending on tissue or cell type; the same modification can result in distinct functional outcome, thereby determining cell fate. (2) Spatiotemporal arrangement—based on its abundance in different compartments of a cell like nucleus, cytoplasm, or different organelles, a particular modification can govern multiple processes. (3) Combination of modifications—Since multiple modifications can be decorated on an RNA transcript, depending on the combination of modifications a different set of functions or interactions can be attained (Fig. 3). These scenarios indicate that the potential epitranscriptome could be far more complex in most tissue types than we currently understand. Hence, in order to map modifications and study their functional role, all three factors need to be considered in the future to develop methods. Methods that can locate multiple modifications to a particular transcript expressed in a specific cell type and abundant in various compartments of cell are needed to understand the functionality of modifications and their contributions to disease phenotypes.

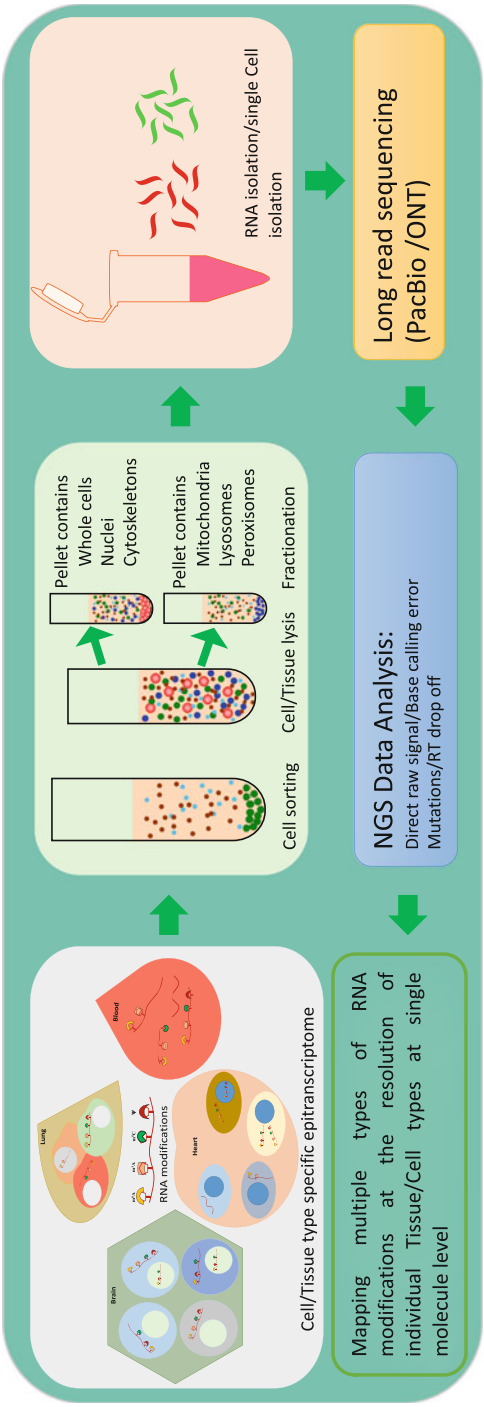


Fig. 3 Future directions to study RNA modifications at a single-molecule level. Schematic elaborates emerging single-molecule and single-cell approaches for simultaneous mapping of multiple RNA modifications in individual cell types

6 Conclusion

In the last few years, rapid growth has been seen in developing methods to study RNA modifications, but our understanding of their functions on mRNA is far from complete. Indeed, most of functional understanding of RNA modifications still largely comes from stable regulatory RNAs. Development of multiple methods has helped us understand modifications at site-specific level as well as on a global scale. However, good quantitative methods at base resolution and ways to simultaneously detect multiple types of modifications in the same RNA molecule or transcriptome-wide are still awaiting development. Long read DRS by ONT or from PacBio now opens more doors to develop better and sensitive methods. Long read sequencing platform from ONT is synthesis independent along with a better computational approach; it can be used to study multiple modifications on a single transcript. With technology development and single cell being explored to study cell-specific transcriptome details, it is possible to combine long read sequencing approach with single cell to identify cell-specific modification maps and functions. DRS from ONT could be a better approach when combined with single-cell sequencing technique to study RNA modification. But both techniques currently require higher input levels on experimental side posing a limitation and nanopore data analysis needs stronger and better computational approaches to reduce errors. For these techniques to be used widely, multiple-level improvement is required on experimental side but mostly on computational side to interpret signals with minimum errors possible.

References

Akbar S, Hayat M (2018) iMethyl-STTNC: identification of N(6)-methyladenosine sites by extending the idea of SAAC into Chou's PseAAC to formulate RNA sequences. J Theor Biol 455:205–211

Alarcon CR, Goodarzi H, Lee H et al (2015a) HNRNPA2B1 is a mediator of m(6)A-dependent nuclear RNA processing events. Cell 162:1299–1308

Alarcon CR, Lee H, Goodarzi H et al (2015b) N6-methyladenosine marks primary microRNAs for processing. Nature 519:482–485

Anonymous (2018) Corrigendum: Nm-seq maps 2′-O-methylation sites in human mRNA with base precision. Nat Methods 15:226–227

Anreiter I, Mir Q, Simpson JT et al (2021) New twists in detecting mRNA modification dynamics. Trends Biotechnol 39:72–89

Banerjee A, Mikhailova E, Cheley S et al (2010) Molecular bases of cyclodextrin adapter interactions with engineered protein nanopores. Proc Natl Acad Sci USA 107:8165–8170

Basturea GN, Rudd KE, Deutscher MP (2006) Identification and characterization of RsmE, the founding member of a new RNA base methyltransferase family. RNA 12:426–434

Basturea GN, Dague DR, Deutscher MP et al (2012) YhiQ is RsmJ, the methyltransferase responsible for methylation of G1516 in 16S rRNA of *E. coli*. J Mol Biol 415:16–21

Behm-Ansmant I, Helm M, Motorin Y (2011) Use of specific chemical reagents for detection of modified nucleotides in RNA. J Nucleic Acids 2011:408053

Boulias K, Toczydlowska-Socha D, Hawley BR et al (2019) Identification of the m(6)Am methyltransferase PCIF1 reveals the location and functions of m(6)Am in the transcriptome. Mol Cell 75:631–643 e638

Buchhaupt M, Peifer C, Entian KD (2007) Analysis of 2'-O-methylated nucleosides and pseudouridines in ribosomal RNAs using DNAzymes. Anal Biochem 361:102–108

Burtis CA (1970) The determination of the base composition of RNA by high-pressure cation-exchange chromatography. J Chromatogr 51:183–194

Carlile TM, Rojas-Duran MF, Zinshteyn B et al (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515:143–146

Chang M, Lv H, Zhang W et al (2017) Region-specific RNA m(6)A methylation represents a new layer of control in the gene regulatory network in the mouse brain. Open Biol 7

Chen W, Feng P, Ding H et al (2015a) iRNA-methyl: identifying N(6)-methyladenosine sites using pseudo nucleotide composition. Anal Biochem 490:26–33

Chen K, Lu Z, Wang X et al (2015b) High-resolution N(6)-methyladenosine (m(6) A) map using photo-crosslinking-assisted m(6) A sequencing. Angew Chem 54:1587–1590

Chen W, Tran H, Liang Z et al (2015c) Identification and analysis of the N(6)-methyladenosine in the *Saccharomyces cerevisiae* transcriptome. Sci Rep 5:13859

Chen W, Feng P, Ding H et al (2016a) Identifying N (6)-methyladenosine sites in the *Arabidopsis thaliana* transcriptome. Mol Gen Genomics 291:2225–2229

Chen W, Feng P, Tang H et al (2016b) RAMPred: identifying the N(1)-methyladenosine sites in eukaryotic transcriptomes. Sci Rep 6:31080

Chen W, Tang H, Lin H (2017a) MethyRNA: a web server for identification of N(6)-methyladenosine sites. J Biomol Struct Dyn 35:683–687

Chen W, Xing P, Zou Q (2017b) Detecting N(6)-methyladenosine sites from RNA transcriptomes using ensemble support vector machines. Sci Rep 7:40242

Chen W, Ding H, Zhou X et al (2018a) iRNA(m6A)-PseDNC: identifying N(6)-methyladenosine sites using pseudo dinucleotide composition. Anal Biochem 561–562:59–65

Chen W, Feng P, Yang H et al (2018b) iRNA-3typeA: identifying three types of modification at RNA's adenosine sites. Mol Ther Nucleic Acids 11:468–474

Chen YR, Yu S, Zhong S (2018c) Profiling DNA methylation using bisulfite sequencing (BS-Seq). Methods Mol Biol 1675:31–43

Chen YS, Ma HL, Yang Y et al (2019a) 5-Methylcytosine analysis by RNA-BisSeq. Methods Mol Biol 1870:237–248

Chen K, Wei Z, Zhang Q et al (2019b) WHISTLE: a high-accuracy map of the human N6-methyladenosine (m6A) epitranscriptome predicted using a machine learning approach. Nucleic Acids Res 47:e41

Chen Z, Zhao P, Li F et al (2020) Comprehensive review and assessment of computational methods for predicting RNA post-transcriptional modification sites from RNA sequences. Brief Bioinform 21:1676–1696

Cozen AE, Quartley E, Holmes AD et al (2015) ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. Nat Methods 12:879–884

Dai Q, Fong R, Saikia M et al (2007) Identification of recognition residues for ligation-based detection and quantitation of pseudouridine and N6-methyladenosine. Nucleic Acids Res 35:6322–6329

Dai Q, Moshitch-Moshkovitz S, Han D et al (2017) Nm-seq maps 2′-O-methylation sites in human mRNA with base precision. Nat Methods 14:695–698

Delatte B, Wang F, Ngoc LV et al (2016) RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. Science 351:282–285

Delaunay S, Frye M (2019) RNA modifications regulating cell fate in cancer. Nat Cell Biol 21:552–559

Dimitrova DG, Teysset L, Carre C (2019) RNA 2′-O-methylation (nm) modification in human diseases. Genes (Basel) 10

Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485:201–206
Dominissini D, Moshitch-Moshkovitz S, Salmon-Divon M et al (2013) Transcriptome-wide mapping of N(6)-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing. Nat Protoc 8:176–189
Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S et al (2016) The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. Nature 530:441–446
Dong ZW, Shao P, Diao LT et al (2012) RTL-P: a sensitive approach for detecting sites of 2′-O-methylation in RNA molecules. Nucleic Acids Res 40:e157
Edelheit S, Schwartz S, Mumbach MR et al (2013) Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5C within archaeal mRNAs. PLoS Genet 9:e1003602
Eisenberg E, Li JB, Levanon EY (2010) Sequence based identification of RNA editing sites. RNA Biol 7:248–252
Elliott BA, Ho HT, Ranganathan SV et al (2019) Modification of messenger RNA by 2′-O-methylation regulates gene expression in vivo. Nat Commun 10:3401
Enroth C, Poulsen LD, Iversen S et al (2019) Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing. Nucleic Acids Res 47:e126
Feng P, Ding H, Yang H et al (2017) iRNA-PseColl: identifying the occurrence sites of different RNA modifications by incorporating collective effects of nucleotides into PseKNC. Mol Ther Nucleic Acids 7:155–163
Garcia-Campos MA, Edelheit S, Toth U et al (2019) Deciphering the "m(6)A code" via antibody-independent quantitative profiling. Cell 178(731–747):e716
Giessing AM, Kirpekar F (2012) Mass spectrometry in the biology of RNA and its modifications. J Proteome 75:3434–3449
Gilbert WV, Bell TA, Schaening C (2016) Messenger RNA modifications: form, distribution, and function. Science 352:1408–1412
Grosjean H, Keith G, Droogmans L (2004) Detection and quantification of modified nucleotides in RNA using thin-layer chromatography. Methods Mol Biol 265:357–391
Grosjean H, Droogmans L, Roovers M et al (2007) Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. Methods Enzymol 425:55–101
Grozhik AV, Linder B, Olarerin-George AO et al (2017) Mapping m(6)A at individual-nucleotide resolution using crosslinking and immunoprecipitation (miCLIP). Methods Mol Biol 1562:55–78
Gustafsson C, Persson BC (1998) Identification of the rrmA gene encoding the 23S rRNA m1G745 methyltransferase in Escherichia coli and characterization of an m1G745-deficient mutant. J Bacteriol 180:359–365
Hengesbach M, Meusburger M, Lyko F et al (2008) Use of DNAzymes for site-specific analysis of ribonucleotide modifications. RNA 14:180–187
Huang Y, He N, Chen Y et al (2018) BERMP: a cross-species classifier for predicting m(6)A sites by integrating a deep learning algorithm and a random forest approach. Int J Biol Sci 14:1669–1677
Hussain S, Sajini AA, Blanco S et al (2013) NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep 4:255–261
Incarnato D, Anselmi F, Morandi E et al (2017) High-throughput single-base resolution mapping of RNA 2-O-methylated residues. Nucleic Acids Res 45:1433–1441
Jenjaroenpun P, Wongsurawat T, Wadley TD et al (2021) Decoding the epitranscriptional landscape from native RNA sequences. Nucleic Acids Res 49:e7
Jia CZ, Zhang JJ, Gu WZ (2016) RNA-MethylPred: a high-accuracy predictor to identify N6-methyladenosine in RNA. Anal Biochem 510:72–75
Jiang T, Yu N, Kim J et al (2019) Oligonucleotide sequence mapping of large therapeutic mRNAs via parallel ribonuclease digestions and LC-MS/MS. Anal Chem 91:8500–8506

Jonkhout N, Tran J, Smith MA et al (2017) The RNA modification landscape in human disease. RNA 23:1754–1769

Jora M, Lobue PA, Ross RL et al (2019) Detection of ribonucleoside modifications by liquid chromatography coupled with mass spectrometry. Biochim Biophys Acta Gene Regul Mech 1862:280–290

Kadumuri RV, Janga SC (2018) Epitranscriptomic code and its alterations in human disease. Trends Mol Med 24:886–903

Kellner S, Burhenne J, Helm M (2010) Detection of RNA modifications. RNA Biol 7:237–247

Khoddami V, Cairns BR (2013) Identification of direct targets and modified bases of RNA cytosine methyltransferases. Nat Biotechnol 31:458–464

Khoddami V, Yerra A, Mosbruger TL et al (2019) Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. Proc Natl Acad Sci USA 116:6784–6789

Kirpekar F, Hansen LH, Rasmussen A et al (2005) The archaeon Haloarcula marismortui has few modifications in the central parts of its 23S ribosomal RNA. J Mol Biol 348:563–573

Lakings DB, Gehrke CW (1971) Analysis of base composition of RNA and DNA hydrolysates by gas-liquid chromatography. J Chromatogr 62:347–367

Lapham J, Crothers DM (2000) Site-specific cleavage of transcript RNA. Methods Enzymol 317:132–139

Larsen MR, Roepstorff P (2000) Mass spectrometric identification of proteins and characterization of their post-translational modifications in proteome analysis. Fresenius J Anal Chem 366:677–690

Li X, Zhu P, Ma S et al (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol 11:592–597

Li GQ, Liu Z, Shen HB et al (2016) TargetM6A: identifying N(6)-methyladenosine sites from RNA sequences via position-specific nucleotide propensities and a support vector machine. IEEE Trans Nanobioscience 15:674–682

Li X, Peng J, Yi C (2017a) Transcriptome-wide mapping of N (1)-methyladenosine methylome. Methods Mol Biol 1562:245–255

Li X, Xiong X, Zhang M et al (2017b) Base-resolution mapping reveals distinct m(1)a methylome in nuclear- and mitochondrial-encoded transcripts. Mol Cell 68:993–1005 e1009

Li J, Huang Y, Yang X et al (2018) RNAm5Cfinder: a web-server for predicting RNA 5-methylcytosine (m5C) sites based on random forest. Sci Rep 8:17299

Limbach PA, Paulines MJ (2017) Going global: the new era of mapping modifications in RNA. Wiley Interdiscip Rev RNA 8

Linder B, Grozhik AV, Olarerin-George AO et al (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods 12:767–772

Liu N, Pan T (2016) Probing N(6)-methyladenosine (m(6)A) RNA modification in total RNA with SCARLET. Methods Mol Biol 1358:285–292

Liu N, Dai Q, Zheng G et al (2015) N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature 518:560–564

Liu Z, Xiao X, Yu DJ et al (2016) pRNAm-PC: predicting N(6)-methyladenosine sites in RNA sequences via physical-chemical properties. Anal Biochem 497:60–67

Liu H, Begik O, Lucas MC et al (2019) Accurate detection of m(6)A RNA modifications in native RNA sequences. Nat Commun 10:4079

Liu J, Dou X, Chen C et al (2020) N (6)-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. Science 367:580–586

Lorenz DA, Sathe S, Einstein JM et al (2020) Direct RNA sequencing enables m(6)A detection in endogenous transcript isoforms at base-specific resolution. RNA 26:19–28

Macon JB, Wolfenden R (1968) 1-Methyladenosine. Dimroth rearrangement and reversible reduction. Biochemistry 7:3453–3458

Mahmoudi O, Wahab A, Chong KT (2020) iMethyl-deep: N6 methyladenosine identification of yeast genome with automatic feature extraction technique by using deep learning algorithm. Genes (Basel) 11

Marchand V, Blanloeil-Oillo F, Helm M et al (2016) Illumina-based RiboMethSeq approach for mapping of 2′-O-Me residues in RNA. Nucleic Acids Res 44:e135

Marchand V, Ayadi L, Ernst FGM et al (2018) AlkAniline-Seq: profiling of m(7) G and m(3) C RNA modifications at single nucleotide resolution. Angew Chem 57:16785–16790

Meyer KD (2019) DART-seq: an antibody-free method for global m(6)A detection. Nat Methods 16:1275–1280

Molinie B, Wang J, Lim KS et al (2016) m(6)A-LAIC-seq reveals the census and complexity of the m(6)A epitranscriptome. Nat Methods 13:692–698

Motorin Y, Marchand V (2018) Detection and analysis of RNA ribose 2′-O-methylations: challenges and solutions. Genes (Basel) 9

Motorin Y, Muller S, Behm-Ansmant I et al (2007) Identification of modified residues in RNAs by reverse transcription-based methods. Methods Enzymol 425:21–53

Motorin Y, Lyko F, Helm M (2010) 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. Nucleic Acids Res 38:1415–1430

Ofengand J, Del Campo M, Kaya Y (2001) Mapping pseudouridines in RNA molecules. Methods 25:365–373

Pandolfini L, Barbieri I, Bannister AJ et al (2019) METTL1 promotes let-7 microRNA processing via m7G methylation. Mol Cell 74:1278–1290 e1279

Pollex T, Hanna K, Schaefer M (2010) Detection of cytosine methylation in RNA using bisulfite sequencing. Cold Spring Harb Protoc 2010:pdb prot5505

Qiang X, Chen H, Ye X et al (2018) M6AMRFS: robust prediction of N6-methyladenosine sites with sequence-based features in multiple species. Front Genet 9:495

Qiu WR, Jiang SY, Xu ZC et al (2017) iRNAm5C-PseDNC: identifying RNA 5-methylcytosine sites by incorporating physical-chemical properties into pseudo dinucleotide composition. Oncotarget 8:41178–41188

Roberts RJ, Carneiro MO, Schatz MC (2017) Erratum to: the advantages of SMRT sequencing. Genome Biol 18:156

Saikia M, Dai Q, Decatur WA et al (2006) A systematic, ligation-based approach to study RNA modifications. RNA 12:2025–2033

Santoro SW, Joyce GF (1997) A general purpose RNA-cleaving DNA enzyme. Proc Natl Acad Sci USA 94:4262–4266

Schaefer M, Pollex T, Hanna K et al (2009) RNA cytosine methylation analysis by bisulfite sequencing. Nucleic Acids Res 37:e12

Schwartz S, Agarwala SD, Mumbach MR et al (2013) High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell 155:1409–1421

Schwartz S, Bernstein DA, Mumbach MR et al (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell 159:148–162

Shi H, Chai P, Jia R et al (2020) Novel insight into the regulatory roles of diverse RNA modifications: re-defining the bridge between transcription and translation. Mol Cancer 19:78

Shu X, Cao J, Cheng M et al (2020) A metabolic labeling method detects m(6)A transcriptome-wide at single base resolution. Nat Chem Biol 16:887–895

Song J, Zhai J, Bian E et al (2018) Transcriptome-wide annotation of m(5)C RNA modifications using machine learning. Front Plant Sci 9:519

Stoddart D, Maglia G, Mikhailova E et al (2010) Multiple base-recognition sites in a biological nanopore: two heads are better than one. Angew Chem 49:556–559

Strobel EJ, Yu AM, Lucks JB (2018) High-throughput determination of RNA structures. Nat Rev Genet 19:615–634

Tan MH, Li Q, Shanmugam R et al (2017) Dynamic landscape and regulation of RNA editing in mammals. Nature 550:249–254

Tanzer A, Hofacker IL, Lorenz R (2019) RNA modifications in structure prediction – status quo and future challenges. Methods 156:32–39
Toh SM, Xiong L, Bae T et al (2008) The methyltransferase YfgB/RlmN is responsible for modification of adenosine 2503 in 23S rRNA. RNA 14:98–106
Uddin MB, Wang Z, Yang C (2020) Dysregulations of functional RNA modifications in cancer, cancer stemness and cancer therapeutics. Theranostics 10:3164–3189
Uziel M, Koh CK, Cohn WE (1968) Rapid ion-exchange chromatographic microanalysis of ultraviolet-absorbing materials and its application to nucleosides. Anal Biochem 25:77–98
Wang X, Yan R (2018) RFAthM6A: a new tool for predicting m(6)A sites in *Arabidopsis thaliana*. Plant Mol Biol 96:327–337
Wang X, Lu Z, Gomez A et al (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505:117–120
Wang Y, Xiao Y, Dong S et al (2020) Antibody-free enzyme-assisted chemical approach for detection of N(6)-methyladenosine. Nat Chem Biol 16:896–903
Wei L, Chen H, Su R (2018) M6APred-EL: a sequence-based predictor for identifying N6-methyladenosine sites using ensemble learning. Mol Ther Nucleic Acids 12:635–644
Worpenberg L, Jakobi T, Dieterich C et al (2019) Identification of methylated transcripts using the TRIBE approach. Methods Mol Biol 1870:89–106
Wulff TF, Arguello RJ, Molina Jordan M et al (2017) Detection of a subset of posttranscriptional transfer RNA modifications in vivo with a restriction fragment length polymorphism-based method. Biochemistry 56:4029–4038
Xiang S, Liu K, Yan Z et al (2016a) RNAMethPre: a web server for the prediction and query of mRNA m6A sites. PLoS One 11:e0162707
Xiang S, Yan Z, Liu K et al (2016b) AthMethPre: a web server for the prediction and query of mRNA m(6)A sites in *Arabidopsis thaliana*. Mol BioSyst 12:3333–3337
Xing P, Su R, Guo F et al (2017) Identifying N(6)-methyladenosine sites using multi-interval nucleotide pair position specificity and support vector machine. Sci Rep 7:46757
Yu N, Jora M, Solivio B et al (2019) tRNA modification profiles and codon-decoding strategies in *Methanocaldococcus jannaschii*. J Bacteriol 201
Zeng Y, Wang S, Gao S et al (2018) Refined RIP-seq protocol for epitranscriptome analysis with low input materials. PLoS Biol 16:e2006092
Zhang M, Sun JW, Liu Z et al (2016) Improving N(6)-methyladenosine site prediction with heuristic selection of nucleotide physical-chemical properties. Anal Biochem 508:104–113
Zhang M, Xu Y, Li L et al (2018) Accurate RNA 5-methylcytosine site prediction based on heuristic physical-chemical properties reduction and classifier ensemble. Anal Biochem 550:41–48
Zhang Z, Chen LQ, Zhao YL et al (2019a) Single-base mapping of m(6)A by an antibody-independent method. Sci Adv 5:eaax0250
Zhang LS, Liu C, Ma H et al (2019b) Transcriptome-wide mapping of internal N(7)-methylguanosine methylome in mammalian mRNA. Mol Cell 74:1304–1316 e1308
Zhao X, Yu YT (2004) Detection and quantitation of RNA base modifications. RNA 10:996–1002
Zhao Z, Peng H, Lan C et al (2018) Imbalance learning for the prediction of N(6)-methylation sites in mRNAs. BMC Genomics 19:574
Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29
Zhou Y, Zeng P, Li YH et al (2016) SRAMP: prediction of mammalian N6-methyladenosine (m6A) sites based on sequence-derived features. Nucleic Acids Res 44:e91
Zhu Y, Pirnie SP, Carmichael GG (2017) High-throughput and site-specific identification of 2′-O-methylation sites using ribose oxidation sequencing (RibOxi-seq). RNA 23:1303–1314
Zhu W, Wang JZ, Xu Z et al (2019) Detection of N6methyladenosine modification residues (review). Int J Mol Med 43:2267–2278
Zou Q, Xing P, Wei L et al (2019) Gene2vec: gene subsequence embedding for prediction of mammalian N (6)-methyladenosine sites from mRNA. RNA 25:205–218

Regulation of RNA Stability Through RNA Modification

Kentaro Kawata and Nobuyoshi Akimitsu

Contents

Abstract As intracellular RNA levels are determined by the balance between RNA synthesis and decay, RNA stability is a critical factor in the quantity of intracellular RNA. Recent advances in epitranscriptomic measuring techniques have shown that RNA modifications, such as methylation and acetylation, affect RNA stability, including those of messenger RNA (mRNA) and noncoding RNA (ncRNA) such as transfer RNA (tRNA) and ribosomal RNA (rRNA). In addition, some viruses infected into cells avoid innate immunogenicity through the RIG-I signaling pathway by modifying their genomic RNAs. In this chapter, we focus on epitranscriptomic regulation for RNA stability. These epitranscriptomic regulations of RNA stability have also been utilized in commercial applications, such as development of gene therapy agents. This chapter also addresses biotechnological

K. Kawata · N. Akimitsu (✉)
Isotope Science Center, The University of Tokyo, Tokyo, Japan
e-mail: akimitsu@ric.u-tokyo.ac.jp

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_8

applications of RNA stabilization that use modified nucleosides and measurement technologies that evaluate RNA stability.

Keywords Regulation of RNA stability · Messenger RNA · Transfer RNA · Ribosomal RNA · Innate immunity · Measurement of RNA metabolism

Abbreviations

8-oxoG	8-oxo-7,8-dihydroguanosine
Ψ	Pseudouridine
ac^4C	N4-acetylcytidine
ADAR	Adenosine deaminase acting on RNA
Ago	Argonaut
ALKBH5	ALK B homolog-5
A_m	2′-*O*-methyladenosine
A-to-I editing	Adenosine to inosine editing
box H/ACA RNP	box H/ACA snoRNA-protein complex
BRIC-seq	BrU immunoprecipitation chase-deep sequencing analysis
BrU	5-bromouridine
CBC	Cap-binding complex
CBLL1	Cbl proto-oncogene-like protein-1
DNMT2	DNA methyltransferase 2
E^5U	5-ethynyluridine
eIF4E	Eukaryotic initiation factor 4E
FTO	Fat mass and obesity-associated protein
GTase	RNA guanylyltransferase
GTP	Guanosine triphosphate
guanine-N7 MTase	Guanine-N7 methyltransferase
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
hMPV	Human metapneumovirus
HRSP12	Heat-responsive protein 12
IAV	Influenza A virus
IFN	Interferon
IL-6	Interleukin 6
LLPS	Liquid–liquid phase separation
lncRNA	Long noncoding RNA
LPS	Lipopolysaccharide
$m^1\Psi$	N1-methylpseudouridine
$m^2{}_2G$	N2,N2-dimethylguanosine
m^5C	C5-methylcytidine
m^6A	N6-methyladenosine
m^6A_m	N6,2′-*O*-dimethyladenosine

m^7G	N7-methylguanosine
METTL3	Methyltransferase-like protein 3
miRNA	MicroRNA
mRNA	Messenger RNA
NAT10	N-acetyltransferase 10
ncRNA	Noncoding RNA
NK cell	Natural killer cell
NMD	Nonsense-mediated decay
NSUN	NOL1/NOP2/SUN domain
OsO4	Osmium tetroxide
PA	Polymerase acidic protein
PAMP	Pathogen-associated molecular pattern
PB2	Polymerase base protein 2
P-body	Processing body
PIN	PilT N-terminus like
polyA	PolyA tail
PRR	Pattern recognition receptor
PUS	Pseudouridine synthases
RAT1	Ribonucleic-acid-trafficking protein 1
RBP	RNA-binding protein
RIG-I	Retinoic acid-inducible gene-I
RISC	RNA-induced silencing complex
RNase	Ribonuclease
rRNA	Ribosomal RNA
RTD pathway	Rapid tRNA decay pathway
s^4U	4-thiouridine
s^6G	6-thioguanosine
SAM	*S*-adenosylmethionine
SLAM-seq	Thiol (SH)-linked alkylation for metabolic RNA sequencing
snRNA	Small nuclear RNA
TLR	Toll-like receptor
TPase	RNA triphosphatase
tRNA	Transfer RNA
TUC-seq	Thiouridine-to-cytidine sequencing
UPF1	Up-frameshift-1
WTAP	Wilms' tumor-associated protein
XRN1	Exoribonuclease 1
YBX1	Y-box binding protein-1
YTHD	YTH domain-containing proteins

1 Introduction

Regulation of intracellular RNA levels is a fundamental factor in cellular states. In mammalian cells, RNA is generated as a transcript of genes in the cell nucleus. The transcribed immature RNAs are processed post-transcriptionally or co-transcriptionally. In many cases, mature RNAs are transported from the nucleus into cytoplasm to regulate cellular functions as messenger RNAs (mRNAs), long noncoding RNAs (lncRNAs), transfer RNAs (tRNAs), or ribosomal RNAs (rRNAs). Strict regulation of RNA degradation is important for various biological processes (Keene 2010; Schott and Stoecklin 2010; Alonso 2012). For example, stabilization of certain RNAs in response to cellular states, such as activation of cytokine receptors and intracellular signal induced by innate immunity, enables RNA levels to increase rapidly (Schott and Stoecklin 2010; Imamura et al. 2018; Shirahama et al. 2020). Regulation of RNA degradation is also a factor in RNA quality control (Lykke-Andersen and Jensen 2015) and is central to the post-transcriptional control of gene expression.

RNA degradation studies have mainly focused on RNA-binding proteins (RBPs) and miRNAs. However, recent studies revealed that the fate of RNAs, including their degradation, is significantly affected by reversible or irreversible chemical modifications. Such chemical modifications in whole RNA are called epitranscriptome. In this chapter, we first summarize known mechanisms of RNA degradation, then describe epitranscriptomic modifications thought to affect RNA degradation, and finally, discuss real-world applications for these RNA modifications.

2 Fundamentals of Intracellular RNA Stability

Intracellular RNAs are mainly degraded by enzymes called ribonucleases (RNases). They consist of three main classes: 5′ exonucleases, which hydrolyze RNA from the 5′ ends; 3′ exonucleases, which hydrolyze RNA from the 3′ ends; and endonucleases, which cleave RNA internally (Mathy et al. 2007; De La Sierra-Gallay et al. 2008; Houseley and Tollervey 2009). Usually, eukaryotic mRNAs and parts of lncRNAs are protected from nucleases by 5′-cap structure and 3′-polyA tail (polyA). A degradation signal initiates RNA degradation via one of the following pathways: a deadenylation-independent decay pathway that starts with removal of the 5′-cap; a deadenylation-dependent degradation pathway that starts by shortening of polyA; and an endonuclease-mediated decay pathway initiated by internal cleavage of the RNA. The deadenylation-dependent decay pathway degrades many eukaryotic mRNAs or lncRNAs. For example, many RBPs, such as CCR4-NOT (which regulates degradation), promote deadenylation of target RNAs (Lau et al. 2009). The RNAs deadenylated by these deadenylation-dependent decay pathways aggregate and are captured by cytoplasmic granules, such as processing bodies

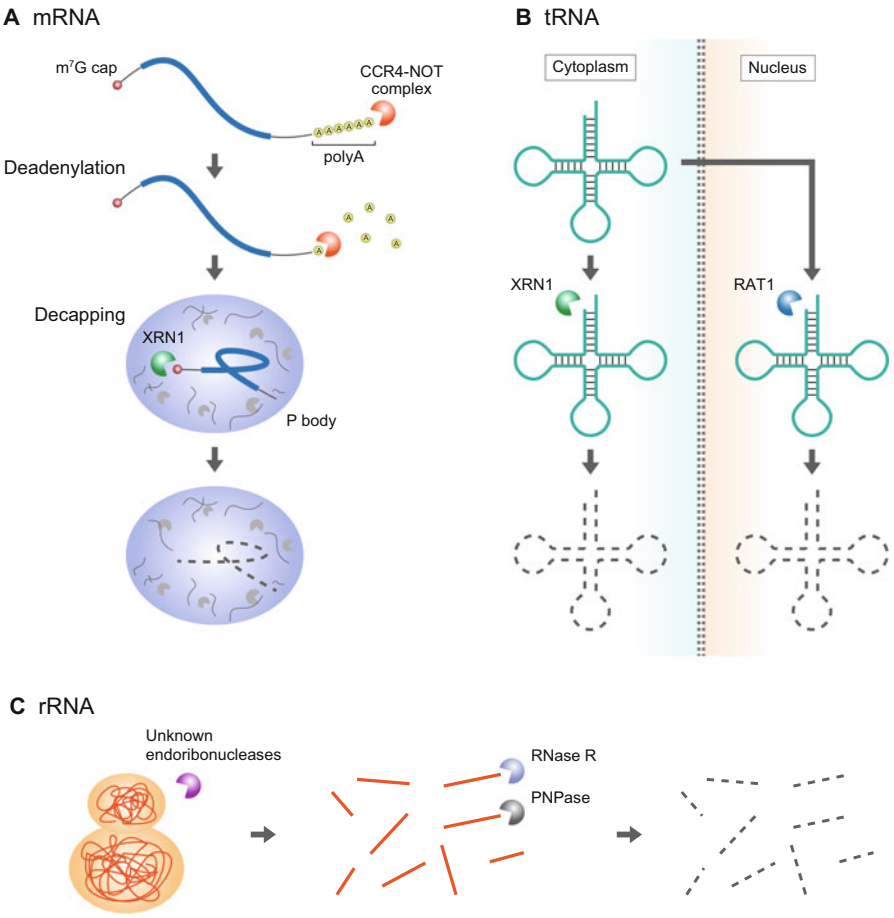


Fig. 1 Schemes of degradation. (**a**) Degradation of mRNA and poly-A-tailed lncRNA usually starts with deadenylation of poly-A-tail promoted by RBPs such as CCR4-NOT. The deadenylated RNAs aggregate and are captured by cytoplasmic granules, such as P-bodies. In which, the 5′-cap of RNAs is removed, and degraded by exonucleases such as XRN1. (**b**) In eukaryotic cells, the tRNA is rapidly degraded by RTD pathway that includes XRN1, which is initiated by exposing the 5′ end. Parts of tRNAs are re-transported to the nucleus and degraded by RAT1. (**c**) The rRNAs damaged by chemicals and UV are first cleaved by endoribonucleases to fragments. The rRNA fragments are further degraded to mononucleotides by exoribonucleases such as RNase R and PNPase

(P-bodies), to be degraded by exonucleases such as exoribonuclease 1 (XRN1). XRN1 is a representative component of the deadenylation-independent decay pathways (Nagarajan et al. 2013), which causes hydrolysis from the 5′ end of RNA after the 5′-cap has been removed (Fig. 1a). RNA silencing, caused by the RNA-induced silencing complex (RISC), is guided by microRNA (miRNA) and is also an important RNA degradation mechanism. The miRNAs are small (20–25 nucleotides), single-stranded RNAs that are generated by incorporation of pre-miRNAs into an

Argonaute (Ago) protein after being cleaved by the endoribonuclease Dicer in the cytoplasm (Ghildiyal and Zamore 2009). In mammalian cells, the Ago family consists of four types (Ago1–4); Ago2 is considered to have RNA cleavage activity (Hutvagner and Simard 2008; Kawamata and Tomari 2010). The RISC is a complex of Ago and miRNA, and often cleaves target mRNAs in regions that have imperfect homologous sequences with the incorporated miRNAs (Tomari and Zamore 2005; Hutvagner and Simard 2008; Ghildiyal and Zamore 2009). The tRNAs are relatively stable; their half-lives are thought to be a few days (Phizicky and Hopper 2010). In eukaryotic cells, the rapid tRNA decay (RTD) pathway includes XRN1 and degrades mature tRNAs (Megel et al. 2015). The RTD pathway is initiated by exposing the 5′ end of the tRNA. Some tRNA is re-transported to the nucleus and degraded by ribonucleic-acid-trafficking protein 1 (RAT1) (Krzyszton et al. 2012), but its regulation is not well understood (Fig. 1b). Although rRNAs are also relatively stable, they are susceptible to damage by chemicals and UV. The nonfunctional rRNA decay pathway is a frequently studied mechanism for rRNAs degradation. Damaged rRNAs are first cleaved by endoribonucleases, and the resulting rRNA fragments are further degraded to mononucleotides by exoribonucleases such as RNase R and PNPase (Basturea et al. 2011). The rRNA fragments cannot degrade in the absence of these exoribonucleases and therefore form large accumulations. However, rRNAs are degraded in starved cells, mainly by exoribonucleases such as RNase II and RNase R (Basturea et al. 2011). Although the major exoribonucleases involved in these pathways have been identified, the type of endoribonuclease and its cleavage site remain unknown (Fig. 1c).

Inflammatory cytokine-coding mRNAs are an important target of the RNA degradation mechanisms. Cells use sensor protein complexes called pattern recognition receptors (PRRs) to recognize common molecular patterns, called pathogen-associated molecular patterns (PAMPs), which are required for pathogen survival and infection (Takeuchi and Akira 2010). Retinoic acid-inducible gene-I (RIG-I), RIG-I-like receptor MDA5 (see Sect. 4.1), and Toll-like receptors (TLRs) are well-known PRRs. These PRRs recognize their ligands, such as genomic RNAs from viruses that have infected the cell, and activate downstream signals to induce inflammatory cytokines such as interleukin 6 (IL-6). As these inflammatory cytokines induce apoptosis and attract cytotoxic cells such as natural killer cells (NK cells), quantities of these cytokines are strictly regulated at the RNA level. These inflammatory cytokine-coding RNAs are unstable and their half-lives are reportedly quite short (Hao and Baltimore 2009), which indicates that their quantities are rapidly regulated (Kawata et al. 2020). Regnase-1 and Roquin, which are endonucleases, are key regulators of inflammatory cytokine-coding RNAs. Regnase-1 is a cytoplasmic protein with a CCCH-type zinc-finger domain and a PilT N-terminus-like (PIN) domain that functions as an RNA-degrading enzyme; it degrades target RNAs by binding to them through the zinc-finger and PIN domains (Uehata and Takeuchi 2017). Regnase-1 recognizes the stem-loop structure in the 3′ UTR of target RNAs and associates with up-frameshift-1 (UPF1) (Mino et al. 2015)—a central element of nonsense-mediated decay (NMD), which targets translated RNAs (Wei-Lin Popp and Maquat 2013). Roquin includes RING, ROQ, and

zinc-finger domains (Schlundt et al. 2014; Tan et al. 2014). The RNA degradation by Roquin is induced by recognition of stem-loops formed by the constitutive decay element in 3′ UTR of target RNAs. Reportedly, when Roquin recognizes stem-loops, it recruits the CCR4-NOT complex to promote deadenylation (Leppek et al. 2013), and removes the 5′-cap by association with Edc4 and Rck (Glasmacher et al. 2010). Although Regnase-1 and Roquin recognize common stem-loop structures, Regnase-1 mainly localizes to the endoplasmic reticulum, thus co-localizing with ribosomes, whereas Roquin mainly localizes among stress granules and P-bodies. Moreover, lipopolysaccharide (LPS) stimulation showed that Regnase-1 and Roquin function in the early and late stages of inflammation, respectively. These results suggest that Regnase-1 and Roquin regulate overlapping targets spatiotemporally to regulate inflammatory cytokines precisely (Mino et al. 2015).

3 RNA Modifications that Affect Stability

Modifications of intracellular RNAs can contribute to their degradation (Fig. 2) (Boo and Kim 2020). According to the MODOMICS database, which registers RNA modification pathways (Boccaletto et al. 2018), more than 170 modifications have been identified for mRNAs and noncoding RNAs (ncRNAs), including tRNAs and rRNAs (Helm and Motorin 2017; Nachtergaele and He 2018). For example, N6-methyladenosine (m^6A) is the best-known and apparently most common methylation modification in various eukaryotic RNA species. C5-methylcytidine (m^5C) is also a common RNA modification. RNA methylation, as with m^6A and m^5C, affects RNA stability. In general, RNA is modified by a protein/complex called a "writer," de-modified by a protein/complex called an "eraser," and recognized by a specific protein/complex called a "reader."

One of the most versatile RNA modification factors is snoRNA, which is one of the ncRNAs locating in nucleolus and regulates RNA modification by directly forming Watson–Crick base pairs with the target RNA. The snoRNAs include two major families, box H/ACA snoRNA and box C/D snoRNA, which catalyze pseudouridination, which affects degradation of target RNAs (see Sect. 3.3), and 2′-*O*-ribose methylation, respectively. The snoRNA was previously considered to target rRNA and small nuclear RNA (snRNA). However, recent high-throughput approaches have revealed that snoRNA makes wide kinds of modifications to various RNA species including mRNA. For example, C/D snoRNA usually catalyzes 2′-*O*-ribose methylation, while its families in yeast, snR4 and snR45, catalyze the acetylation of cytosine residues in 18S rRNA (Sharma et al. 2017), while recent studies revealed that SNORD32A and SNORD97, which are members of C/D snoRNA, also catalyze 2′-*O*-ribose methylation of *Pxdn* mRNA encoding peroxidase and tRNA, respectively (Vitali and Kiss 2019; Elliott et al. 2019). However, the effects of these on RNA stability have not been clarified well.

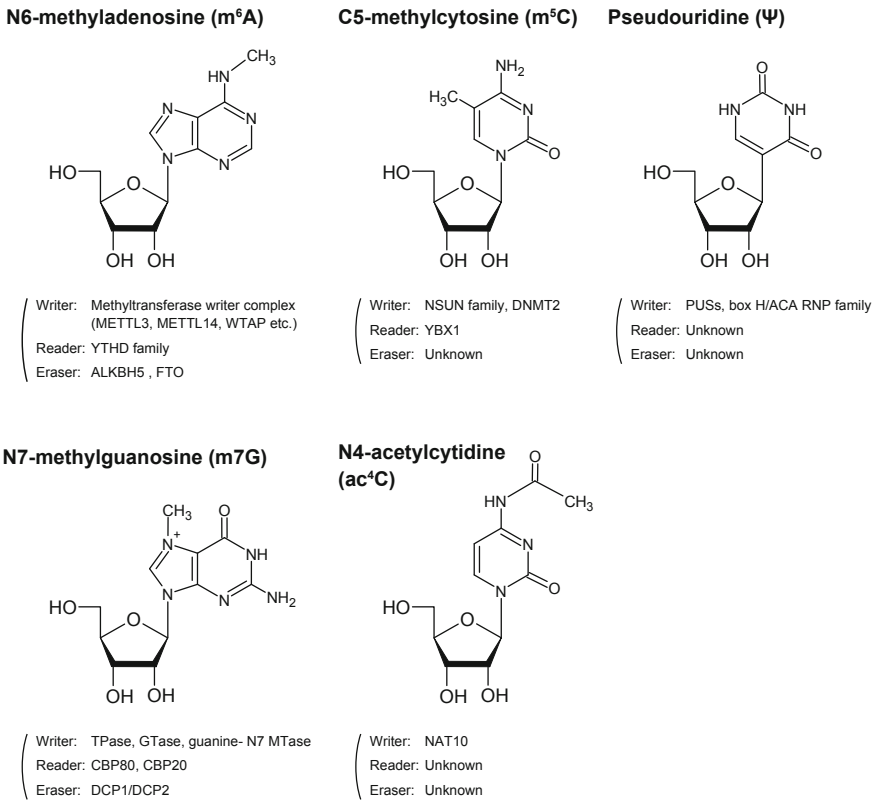


Fig. 2 Structures of representative modified nucleic acids that regulate RNA degradation. Each modification is regulated by specific modification (writer), recognition factor (reader), and de-modification factor (eraser)

3.1 N6-methyladenosine (m^6A)

3.1.1 Distribution and Function of m^6A

One well-studied stabilizing RNA modification is m^6A, which is present in various mammalian cells and accounts for more than 80% of all RNA modifications (Lee et al. 2020). It has also been identified in tRNAs and rRNAs and accounts for 0.1%–0.4% of adenosine residues in intracellular RNAs (Niu et al. 2013). Although the m^6A is found mainly in the mRNA 5′ UTR, the coding sequence, and the 3′ UTR (Zaccara et al. 2019), it is especially enriched around the translation stop codon and in the 3′ UTR (Dominissini et al. 2012; Meyer et al. 2012; Linder et al. 2015; Lee et al. 2020). Whereas m^6A regulates various processing and metabolic characteristics in mRNAs (Lee et al. 2020), it is rare in tRNAs and rRNAs. Reportedly, A1832 in 18S rRNA and A4190 in 28S rRNA can be methylated in human rRNAs

(Piekna-Przybylska et al. 2008), but its function in these ncRNAs is not well understood (Piekna-Przybylska et al. 2008).

3.1.2 Regulation of m^{6}A Modification

The m^{6}A modification is co-transcriptionally added to nascent RNAs in the nucleus by the methyltransferase writer complex (Fig. 3a). This complex recognizes DRACH motifs on RNA (where D = A, G, or U; R = purine; and H = A, C, or U) and introduces the m^{6}A into nascent RNA (Lee et al. 2020). However, only parts of RNAs are actually modified, and only a few potential m^{6}A consensus sites are methylated. How these sites are specified is poorly understood (Zaccara et al. 2019). The m^{6}A in mRNA and other polymerase II-derived RNA is mainly regulated by heterodimers formed by methyltransferase-like protein 3 (METTL3, also known as MT-A70) and METTL14. METTL3 is the central component of this enzymatic activity, and METTL14 increases the enzymatic activity of METTL3 by binding to the target RNA (Lee et al. 2020). The METTL3–METTL14 heterodimer is responsible for most m^{6}A modifications in mRNA. *METTL3* deletion experiments with mouse embryonic stem cell and inactivation of *METTL14* by CRISPR resulted

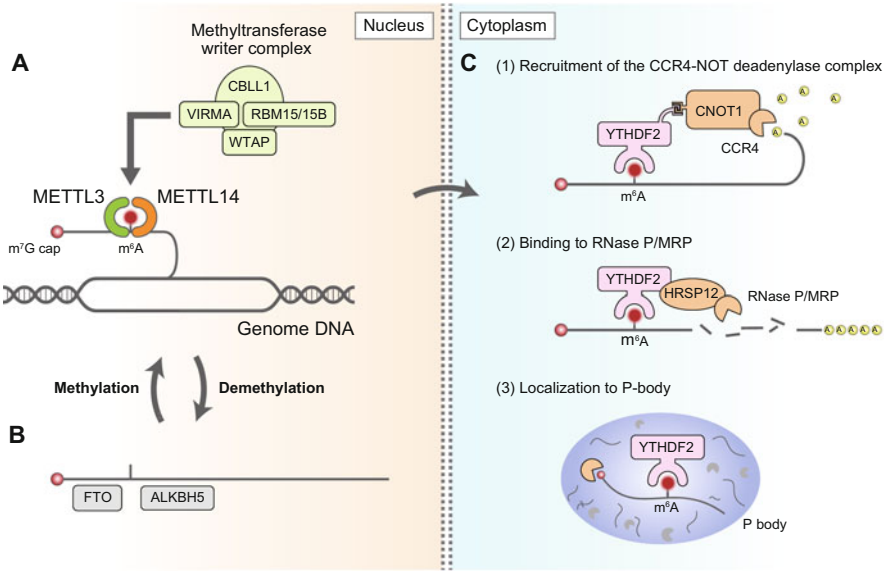


Fig. 3 Regulating RNA degradation by m^{6}A. (**a**) Introduction of m^{6}A modification. The m^{6}A modification is added by the m^{6}A writer complex, which includes METTL3, METTL14, WTAP, VIRMA, CBLL, and PBM/15/15B. METTL3 is the central enzymatic component, and METTL14 increases METTL3′s activity rate. (**b**) Removing m^{6}A modification. ALKBH5 and FTO act to demethylate m^{6}A modifications. (**c**) Degradation of m^{6}A-modified RNAs. YTHDF2 mediates degradation of m^{6}A-modified mRNA through any of three major pathways: recruitment of the CCR4–NOT deadenylase complex, binding to RNase P/MRP, and localization to the P-body

in the loss of more than 99% of total m^6A in polyA RNA (Geula et al. 2015). In addition to METTL3 and METTL14, this writer complex also includes other components, such as Wilms' tumor-associated protein (WTAP) which is a METTL3 adaptor (Zhong et al. 2008; Agarwala et al. 2012; Liu et al. 2014; Ping et al. 2014; Schwartz et al. 2014), VIRMA (also known as KIAA1429) which interacts with WATP (Horiuchi et al. 2013; Liu et al. 2014; Schwartz et al. 2014; Lence et al. 2016), RBM15/15B which affects the m^6A specificity of RBM (Horiuchi et al. 2013; Patil et al. 2016; Lence et al. 2016), and Cbl proto-oncogene-like protein-1 (CBLL1; also known as Hakai) (Lence et al. 2016). How these components interact with each other is unclear; however, they can receive a wide range of signals and may contribute to the strict regulation of RNA-specific methylation. RNA silencing machinery has been suggested to affect m^6A modification (Erson-Bensan and Begik 2017). The m^6A is enriched in regions recognized by miRNA on the 3′ UTR of target RNAs, and RISC-miRNA binds on METTL3 on RNA to regulate m^6A modifications (Chen et al. 2015). The m^6A "erasers," which demethylate m^6A, also participate in these changes. The m^6A erasers, α-ketoglutarate-dependent dioxygenase ALK B homolog-5 (ALKBH5) and fat mass and obesity-associated protein (FTO), are well known (Fig. 3b) (Fu et al. 2013; Zheng et al. 2013; Liu et al. 2014; Zhao et al. 2014). ALKBH5 demethylates m^6A with DRACH motifs, and FTO is a global eraser that also demethylates other RNA modifications, including N6,2′-*O*-dimethyladenosine (m^6A_m) (Wei et al. 2018).

3.1.3 Regulation of RNA Degradation with m^6A

The original report of RNA destabilization by the m^6A modification found lengthened mRNA half-lives in human and mouse cells upon suppression of METTL3 and WTAP (Liu et al. 2014; Schwartz et al. 2014; Batista et al. 2014). YTH domain-containing proteins (YTHD) family (YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2) is a major "reader" protein that destabilizes RNA via the m^6A modification (Liao et al. 2018). Although YTHDFs have three subtypes (YTHDF 1–3), all of which affect RNA degradation (Tirumuru et al. 2016; Shi et al. 2017; Lu et al. 2018), YTHDF2 is particularly considered to have central function in RNA degradation (Park et al. 2019).

Three major pathways have been suggested for YTHDF2-mediated degradation of m^6A-modified mRNA (Fig. 3c): recruitment of the CCR4-NOT deadenylase complex (Du et al. 2016); binding to RNase P/MRP (Park et al. 2019); and localization to P-bodies (Wang et al. 2014; Ries et al. 2019). The CCR4-NOT complex is a large protein complex that is highly conserved evolutionarily, from yeast to humans, and has been identified as responsible for mRNA deadenylation (Denis and Chen 2003; Collart and Timmers 2004; Lau et al. 2009). This complex is recruited to mRNA or polyA-added lncRNAs by multiple signals to initiate RNA degradation with deadenylation; YTHDF2 recognizes and directly interacts with m^6A in RNAs through the N-terminal region to the SH domain of CNOT1 which is a

component of the CCR4-NOT complex (Du et al. 2016). When recruited to the m^6A-modified RNA, the CCR4-NOT complex triggers RNA degradation by deadenylating the RNA (Jarrous 2017). RNase P/MRP is an RNP complex that mainly contributes to degradation of tRNAs, but also targets some mRNAs and lncRNAs. When YTHDF2 recognizes m^6A on RNAs, it is bridged by heat-responsive protein 12 (HRSP12; also known as reactive intermediate imine deaminase A homolog, UK114 antigen homolog, or 14.5 kDa translational inhibitor protein) to bind with RNase P/MRP. The RNase P/MRP recruited to the RNA causes endoribonucleolytic cleavage (Park et al. 2019). The HRSP12 enhances binding of YTHDF2 to mRNA, and the YTHDF2 enhances binding of HRSP12 to mRNA. Through these coordinated effects, HRSP12 promotes RNA degradation. YTHDF2 can also localize RNAs to intracellular granules, such as P-bodies, through liquid–liquid phase separation (LLPS) (Ries et al. 2019). P-bodies are characteristic granules in the cytoplasm, and many RNA processes take place within them. It is showed that YTHDF2-bound RNAs which contain m^6A are localized to P-bodies, depending on stresses such as heat shock, and the number of m^6A on the RNA (Ries et al. 2019). Fu et al. also showed that in oxidatively stressed cells, m^6A-modified mRNA localized to cytoplasmic stress granules, in a YTHDF protein-dependent manner (Fu and Zhuang 2019). Ries et al. also reported that *METTL14* knockout had no effect on stress granule formation after heat shock. Further studies are required to clarify the mechanism of m^6A-containing RNA degradation of YTHDF2 through LLPS (Liu et al. 2020).

3.2 C5-methylcytosine (m^5C)

3.2.1 Distribution and Function of m^5C

The study of m^5C initially began as an epigenetic modification of DNA (Suzuki and Bird 2008). However, many RNA species are found to be rich in m^5C, including many archaeal and eukaryotic tRNAs. This modification stabilizes secondary structures of tRNAs and inhibits their degradation (Helm 2006; Agris 2008; Anderson and Wang 2009; Motorin and Helm 2010; Schaefer et al. 2010; Squires and Preiss 2010). Although rRNAs are also rich in m^5C (Chow et al. 2007), whether m^5C affects their stability is unclear. Improved and more comprehensive m^5C detection techniques that use bisulfite sequencing (Lister et al. 2009) have identified m^5C modifications in mRNAs with high resolution (Squires et al. 2012). The mRNAs with no specific m^5C modification sites and the function of m^5C in mRNA and lncRNA still require clarification.

3.2.2 Regulation of m^5C Modification

The m^5C modification is mediated by two types of methyltransferases: the NOL1/NOP2/SUN domain (NSUN) family and DNA methyltransferase 2 (DNMT2) which is a homolog of DNA methyltransferase (Trixl and Lusser 2019; Bohnsack et al. 2019; Boo and Kim 2020). The NSUN family catalyzes the methylation of diverse RNA species and consists of NSUN1, NSUN2, NSUN3, NSUN4, NSUN5, NSUN6, and NSUN7 in mammals (Trixl and Lusser 2019). NSUN1, NSUN2, and NSUN5 are conserved throughout eukaryotes, but the remaining NSUN proteins are present only in higher eukaryotes, including humans. NSUN proteins are *S*-adenosylmethionine (SAM)-dependent methyltransferases and include similar structural cores, including SAM-binding sites and catalytic domains (Liu and Santi 2000; Cheng and Roberts 2001; King and Redman 2002). Each member of the NSUN family catalyzes the methylation of unique substrates (Trixl and Lusser 2019). NSUN1, NSUN4, and NSUN5 recognize unique cytidines on cytosolic and mitochondrial rRNAs as substrates. NSUN3 and NSUN6 are methyltransferases that recognize cytosolic and mitochondrial tRNA as substrates, as well as DNMT2 (described below). NSUN2 has the most versatile target specificity among the NSUN family, and recognizes cytidines in all mRNAs, tRNAs, and rRNAs. Although DNMT2 was originally thought to be a DNA methyltransferase (Dong et al. 2001), it is reported to be an RNA methyltransferase that recognizes C38 on several tRNAs as substrates (Goll et al. 2006). Over 90% of tRNA methylation was reduced in *Dnmt2*$^{-/-}$/*Nsun2*$^{-/-}$ double-knockout mice (Blanco et al. 2014), which suggests that most m^5C in tRNAs is a substrate for NSUN2 and DNMT2. However, its m^5C eraser has not been identified (Boo and Kim 2020); whether this modification is a reversible process is unclear.

3.2.3 Regulation of RNA Degradation with m^5C

At the molecular level, m^5C enhances the thermal stability of hydrogen bonding with guanine to stabilize the RNA structure (Hayrapetyan et al. 2009; Motorin and Helm 2010). Many studies of m^5C have focused on tRNAs. This modification occurs most often at a few cytidines that span positions 47–50 in tRNAs and increase the hydrophobicity of the base pairs to stabilize the interaction (Väre et al. 2017). A highly susceptible position for this modification is C38 of tRNA, which increases the accuracy and efficiency of translation, and stabilizes tRNAs by protecting them from stress-induced endonuclease-mediated fragmentation (Schaefer et al. 2010; Tuorto et al. 2012). The rRNAs are also rich in m^5C modifications, which are highly conserved evolutionarily (Motorin and Helm 2010). For example, methylation of C2870 and C2278 is conserved in humans and yeast (Schaefer et al. 2009; Motorin and Helm 2010; Sharma et al. 2013; Schosserer et al. 2015; Bourgeois et al. 2015). In contrast, m^5C has not been detected in eukaryotic 18S RNA (Edelheit et al. 2013). Although m^5C seems to stabilize the RNA structure by promoting base stacking, this

relationship has not reported between m^5C modification and rRNA degradation regulation. m^5C helps regulation of mRNA degradation. It was showed that downregulating NSUN2 reduced the quantity and half-life of $p16^{INK4}$ mRNA, which suggests that NSUN2 stabilizes $p16^{INK4}$ mRNA (Zhang et al. 2012). In addition, m^5C shares its Y-box binding protein-1 (YBX1) as a reader protein with 8-oxo-7,8-dihydroguanosine (8-oxoG) (Chen et al. 2019; Yang et al. 2019). The m^5C-recognizing YBX1 can contribute to mRNA stabilization, by recruiting HuR, a stabilizing RBP (Chen et al. 2019).

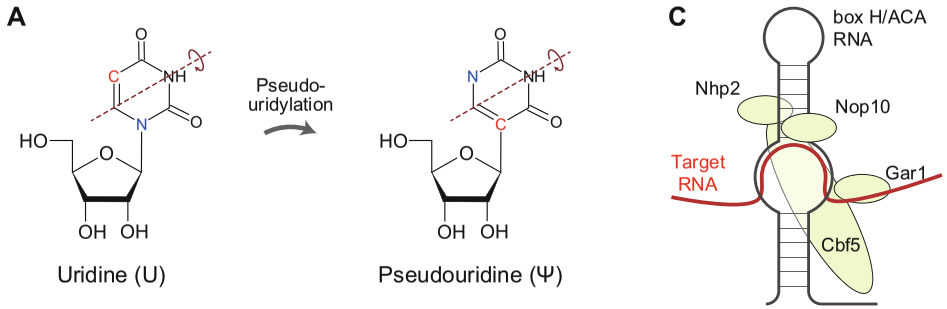
3.3 Pseudouridine (Ψ)

3.3.1 Distribution and Function of Ψ

Ψ was the earliest modified base, discovered in 1951, and is also called the "fifth nucleotide." It was initially found in rRNAs and tRNAs, but recent transcriptome-wide analyses of human and yeast revealed that many human and yeast mRNAs contain Ψ modification (Schwartz et al. 2014; Carlile et al. 2014; Li et al. 2015). Thus, Ψ is one of the most common modified bases, and found in all RNA and species (Charette and Gray 2000; Liang et al. 2002; Hamma and Ferré-D'Amaré 2006).

3.3.2 Regulation of Ψ Modification

Ψ is an isomer of uridine, in which the uracil is attached to ribose ring via a C–C bond instead of an N–C glycosidic bond (Fig. 4a). Two pathways for converting uridine to Ψ are known—the RNA-independent and RNA-dependent pathways (Penzo et al. 2017; Adachi et al. 2019). The RNA-independent pathway uses a stand-alone enzyme called pseudouridine synthases (PUS). In eukaryotes, 10 identified PUSs have been classified into five families (TruA, TruB, RluA, TruD, and Pus10) (Fig. 4b) (Hamma and Ferré-D'Amaré 2006). All PUSs share a common catalytic core domain with an aspartic acid residue that is essential for catalytic reactions. They recognize specific substrates and catalyze chemical reactions. For example, PUS3, a member of the TruA family, specifically modifies tRNAs U38 and U39 in the cytoplasm and mitochondria. RNA-dependent pseudo-uridination is catalyzed by the box H/ACA snoRNA–protein complex (box H/ACA RNP) family (see Sect. 3) (Fig. 4c) (Yu and Meier 2014). These RNPs consist of one unique box H/ACA snoRNA and four common proteins, such as Cbf5 (called as Dyskerin in humans), Nhp2, Nop10, and Gar1. Cbf5 acts as a catalytic subunit and converts uridine recognized by the box H/ACA snoRNP to Ψ.



B

Family	Synthase	Location	Known target
TruA	PUS1	Nucleus	mRNA, cyto-tRNA
	PUS2	Nucleus, Cytoplasm	cyto-tRNA
	PUS3	Mitochondria	mt-tRNA
TruB	PUS4	Nucleus, Cytoplasm	mRNA, cyto-tRNA, mt-tRNA
RluA	PUS5	Nucleus, Mitochondria	mRNA
	PUS6	Mitochondria	mt-rRNA)
	PUS8	Nucleus, Cytoplasm	mRNA, cyto-tRNA
	PUS9	Cytoplasm	cyto-tRNA
TruD	PUS7	Cytoplasm, Mitochondria	mRNA, mt-tRNA
Pus10	PUS10	Nucleus, Cytoplasm, Mitochondria	mRNA, mt-tRNA

Fig. 4 Regulation of RNA degradation by Ψ. (**a**) Structures of uridine and its isomer, Ψ. (**b**) List of PUSs (key factors of RNA-independent pathway). (**c**) The H/ACA RNP box structure consists of a H/ACA RNA box and four proteins—Cbf5, Nhp2, Nop10, and Gar1

3.3.3 Regulation of RNA Degradation with Ψ

Compared with uridine, Ψ forms a more robust base pair with adenine (Penzo et al. 2017; Adachi et al. 2019). Therefore, Ψ in mRNAs may affect local secondary structure and stability. Karijolich et al. showed both in vitro and in vivo that artificial U-to-Ψ conversion in translation termination codons (UAA, UGA, and UAG) converts them to missense codons, and U-to-Ψ conversion at the premature termination codon suppressed RNA degradation by NMD (Karijolich and Yu 2011), Karikó et al. showed that in mammalian cells, transcribed Ψ-containing mRNAs are more stable than mRNAs without Ψ of the same sequence (Karikó et al. 2008). The RNA-stabilizing mechanism of Ψ is not sufficiently understood, however, and warrants further study.

3.4 N7-methylguanosine (m^7G)

3.4.1 Distribution and Function of m^7G

In contrast to the internal distributions of m^6A and m^5C within RNA, m^7G is constitutively located at the 5′ end of nearly all eukaryotic mRNAs (Furuichi 2015; Ramanathan et al. 2016). The m^7G binds to the 5′ end of nascent RNAs via a triphosphate bond, mainly co-transcriptionally, and is called m^7G-cap or 5′-cap. The m^7G-cap protects against exonuclease cleavage, is an initiating signal to cap-dependent protein synthesis, identifies the RNA for nuclear pre-mRNA splicing, polyadenylation, and nuclear extracellular transport (Ramanathan et al. 2016), and acts as a center for non-self-identification for innate immunity responses to foreign RNAs (Daffis et al. 2010; Ramanathan et al. 2016). The m^7G is also found internally in variable region and anticodon region of tRNA (Guy and Phizicky 2014), which play a major role in stabilizing the tRNA (Tomikawa 2018) and ensuring efficient and accurate translation (Huang and Hopper 2016; Duechler et al. 2016), respectively.

3.4.2 Regulation of m^7G Modification

Capping is the first RNA modification transcribed by RNA polymerase II, and it occurs co-transcriptionally in the nucleus when the 25–30 nts of nascent RNA are transcribed (Shatkin and Manley 2000; Moteki and Price 2002). The m^7G-cap in eukaryotes is attached by three steps (Fig. 5a): First, RNA triphosphatase (TPase) removes a phosphate group at the 5′ end of the nascent RNA; RNA guanylyltransferase (GTase) then removes two phosphate groups from guanosine triphosphate (GTP) and adds the remaining guanosine monophosphate to the 5′ end of the RNA; and finally, the guanine-N7 methyltransferase (guanine-N7 MTase) methylates the seventh position of the added guanine (Ramanathan et al. 2016). The

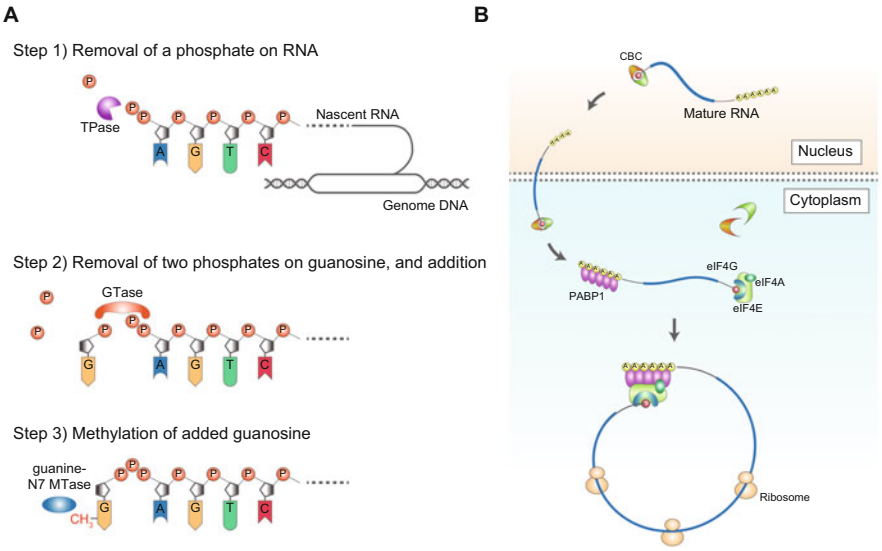


Fig. 5 Regulation of RNA degradation by m^7G (**a**) Three steps of m^7G capping: A phosphate on the 5′ RNA end is removed by TPase; guanosine is dephosphorylated, removed, and attached to the RNA by GTase; and the attached guanosine is methylated. (**b**) Stabilization of m^7G-capped RNA. The capped RNA is pseudo-circulated by binding the eIF4F complex to the PAPB1

added m^7G-cap functions as a signal to promote splicing in the nucleus and extracellular transport, inhibits exonuclease degradation in the cytoplasm, and promotes translation (Evdokimova et al. 2001). Re-capping of uncapped mRNAs is reported to occur in P-bodies (cytoplasmic RNA [re]-capping) (Mukherjee et al. 2014; Ramanathan et al. 2016). This is thought to regulate RNA translation efficiency, but its biological significance requires further study.

3.4.3 Regulation of RNA Degradation with m^7G

In the nucleus, m^7G is co-transcriptionally recognized by the cap-binding complex (CBC), a heterodimer composed of CBP80 and CBP20, which functions as a signal for splicing and nuclear transport (Parker and Sheth 2007; Isken and Maquat 2007; Maquat et al. 2010). In the cytoplasm, this CBC is replaced by eukaryotic initiation factor 4E (eIF4E) in the eIF4F complex. eIF4E and eIF4G, which are also part of the eIF4F complex, interact with the mRNA polyA-binding protein PABP1 and pseudo-circulate the mRNA (Fig. 5b). This pseudo-circularization increases translation efficiency and stabilizes the mRNA by protecting it from exonucleases. The addition of m^7G also causes 2′-OH methylation of the first and second RNA bases, resulting in 2′-*O*-methyladenosine (A_m) when the first base is adenine, which can be further methylated at position 6 to form m^6A_m. m^6A_m is a more preferred substrate for FTO and further stabilizes RNAs that contain this modification (Mauer et al. 2017).

The m^7G is also found internally in mammalian mRNA (Zhang et al. 2019), tRNA (Guy and Phizicky 2014), and rRNA (Sloan et al. 2017). Most of the m^7G modification in tRNA occurs at position 46 in the variable region, which forms tertiary base pair with C13-G22 to stabilize the tRNA (Tomikawa 2018). The m^7G modification in tRNA is also found in anticodon loop, which can stabilize specific base pairs to ensure efficient and accurate translation (Huang and Hopper 2016; Duechler et al. 2016). The functions of internal m^7G in mammalian rRNA and mRNA have been not well clarified and further study is required.

3.5 N4-acetylcytidine (ac^4C)

3.5.1 Distribution and Function of ac^4C

The ac^4C modification is the only currently known acetylation event in eukaryotic RNA. This modification was initially found in the bacterial $tRNA^{met}$ anticodon, but was later detected in eukaryotic $tRNA^{Ser}$, $tRNA^{Lue}$, and 18S rRNA (Boccaletto et al. 2018). Dong et al. also used LC-MS/MS to show that ac^4C is installed in human mRNA (Dong et al. 2016).

3.5.2 Regulation of ac^4C Modification

N-acetyltransferase 10 (NAT10) is a writer protein for ac^4C modification. Analysis of RNA modifications in *NAT10* knockout mice confirmed that *NAT10* deficiency significantly reduces ac^4C modification in RNA (Arango et al. 2018). For all RNA species, ac^4C modification is apparently catalyzed by the NAT10 enzyme or its homologs (Chimnaronk et al. 2009; Suzuki et al. 2014; Sharma et al. 2015). Thus far, neither the ac^4C reader nor the eraser protein has been identified. Whether ac^4C modification is a reversible process is also unclear. Therefore, future studies should elucidate the molecular mechanisms of ac^4C modification in RNAs.

3.5.3 Regulation of RNA Degradation with ac^4C

Analysis of the half-lives of ac^4C-modified mRNA using *NAT10* knockout mice showed that acetylation levels were positively correlated with target mRNA stability and that translation was enhanced when ac^4C exists in wobble cytidine (Arango et al. 2018). This suggests that ac^4C contributes to RNA stabilization and promotion of translational efficiency. However, the specific mechanism at work here is still unclear.

3.6 Others

RNA degradation is a fundamental determinant of cellular states; however, several other RNA modifications besides those described above are predicted to affect stability. For example, 8-oxoG shares a reader protein YBX1 with m^5C, which has been suggested to stabilize RNA by recruiting HuR, an RBP that stabilizes RNA (Chen et al. 2019). N2,N2-dimethylguanosine ($m^2{}_2G$) is also reported to pair with adenine instead of cytosine (Pallan et al. 2008), which suggests that it affects the structure and stability of rRNAs and tRNAs.

One of the specific RNA modifications is adenosine to inosine (A-to-I) editing. Adenosines in double-stranded RNAs are converted to inosine through hydrolytic deamination by adenosine deaminase acting on RNA (ADAR). Since structure of inosine is similar to that of guanosine, the converted inosine is recognized as guanosine. Although A-to-I editing affects protein structure during translation, it can affect secondary structure of RNA and 3′UTR in mRNA, where RNA degradation machineries recognize to regulate RNA stability.

Comprehensive measurement techniques have revealed details about RNA modification. For example, Oxford Nanopore direct-RNA sequencing technology, a long-read sequencing technique, is used to directly identify multiple RNA modifications installed on polyA RNA and targeted lncRNAs (Leger et al. 2019). These technological developments may reveal novel RNA modifications and their functions. Many of these novel RNA modifications might also affect RNA stability, in addition to the RNA modifications described above.

4 Mechanism of Innate Immunity Evasion of Viruses by RNA Modification

As described above, many chemical modifications regulate RNA stability. Cells use RNA recognition and degradation machinery, both to regulate ordinary cellular processes and as an immunity process that restricts viral RNAs. However, some viruses inhibit RNA degradation mechanisms through their own RNA modifications.

4.1 Mechanisms of Foreign RNA Suppression in Host Cells

To eliminate virus, host cells evolved mechanisms to degrade viral RNA and DNA. This is a part of innate immunity, and its major cellular factors are cytoplasmic PRRs, such as RIG-I and MDA5. These receptors exist in the cytoplasm as pathogen recognition receptors and monitor viral infection by recognizing phosphate groups and motif sequences at the 5′ end of RNA. Upon recognizing viral infections, RIG-I is ubiquitinated by TRIM25, which activates downstream signaling (Gack et al.

2007), resulting in a cellular type-I interferon (IFN) response via the NF-κB pathway (Pichlmair et al. 2006; Kell and Gale 2015). The OAS/RNaseL pathway is another innate immune system. OAS is a metabolic enzyme that recognizes exogenous RNA and synthesizes 2′,5′-linked oligoadenylate (2–5A) from ATP as a substrate; 2-5A activates the RNA-degrading enzyme RNase L, which activates the exogenous RNA degradation mechanism (Nogimori et al. 2019).

4.2 Suppression of Innate Immunity by RNA Modification

The m^6A modification has been suggested to act as an immune evasion mechanism for certain viruses (Tan and Gao 2018). Human metapneumovirus (hMPV), which causes respiratory syndrome in children (Falsey et al. 2003; Greensill et al. 2003; Williams et al. 2004), is a single-stranded RNA virus. In hMPV as a model, the m^6A modification installed in viral genomic RNA allows the RNA to escape recognition by RIG-I, and attenuates the RIG-I-dependent production of type-I IFN (Lu et al. 2020). Knocking out *METTL3*, part of the m^6A writer, reduced mRNA levels of influenza A virus (IAV) in A549 lung cancer cells (Courtney et al. 2017). This m^6A modification is also found in other viruses that replicate in the nucleus, such as human immunodeficiency virus (HIV) and adenovirus (Tirumuru et al. 2016; Riquelme-Barrios et al. 2018), which implies that this modification helps avoid innate immunity. Hepatitis C virus (HCV) RNA in infected liver cancer cells is reported to use m^6A methylation and demethylation mechanisms to relocalize to the cytoplasm (Gokhale et al. 2016), which suggests that some cytoplasmic replicating viruses take over m^6A modification strategies for survival.

Unexpectedly, m^7G-cap itself has no apparent effect on the RIG-I recognition system (Devarkar et al. 2016), whereas methylation of the first position held by the m^7G-cap (called cap1) attenuates its interactions with RIG-I and MDA5 and suppresses the IFN signaling pathway (Züst et al. 2011; Schuberth-Wagner et al. 2015). To take over this m^7G capping innate immunity system, some viruses have their own capping mechanisms that are completely different from those of host cells. For example, in host cells, GTP is methylated after binding to monophosphorylated RNA, whereas in Alphaviruses, which have single-stranded RNA as their genomes, GTP is methylated by nsP1, a GTase, and then bound to monophosphorylated RNA by nsP2, an RNA TPase (Fata et al. 2002). Instead of developing their own capping mechanism, some RNA viruses "steal" cap structures from host RNAs, which is called cap-snatching. For example, the RNA-dependent RNA polymerase complex of IAV uses a typical cap-snatching mechanism (Sugiyama et al. 2009; Huet et al. 2010; Reich et al. 2014). It binds to the host's capped RNA using polymerase base protein 2 (PB2), and its polymerase acidic protein (PA) cleaves the first 10–15 nt of the capped RNA. The cleaved capped RNA is used to prime viral RNA (Ruigrok et al. 2010; Gu et al. 2015). Thus, most viruses use m^7G-cap to stabilize their own genomes. The m^7G is also identified in HCV and polioviruses and lacks their own

m^7G-cap structures (McIntyre et al. 2018). These results suggest that guanosine in viral genomes may be modified, but the mechanism and biological role are unclear.

Methods of identifying RNA modifications at the transcriptome level enable us to detect modifications on many viral genomic or viral-derived RNAs. The RNA modification landscapes of whole transcriptomes obtained from virus-infected cells were investigated based on mass spectrometry analysis (McIntyre et al. 2018). In the study, 2′-*O*-methylated nucleotides, N1-methyladenosine, m5C, and N4-acetylcytosine were found on genomes isolated from Zika virus, dengue virus, hepatitis C virus, poliovirus, and HIV-1, which suggests that viruses use these RNA modifications to regulate several activities, including stability.

5 Applications for Modified Nucleotides

5.1 Expansion of RNA Modification to mRNA Medicine

Insight from viral stability regulation based on RNA modification has had an unexpected effect on the RNA industry. Therapies that use mRNA medicine can modify cellular conditions or add specific functions by incorporating in vitro transcribed mRNA that encodes a target protein into cells to express the protein. As this technique does not require the genetic entity to be transported to the nucleus, it decreases the risk of oncogenesis that could be caused by inserting mutations into the genome.

However, in vitro transcribed mRNAs are subject to degradation by exoribonucleases (Nagarajan et al. 2013) and may elicit innate immunity responses similar to those of viruses (Sahin et al. 2014). Several approaches to this problem have been tried, which use studies of innate immunity to virus. Innate immunity to foreign mRNAs can be suppressed by modified nucleic acids: m^5C and Ψ incorporated in in vitro have led to transcription of inactive *RIG-I* mRNA, which stifles degradation by innate immunity (Durbin et al. 2016); Ψ incorporated into mRNA represses protein kinase R, which represses translation of exogenous mRNA (Anderson et al. 2010).

These modified RNAs are also useful in the quality control of biologically derived medicines; Parr et al. created N1-methylpseudouridine ($m^1Ψ$)-modified mRNA switches that allow more efficient removal of undifferentiated induced pluripotent stem cells (iPS cells) (Parr et al. 2020). These industrial applications could lead to safer cellular medicine and highly specific mRNA therapies.

5.2 *Measurements of RNA Kinetics Using Modified Nucleotides*

Modified nucleotides are useful for comprehensive kinetic measurements, including RNA stability. Artificially modified nucleotides are incorporated in nascent RNAs by adding them to cell media (Fig. 6a). Changes in the labeled RNAs are obtained

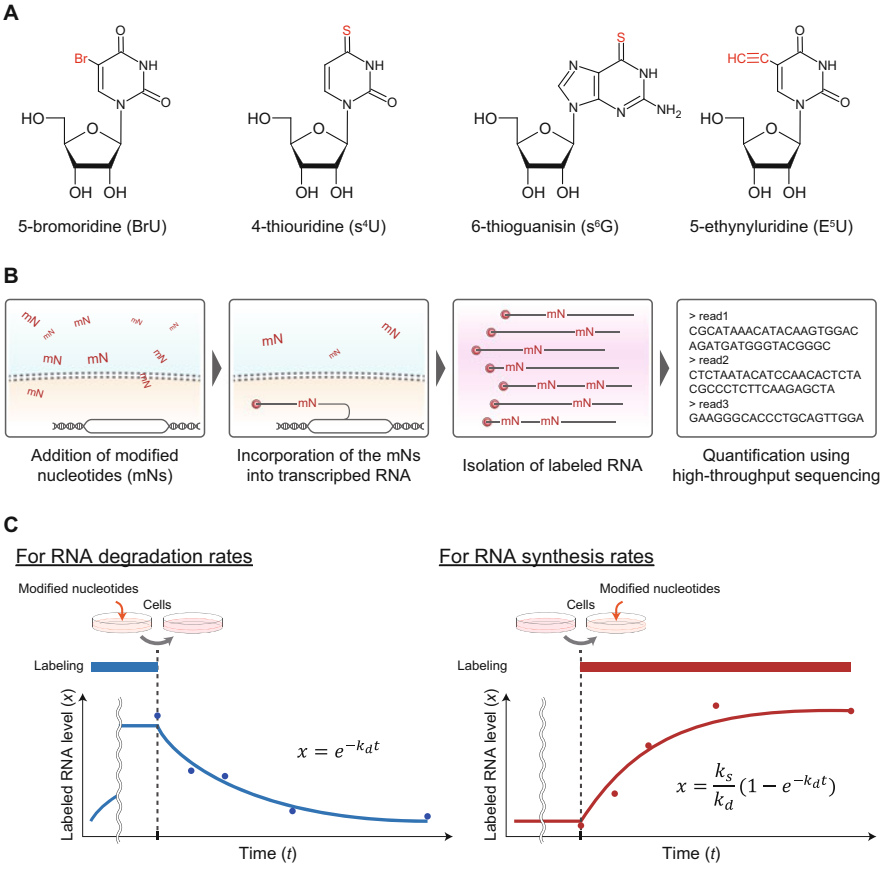


Fig. 6 Measuring of intracellular RNA stability using modified nucleosides. (**a**) Molecular structures of representative modified nucleosides used to measure RNA stability: 5-bromoridine (BrU), 4-thiouridine (s^4U), 6-thioguanisine (s^6G), and 5-ethynyluridine (E^5U). (**b**) Scheme of RNA stability measurement. RNAs that incorporate modified nucleotides are isolated and quantified using high-throughput sequencing. (**c**) Kinetics of RNAs that are labeled with modified nucleotides (k_s: RNA synthesis rate; k_d: RNA synthesis rate). Left: to measure the RNA degradation rate, modified nucleotides are added transiently, and the labeled RNA is quantified over a period of time. Labeled RNA decreases time dependently. Right: to measure the RNA synthesis rate, modified nucleotides are added, and the labeled RNAs are quantified over a period of time. The labeled RNA increases time dependently

over time and are quantified using high-throughput sequencing to estimate RNA synthesis and/or degradation comprehensively (Fig. 6b, c) (Yamada and Akimitsu 2019; Wolfe et al. 2019). Uridine analogs, such as 5-bromouridine (BrU) and 4-thiouridine (s^4U), are frequently used for this purpose. BrU immunoprecipitation chase-deep sequencing analysis (BRIC-seq) was developed to estimate RNA degradation by fitting changes in amounts of individual BrU-labeled RNA transcripts over time, measured by massive parallel sequencing, to a mathematical model (Tani et al. 2012; Imamachi et al. 2014; Yamada et al. 2018). To address this point, techniques based on chemical reactions were developed. Some groups developed methods to conjugate thiol-specific biotin reversibly to the s^4U incorporated in the RNA (Cleary et al. 2005; Kenzelmann et al. 2007; Dölken et al. 2008; Friedel et al. 2009; Schwalb et al. 2016; Michel et al. 2017). These RNAs including biotinylated s^4Us can be enriched by binding to streptavidin beads efficiently. Similar methods to induce mismatches in s^4U-incorporating RNA have been developed, including thiol (SH)-linked alkylation for metabolic RNA sequencing (SLAM-seq), thiouridine-to-cytidine sequencing (TUC-seq), and TimeLapse-seq. These methods allow informatic determination of s^4U-labeled RNAs by reverse transcription-dependent s^4U-to-C conversion. SLAM-seq alkylates s^4U using the primary thiol-reactive compound iodoacetamide (Herzog et al. 2017; Muhar et al. 2018). As the alkylated s^4U pairs with guanine instead of adenine during reverse transcription, the alkylated s^4U is detected as cytosine. Identification of the T-to-C mutation in the RNA sequence allows us to distinguish between s^4U-labeled and -unlabeled RNAs. Similarly, TUC-seq (Riml et al. 2017; Lusser et al. 2020) and TimeLapse-seq (Schofield et al. 2018) use osmium tetroxide (OsO_4)-mediated oxidation and oxidative-nucleophilic-aromatic substitution, respectively, to induce T-to-C mutations. Other methods measure multiple phases of RNA metabolism (such as RNA synthesis and degradation) simultaneously, by combining multiple modified nucleotide labeling. For instance, we have developed Dyrec-seq, which combines BrU and s^4U labeling to measure RNA synthesis and degradation simultaneously (Kawata et al. 2020). Gasser et al. developed a TUC-seq DUAL that simultaneously causes T-to-C and G-to-A mutations in a single reaction, using OsO_4-mediated oxidation on RNAs labeled simultaneously with s^4U and 6-thioguanosine (Gasser et al. 2020).

Finally, the variety of modified nucleotides and distinguished methods for analyzing these RNA dynamics has expanded over the years. For example, 5-ethynylcytidine has been used marginally to measure RNA metabolism (Qu et al. 2013), but this modified nucleotide was diverted for Oxford Nanopore direct-RNA sequencing technology to widen its scope of usage (Maier et al. 2020). However, individual modified nucleotides have unique biological and chemical properties, such as cellular uptake efficiency and cytotoxicity. These properties warrant further investigation.

6 Conclusion

Interest in RNAs is generally based on their roles in carrying and decoding information encoded in genomic DNA. However, RNAs not only carry information but also reflect the cellular state. Moreover, they participate in cellular functions by forming complex secondary and tertiary structures. RNA modifications what we now call as epitranscriptome were initially established in the 2010s. Advances in measurement technology have found more than 170 RNA modifications, only a few of which have been shown to affect RNA degradation. Although RNA degradation has been regarded as the end of the RNA life cycle—i.e., the death of information—RNA degradation affects both temporal and spatial cell fates through coordinated regulation by various players (Alonso 2012). In other words, RNA degradation participates in cellular function as a biological code. Elucidating the mechanisms of the epitranscriptome's functions will open an approach to various cellular physiological mechanisms.

Acknowledgments We thank our laboratory members for critically reading this manuscript. This manuscript was edited by Edanz (https://en-author-services.edanzgroup.com/ac). This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Number: 17KK0163, 18H02570, 18KT0016, 16H06279, and 20H04838) and the Uehara Memorial Foundation. K.K. received funding from JSPS KAKENHI (Grant Number: 19K16635), Kowa Life Science Foundation, and Takeda Science Foundation.

References

Adachi H, De Zoysa MD, Yu YT (2019) Post-transcriptional pseudouridylation in mRNA as well as in some major types of noncoding RNAs. Biochim Biophys Acta 1862:230–239

Agarwala SD, Blitzblau HG, Hochwagen A et al (2012) RNA methylation by the MIS complex regulates a cell fate decision in yeast. PLoS Genet 8(6):e1002732

Agris PF (2008) Bringing order to translation: the contributions of transfer RNA anticodon-domain modifications. EMBO Rep 9:629–635

Alonso CR (2012) A complex "mRNA degradation code" controls gene expression during animal development. Trends Genet 28:78–88

Anderson JT, Wang X (2009) Nuclear RNA surveillance: no sign of substrates tailing off nuclear RNA surveillance J. T. Anderson and X. Wang. Crit Rev Biochem Mol Biol 44:16–24

Anderson BR, Muramatsu H, Nallagatla SR et al (2010) Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. Nucleic Acids Res 38:5884–5892

Arango D, Sturgill D, Alhusaini N et al (2018) Acetylation of cytidine in mRNA promotes translation efficiency. Cell 175 e24:1872–1886

Basturea GN, Zundel MA, Deutscher MP (2011) Degradation of ribosomal RNA during starvation: comparison to quality control during steady-state growth and a role for RNase PH. RNA 17:338–345

Batista PJ, Molinie B, Wang J et al (2014) M^6A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell 15:707–719

Blanco S, Dietmann S, Flores JV et al (2014) Aberrant methylation of tRNA s links cellular stress to neuro-developmental disorders. EMBO J 33:2020–2039

Boccaletto P, MacHnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307

Bohnsack KE, Höbartner C, Bohnsack MT (2019) Eukaryotic 5-methylcytosine (M^5C) RNA methyltransferases: mechanisms, cellular functions, and links to disease. Genes (Basel) 10:102

Boo SH, Kim YK (2020) The emerging role of RNA modifications in the regulation of mRNA stability. Exp Mol Med 52:400–408

Bourgeois G, Ney M, Gaspar I et al (2015) Eukaryotic rRNA modification by yeast 5-methylcytosine-methyltransferases and human proliferation-associated antigen p120. PLoS One 10:e0133321

Carlile TM, Rojas-Duran MF, Zinshteyn B et al (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515:143–146

Charette M, Gray MW (2000) Pseudouridine in RNA: what, where, how and why. IUBMB Life 49:341–351

Chen T, Hao YJ, Zhang Y et al (2015) M^6A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. Cell Stem Cell 16:289–301

Chen X, Li A, Sun BF et al (2019) 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. Nat Cell Biol 21:978–990

Cheng X, Roberts RJ (2001) AdoMet-dependent methylation DNA methyltransferases and base flipping. Nucleic Acids Res 29:3784–3795

Chimnaronk S, Suzuki TT, Manita T et al (2009) RNA helicase module in an acetyltransferase that modifies a specific tRNA anticodon. EMBO J 28:1362–1373

Chow CS, Lamichhane TN, Mahto SK (2007) Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. ACS Chem Biol 2:610–619

Cleary MD, Meiering CD, Jan E et al (2005) Biosynthetic labeling of RNA with uracil phosphoribosyltransferase allows cell-specific microarray analysis of mRNA synthesis and decay. Nat Biotechnol 23:232–237

Collart MA, Timmers HTM (2004) The eukaryotic Ccr4-not complex: a regulatory platform integrating mRNA metabolism with cellular signaling pathways? Prog Nucleic Acid Res Mol Biol 77:289–322

Courtney DG, Kennedy EM, Dumm RE et al (2017) Epitranscriptomic enhancement of influenza a virus gene expression and replication. Cell Host Microbe 22 e5:377–386

Daffis S, Szretter KJ, Schriewer J et al (2010) 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature 468:452–456

De La Sierra-Gallay IL, Zig L, Jamalli A et al (2008) Structural insights into the dual activity of RNase. J Nat Struct Mol Biol 15:206–212

Denis CL, Chen J (2003) The CCR4-NOT complex plays diverse roles in mRNA metabolism. Prog Nucleic Acid Res Mol Biol 73:221–250

Devarkar SC, Wang C, Miller MT et al (2016) Structural basis for m^7G recognition and 2′-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. Proc Natl Acad Sci USA 113:596–601

Dölken L, Ruzsics Z, Rädle B et al (2008) High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. RNA 14:1959–1972

Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. Nature 485:201–206

Dong A, Yoder JA, Zhang X et al (2001) Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. Nucleic Acids Res 29:439–448

Dong C, Niu L, Song W et al (2016) tRNA modification profiles of the fast-proliferating cancer cells. Biochem Biophys Res Commun 476:340–345

Du H, Zhao Y, He J et al (2016) YTHDF2 destabilizes m^6A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat Commun 7:1–11

Duechler M, Leszczyńska G, Sochacka E et al (2016) Nucleoside modifications in the regulation of gene expression: focus on tRNA. Cell Mol Life Sci 73:3075–3095

Durbin AF, Wang C, Marcotrigiano J et al (2016) RNAs containing modified nucleotides fail to trigger RIG-I conformational changes for innate immune signaling. MBio 7:e00833–e00816
Edelheit S, Schwartz S, Mumbach MR et al (2013) Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m^5C within archaeal mRNAs. PLoS Genet 9:e1003602
Elliott BA, Ho HT, Ranganathan SV et al (2019) Modification of messenger RNA by 2′-O-methylation regulates gene expression in vivo. Nat Commun 10:3401
Erson-Bensan AE, Begik O (2017) m^6A modification and implications for microRNAs. MicroRNA 6:97–101
Evdokimova V, Ruzanov P, Imataka H et al (2001) The major mRNA-associated protein YB-1 is a potent 5′ cap-dependent mRNA stabilizer. EMBO J 20:5491–5502
Falsey AR, Erdman D, Anderson LJ et al (2003) Human metapneumovirus infections in young and elderly adults. J Infect Dis 187:785–790
Fata CL, Sawicki SG, Sawicki DL (2002) Modification of Asn374 of nsP1 suppresses a Sindbis virus nsP4 minus-Strand polymerase mutant. J Virol 76:8641–8649
Friedel CC, Dölken L, Ruzsics Z et al (2009) Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. Nucleic Acids Res 37:e115
Fu Y, Zhuang X (2019) m^6A-binding YTHDF proteins promote stress granule formation by modulating phase separation of stress granule proteins. bioRxiv 694455
Fu Y, Jia G, Pang X et al (2013) FTO-mediated formation of N6-hydroxymethyladenosine and N 6-formyladenosine in mammalian RNA. Nat Commun 4:1798
Furuichi Y (2015) Discovery of m^7G-cap in eukaryotic mRNAs. Proc Japan Acad Ser B Phys Biol Sci 91:394–409
Gack MU, Shin YC, Joo CH et al (2007) TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. Nature 446:916–920
Gasser C, Delazer I, Neuner E et al (2020) Thioguanosine conversion enables mRNA-lifetime evaluation by RNA sequencing using double metabolic Labeling (TUC-seq DUAL). Angew Chem – Int Ed Engl 59:6881–6886
Geula S, Moshitch-Moshkovitz S, Dominissini D et al (2015) m^6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science 347:1002–1006
Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. Nat Rev Genet 10:94–108
Glasmacher E, Hoefig KP, Vogel KU et al (2010) Roquin binds inducible costimulator mRNA and effectors of mRNA decay to induce microRNA-independent post-transcriptional repression. Nat Immunol 11:725–733
Gokhale NS, McIntyre ABR, McFadden MJ et al (2016) N6-methyladenosine in flaviviridae viral RNA genomes regulates infection. Cell Host Microbe 20:654–665
Goll MG, Kirpekar F, Maggert KA et al (2006) Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science 311:395–398
Greensill J, McNamara PS, Dove W et al (2003) Human metapneumovirus in severe respiratory syncytial virus bronchiolitis. Emerg Infect Dis 9:372–375
Gu W, Gallagher GR, Dai W et al (2015) Influenza a virus preferentially snatches noncoding RNA caps. RNA 21:2067–2075
Guy MP, Phizicky EM (2014) Two-subunit enzymes involved in eukaryotic post-transcriptional tRNA modification. RNA Biol 11:1608–1618
Hamma T, Ferré-D'Amaré AR (2006) Pseudouridine synthases. Chem Biol 13:1125–1135
Hao S, Baltimore D (2009) The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. Nat Immunol 10:281–288
Hayrapetyan A, Grosjean H, Helm M (2009) Effect of a quaternary pentamine on RNA stabilization and enzymatic methylation. Biol Chem 390:851–861
Helm M (2006) Post-transcriptional nucleotide modification and alternative folding of RNA. Nucleic Acids Res 34:721–733

Helm M, Motorin Y (2017) Detecting RNA modifications in the epitranscriptome: predict and validate. Nat Rev Genet 18:275–291
Herzog VA, Reichholf B, Neumann T et al (2017) Thiol-linked alkylation of RNA to assess expression dynamics. Nat Methods 14:1198–1204
Horiuchi K, Kawamura T, Iwanari H et al (2013) Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle. J Biol Chem 288:33292–33302
Houseley J, Tollervey D (2009) The many pathways of RNA degradation. Cell 136:763–776
Huang HY, Hopper AK (2016) Multiple layers of stress-induced regulation in tRNA biology. Life 6
Huet S, Avilov SV, Ferbitz L et al (2010) Nuclear import and assembly of influenza a virus RNA polymerase studied in live cells by fluorescence cross-correlation spectroscopy. J Virol 84:1254–1264
Hutvagner G, Simard MJ (2008) Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol 9:22–32
Imamachi N, Tani H, Mizutani R et al (2014) BRIC-seq: a genome-wide approach for determining RNA stability in mammalian cells. Methods 67:55–63
Imamura K, Takaya A, Ishida Y et al (2018) Diminished nuclear RNA decay upon Salmonella infection upregulates antibacterial noncoding RNAs. EMBO J 37:e97723
Isken O, Maquat LE (2007) Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev 21:1833–1856
Jarrous N (2017) Roles of RNase P and its subunits. Trends Genet 33:594–603
Karijolich J, Yu YT (2011) Converting nonsense codons into sense codons by targeted pseudouridylation. Nature 474:395–399
Karikó K, Muramatsu H, Welsh FA et al (2008) Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. Mol Ther 16:1833–1840
Kawamata T, Tomari Y (2010) Making RISC. Trends Biochem Sci 35:368–376
Kawata K, Wakida H, Yamada T et al (2020) Metabolic labeling of RNA using multiple ribonucleoside analogs enables the simultaneous evaluation of RNA synthesis and degradation rates. Genome Res gr.264408.120
Keene JD (2010) Minireview: global regulation and dynamics of ribonucleic acid. Endocrinology 151:1391–1397
Kell AM, Gale M (2015) RIG-I in RNA virus recognition. Virology 479–480:110–121
Kenzelmann M, Maertens S, Hergenhahn M et al (2007) Microarray analysis of newly synthesized RNA in cells and animals. Proc Natl Acad Sci USA 104:6164–6169
King MY, Redman KL (2002) RNA methyltransferases utilize two cysteine residues in the formation of 5-methylcytosine. Biochemistry 41:11218–11225
Krzyszton M, Zakrzewska-Placzek M, Koper M et al (2012) Rat1 and Xrn2: the diverse functions of the nuclear Rat1/Xrn2 exonuclease. In: Enzymes. Academic, New York, pp 131–163
Lau NC, Kolkman A, van Schaik FMA et al (2009) Human Ccr4-not complexes contain variable deadenylase subunits. Biochem J 422:443–453
Lee Y, Choe J, Park OH et al (2020) Molecular mechanisms driving mRNA degradation by m6A modification. Trends Genet 36:177–188
Leger A, Amaral P, Pandolfini L et al (2019) RNA modifications detection by comparative Nanopore direct RNA sequencing bioRxiv:843136
Lence T, Akhtar J, Bayer M et al (2016) M6A modulates neuronal functions and sex determination in Drosophila. Nature 540:242–247
Leppek K, Schott J, Reitter S et al (2013) Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. Cell 153:869–881
Li X, Zhu P, Ma S et al (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol 11:592–597

Liang XH, Xu YX, Michael S (2002) The spliced leader-associated RNA is a trypanosome-specific sn(o) RNA that has the potential to guide pseudouridine formation on the SL RNA. RNA 8:237–246
Liao S, Sun H, Xu C (2018) YTH domain: a family of N6-methyladenosine (m^6A) readers. Genomics, Proteomics Bioinforma 16:99–107
Linder B, Grozhik AV, Olarerin-George AO et al (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods 12:767–772
Lister R, Pelizzola M, Dowen RH et al (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462:315–322
Liu Y, Santi DV (2000) m^5C RNA and m^5C DNA methyl transferases use different cysteine residues as catalysts. Proc Natl Acad Sci USA 97:8263–8265
Liu J, Yue Y, Han D et al (2014) A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol 10:93–95
Liu SY, Feng Y, Wu JJ et al (2020) m^6A facilitates YTHDF-independent phase separation. J Cell Mol Med 24:2070–2072
Lu W, Tirumuru N, Gelais CS et al (2018) N6-Methyladenosine– binding proteins suppress HIV-1 infectivity and viral production. J Biol Chem 293:12992–13005
Lu M, Zhang Z, Xue M et al (2020) N6-methyladenosine modification enables viral RNA to escape recognition by RNA sensor RIG-I. Nat Microbiol 5:584–598
Lusser A, Gasser C, Trixl L et al (2020) Thiouridine-to-cytidine conversion sequencing (TUC-Seq) to measure mRNA transcription and degradation rates. Methods Mol Biol 2062:191–211
Lykke-Andersen S, Jensen TH (2015) Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat Rev Mol Cell Biol 16:665–677
Maier KC, Gressel S, Cramer P et al (2020) Native molecule sequencing by nano-ID reveals synthesis and stability of RNA isoforms. Genome Res 30:1332–1344
Maquat LE, Tarn WY, Isken O (2010) The pioneer round of translation: features and functions. Cell 142:368–374
Mathy N, Bénard L, Pellegrini O et al (2007) 5′-to-3′ exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5′ stability of mRNA. Cell 129:681–692
Mauer J, Luo X, Blanjoie A et al (2017) Reversible methylation of m^6Am in the 5′ cap controls mRNA stability. Nature 541:371–375
McIntyre W, Netzband R, Bonenfant G et al (2018) Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection. Nucleic Acids Res 46:5776–5791
Megel C, Morelle G, Lalande S et al (2015) Surveillance and cleavage of eukaryotic tRNAs. Int J Mol Sci 16:1873–1893
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149:1635–1646
Michel M, Demel C, Zacher B et al (2017) TT-seq captures enhancer landscapes immediately after T-cell stimulation. Mol Syst Biol 13:920
Mino T, Murakawa Y, Fukao A et al (2015) Regnase-1 and roquin regulate a common element in inflammatory mRNAs by spatiotemporally distinct mechanisms. Cell 161:1058–1073
Moteki S, Price D (2002) Functional coupling of capping and transcription of mRNA. Mol Cell 10:599–609
Motorin Y, Helm M (2010) TRNA stabilization by modified nucleotides. Biochemistry 49:4934–4944
Muhar M, Ebert A, Neumann T et al (2018) SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. Science 360:800–805
Mukherjee C, Bakthavachalu B, Schoenberg DR (2014) The cytoplasmic capping complex assembles on adapter protein Nck1 bound to the proline-rich C-terminus of mammalian capping enzyme. PLoS Biol 12:e1001933
Nachtergaele S, He C (2018) Chemical modifications in the life of an mRNA transcript. Annu Rev Genet 52:349–372

Nagarajan VK, Jones CI, Newbury SF et al (2013) XRN 5′→3′ exoribonucleases: structure, mechanisms and functions. Biochim Biophys Acta 1829:590–603
Niu Y, Zhao X, Wu YS et al (2013) N6-methyl-adenosine (m^6A) in RNA: an old modification with a novel epigenetic function. Genomics, Proteomics Bioinforma 11:8–17
Nogimori T, Nishiura K, Kawashima S et al (2019) Dom34 mediates targeting of exogenous RNA in the antiviral OAS/RNase L pathway. Nucleic Acids Res 47:432–449
Pallan PS, Kreutz C, Bosio S et al (2008) Effects of N2,N2-dimethylguanosine on RNA structure and stability: crystal structure of an RNA duplex with tandem $m^2{}_2G$:A pairs. RNA 14:2125–2135
Park OH, Ha H, Lee Y et al (2019) Endoribonucleolytic cleavage of m^6A-containing RNAs by RNase P/MRP complex. Mol Cell 74:494–507.e8
Parker R, Sheth U (2007) P bodies and the control of mRNA translation and degradation. Mol Cell 25:635–646
Parr CJC, Wada S, Kotake K et al (2020) N1-Methylpseudouridine substitution enhances the performance of synthetic mRNA switches in cells. Nucleic Acids Res 48:e35
Patil DP, Chen CK, Pickering BF et al (2016) M^6A RNA methylation promotes XIST-mediated transcriptional repression. Nature 537:369–373
Penzo M, Guerrieri AN, Zacchini F et al (2017) RNA pseudouridylation in physiology and medicine: for better and for worse. Genes (Basel) 8:301
Phizicky EM, Hopper AK (2010) tRNA biology charges to the front. Genes Dev 24:1832–1860
Pichlmair A, Schulz O, Tan CP et al (2006) RIG-I-mediated antiviral responses to single-stranded RNA bearing 5′-phosphates. Science 314:997–1001
Piekna-Przybylska DD, Decatur WA, Fournier MJ (2008) The 3D rRNA modification maps database: with interactive tools for ribosome analysis. Nucleic Acids Res 36:D178–D183
Ping XL, Sun BF, Wang L et al (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res 24:177–189
Qu D, Zhou L, Wang W et al (2013) 5-Ethynylcytidine as a new agent for detecting RNA synthesis in live cells by "click" chemistry. Anal Biochem 434:128–135
Ramanathan A, Robb GB, Chan SH (2016) mRNA capping: biological functions and applications. Nucleic Acids Res 44:7511–7526
Reich S, Guilligay D, Pflug A et al (2014) Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature 516:361–366
Ries RJ, Zaccara S, Klein P et al (2019) m^6A enhances the phase separation potential of mRNA. Nature 571:424–428
Riml C, Amort T, Rieder D et al (2017) Osmium-mediated transformation of 4-Thiouridine to cytidine as key to study RNA dynamics by sequencing. Angew Chemie Int Ed Engl 56:13479–13483
Riquelme-Barrios S, Pereira-Montecinos C, Valiente-Echeverría F et al (2018) Emerging roles of N6-methyladenosine on HIV-1 RNA metabolism and viral replication. Front Microbiol 9:576
Ruigrok RW, Crépin T, Hart DJ et al (2010) Towards an atomic resolution understanding of the influenza virus replication machinery. Curr Opin Struct Biol 20:104–113
Sahin U, Karikó K, Türeci Ö (2014) MRNA-based therapeutics-developing a new class of drugs. Nat Rev Drug Discov 13:759–780
Schaefer M, Pollex T, Hanna K et al (2009) RNA cytosine methylation analysis by bisulfite sequencing. Nucleic Acids Res 37:e12
Schaefer M, Pollex T, Hnna K et al (2010) RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. Genes Dev 24:1590–1595
Schlundt A, Heinz GA, Janowski R et al (2014) Structural basis for RNA recognition in roquin-mediated post-transcriptional gene regulation. Nat Struct Mol Biol 21:671–678
Schofield JA, Duffy EE, Kiefer L et al (2018) TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding. Nat Methods 15:221–225
Schosserer M, Minois N, Angerer TB et al (2015) Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. Nat Commun 6:6158

Schott J, Stoecklin G (2010) Networks controlling mRNA decay in the immune system. Wiley Interdiscip Rev RNA 1:432–456
Schuberth-Wagner C, Ludwig J, Bruder AK et al (2015) A conserved histidine in the RNA sensor RIG-I controls immune tolerance to N1-2′O-methylated self RNA. Immunity 43:41–51
Schwalb B, Michel M, Zacher B et al (2016) TT-seq maps the human transient transcriptome. Science 352:1225–1228
Schwartz S, Mumbach MR, Jovanovic M et al (2014) Perturbation of m^6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. Cell Rep 8:284–296
Sharma S, Yang J, Watzinger P et al (2013) Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. Nucleic Acids Res 41:9062–9076
Sharma S, Langhendries JL, Watzinger P et al (2015) Yeast Kre33 and human NAT10 are conserved 18S rRNA cytosine acetyltransferases that modify tRNAs assisted by the adaptor Tan1/THUMPD1. Nucleic Acids Res 43:2242–2258
Sharma S, Yang J, van Nues R et al (2017) Specialized box C/D snoRNPs act as antisense guides to target RNA base acetylation. PLoS Genet 13:e1006804
Shatkin AJ, Manley JL (2000) The ends of the affair: capping and polyadenylation. Nat Struct Biol 7:838–842
Shi H, Wang X, Lu Z et al (2017) YTHDF3 facilitates translation and decay of N 6-methyladenosine-modified RNA. Cell Res 27:315–328
Shirahama S, Miki A, Kaburaki T et al (2020) Long non-coding RNAs involved in pathogenic infection. Front Genet 11:454
Sloan KE, Warda AS, Sharma S et al (2017) Tuning the ribosome: the influence of rRNA modification on eukaryotic ribosome biogenesis and function. RNA Biol 14:1138–1152
Squires JE, Preiss T (2010) Function and detection of 5-methylcytosine in eukaryotic RNA. Epigenomics 2:709–715
Squires JE, Patel HR, Nousch M et al (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40:5023–5033
Sugiyama K, Obayashi E, Kawaguchi A et al (2009) Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. EMBO J 28:1803–1811
Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9:465–476
Suzuki T, Ito S, Horikawa S et al (2014) Human NAT10 is an ATP-dependent rna acetyltransferase responsible for N4-acetylcytidine formation in 18S ribosomal RNA (rRNA). J Biol Chem 289:35724–35730
Takeuchi O, Akira S (2010) Pattern recognition receptors and inflammation. Cell 140:805–820
Tan B, Gao SJ (2018) RNA epitranscriptomics: regulation of infection of RNA and DNA viruses by N6-methyladenosine (m^6A). Rev Med Virol 28:e1983
Tan D, Zhou M, Kiledjian M et al (2014) The ROQ domain of Roquin recognizes mRNA constitutive-decay element and double-stranded RNA. Nat Struct Mol Biol 21:679–685
Tani H, Mizutani R, Salam KA et al (2012) Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. Genome Res 22:947–956
Tirumuru N, Zhao BS, Lu W et al (2016) N6-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 gag protein expression. elife 5:e15528
Tomari Y, Zamore PD (2005) Perspective: machines for RNAi. Genes Dev 19:517–529
Tomikawa C (2018) 7-methylguanosine modifications in transfer RNA (tRNA). Int J Mol Sci 19:4080
Trixl L, Lusser A (2019) The dynamic RNA modification 5-methylcytosine and its emerging role as an epitranscriptomic mark. Wiley Interdiscip Rev RNA 10:e1510
Tuorto F, Liebers R, Musch T et al (2012) RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat Struct Mol Biol 19:900–905
Uehata T, Takeuchi O (2017) Regnase-1 is an endoribonuclease essential for the maintenance of immune homeostasis. J Interf Cytokine Res 37:220–229

Väre VYP, Eruysal ER, Narendran A et al (2017) Chemical and conformational diversity of modified nucleosides affects tRNA structure and function. Biomol Ther 7:29
Vitali P, Kiss T (2019) Cooperative 2′-o-methylation of the wobble cytidine of human elongator $tRNA^{met}$(cat) by a nucleolar and a Cajal body-specific box C/D RNP. Genes Dev 33:741–746
Wang X, Lu Z, Gomez A et al (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505:117–120
Wei J, Liu F, Lu Z et al (2018) Differential m 6 A, m 6 A m , and m 1 A demethylation mediated by FTO in the cell nucleus and cytoplasm. Mol Cell 71 e5:973–985
Wei-Lin Popp M, Maquat LE (2013) Organizing principles of mammalian nonsense-mediated mRNA decay. Annu Rev Genet 47:139–165
Williams JV, Harris PA, Tollefson SJ et al (2004) Human Metapneumovirus and lower respiratory tract disease in otherwise healthy infants and children. N Engl J Med 350:443–450
Wolfe MB, Goldstrohm AC, Freddolino PL (2019) Global analysis of RNA metabolism using bio-orthogonal labeling coupled with next-generation RNA sequencing. Methods 155:88–103
Yamada T, Akimitsu N (2019) Contributions of regulated transcription and mRNA decay to the dynamics of gene expression. Wiley Interdiscip Rev RNA 10:e1508
Yamada T, Imamachi N, Onoguchi-Mizutani R et al (2018) 5′-bromouridine IP chase (BRIC)-seq to determine RNA half-lives. In: Methods in molecular biology. Humana Press, pp 1–13
Yang YGY, Wang L, Han X et al (2019) RNA 5-methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA decay. Mol Cell 75.e11:1188–1202
Yu YT, Meier UT (2014) RNA-guided isomerization of uridine to pseudouridine – Pseudouridylation. RNA Biol 11:1483–1494
Zaccara S, Ries RJ, Jaffrey SR (2019) Reading, writing and erasing mRNA methylation. Nat Rev Mol Cell Biol 20:608–624
Zhang X, Liu Z, Yi J et al (2012) The tRNA methyltransferase NSun2 stabilizes p16 INK4 mRNA by methylating the 3′2-untranslated region of p16. Nat Commun 3:712
Zhang LS, Liu C, Ma H et al (2019) Transcriptome-wide mapping of internal N7-methylguanosine methylome in mammalian mRNA. Mol Cell 74 e8:1304–1316
Zhao X, Yang YGY, Sun BF et al (2014) FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res 24:1403–1419
Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29
Zhong S, Li H, Bodi Z et al (2008) MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. Plant Cell 20:1278–1288
Züst R, Cervantes-Barragan L, Habjan M et al (2011) Ribose 2′-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. Nat Immunol 12:137–143

Deciphering RNA Methylation in Cancer

Daniela Barros-Silva, Elena S. Martens-Uzunova, and Carmen Jerónimo

Contents

Abstract Every single RNA nucleotide may undergo a variety of (post-) transcriptional chemical modifications. Historically, the inefficiency of detection methods and the difficulties in chemical structure elucidation have been a rate-limiting step in the discovery and functional analysis of ribonucleotide modifications. The current substantial progress in RNA modification profiling techniques launched epitranscriptomics as a new research field investigating this additional layer of information influencing cell physiology and disease development. RNA methylation is one of the most common and versatile chemical alterations found in the epitranscriptome, indicating a previously invisible code outside DNA and RNA sequences. Herein, we portray the historical evolution of strategies commonly used for overall and site-specific detection of methylated nucleotides in RNA and provide an overview of the relevance of these approaches for cancer biology research. We

D. Barros-Silva · C. Jerónimo (✉)
Cancer Biology and Epigenetics Group, Research Center of Portuguese Oncology Institute of Porto, Porto Comprehensive Cancer Center, Porto, Portugal
e-mail: carmenjeronimo@ipoporto.min-saude.pt

E. S. Martens-Uzunova
Department of Urology, Erasmus MC-University Medical Center Rotterdam, CA, Rotterdam, The Netherlands

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_9

also discuss the potential of third-generation sequencing methods for direct detection of RNA methylation and prospects of RNA methylation for anticancer therapy.

Keywords Epitranscriptomics · RNA methylation · Methylation-detection strategies · Cancer

1 Introduction

RNA modifications and their impact on RNA structure and function are currently emerging as a new research field in molecular and cellular biology called "epitranscriptomics" (Saletore et al. 2012; Li and Mason 2014; Accornero et al. 2020). So far, more than 150 different types of post-transcriptional modifications at the RNA level have been discovered across all domains of life and some of them are remarkably conserved throughout evolution. Among the vast variety of different RNA modifications, the addition of a methyl group ($-CH_3$) at different positions in RNA nucleosides is, so far, the most versatile and commonly found. Currently, the MODOMICS database (http://genesilico.pl/modomics/) (Boccaletto et al. 2018) lists 72 different modifications that contain a methyl group. Nonetheless, only 11 types of methylated nucleosides are reported in Eukaryotes phylogeny (Fig. 1). These RNA modifications are transcriptionally or post-transcriptionally added either at the nucleoside base or at the sugar-phosphate backbone of both coding and noncoding RNAs altering the properties of the transcript (Kellner et al. 2010). Among the wide variety of known methylated nucleosides, adenosine methylation at position N6 (N6-methyladenosine, m^6A) is the best understood and most frequent internal mark found in eukaryote messenger RNA (mRNA) and ribose methylation at position 2′-*O* (2-*O*-methylation, Nm) is the most common modification in abundant noncoding RNAs, such as ribosomal RNA (rRNAs) and transfer RNAs (tRNAs) (Romano et al. 2018). Altered patterns of these two types of RNA methylation are found in several pathologies, including cancer (Chen et al. 2019b; Dimitrova et al. 2019). Indeed, it has been shown that aberrant RNA methylation can modulate gene expression, leading to translational reprograming and cancer cell survival and proliferation (Delaunay and Frye 2019).

Although most of known RNA modifications were discovered in the last century, the lack of suitable tools to identify and map altered RNA nucleosides with sufficient molecular resolution and precision limited our perception of the spectrum of RNA modification abnormalities that may drive disease. Comprehensive technologies for their accurate detection and quantification began to be developed only in the past decade with the wide implementation of high-throughput next-generation sequencing (NGS) (Motorin and Helm 2019). The recent breakthroughs in new profiling methods for several common RNA modifications opened new possibilities for the exploitation of the epitranscriptomic landscape and the underlying regulatory mechanisms. These technical advances enabled new biological findings regarding the

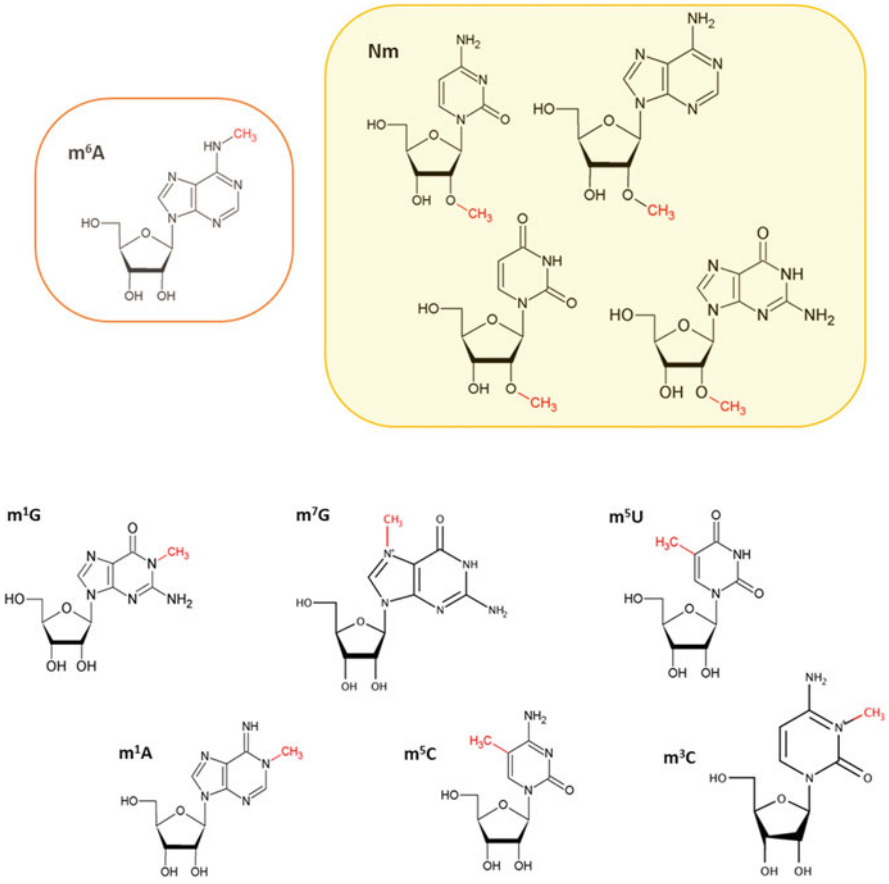


Fig. 1 Main eukaryotic RNA modifications that include the addition of a methyl group (CH_3, red font). M^6A and Nm are highlighted as the most common RNA methylation marks. Chemical structures were retrieved from the MODOMICS database (http://genesilico.pl/modomics/). Nm: 2′-*O*-methylation at any base (N); m^6A: N6-methyladenosine; m^1A: 1-methyladenosine; m^5C: 5-methylcytosine; m^5U: 5-methyluridine; m^1G: 1-methylguanosine; m^7G: 7-methylguanosine; m^3C: 3-methylcytidine

implication of RNA modifications in tumor heterogeneity and may ultimately assist the progress of precision medicine by contributing to the discovery and development of new therapeutic targets and valuable biomarkers for the personalized treatment of cancer patients.

In this chapter, we give a historical review of RNA methylation detection, from a cancer researcher's point of view, with a focus on the two most common types of RNA methylation, m^6A and Nm, in eukaryote RNA. We provide an overview of the currently acquired knowledge on m^6A and Nm RNA methylation patterns relevant in cancer biology and discuss the current most broadly used strategies to uncover functional importance of altered RNA methylation patterns, pointing out their

advantages and limitations. We also discuss epitranscriptomic modifications as potential druggable targets for cancer therapy.

2 RNA Methylation Biology: Focus on m^{6}A and Nm

2.1 N6-Methyladenosine (m^{6}A)

RNA methylation at position N6 in adenosine (m^{6}A) was first discovered in 1974 (Desrosiers et al. 1974) and ever since has been reported as regulating the generation and function of almost every known RNA class. Remarkably, m^{6}A is the most abundant internal modification in eukaryote mRNA and its methylation patterns are not randomly distributed, but rather enriched near stop codons and 3′-untranslated terminal regions (UTRs) (Meyer et al. 2012).

The deposition of m^{6}A is dynamically regulated by two very important classes of catalytic proteins: methyltransferases ("writers") and demethylases ("erasers"). A group of RNA binding proteins ("readers") are critical for the decoding of m^{6}A methylation as they mediate the recruitment of downstream functional complexes (Fig. 2a). A summary of known key players of m^{6}A modifications has been addressed in several recent publications (Meyer and Jaffrey 2017; Lobo et al. 2018).

As the best characterized RNA post-transcriptional alteration, the role of m^{6}A and its multicomponent regulatory complex in splicing, translation, and transcript stability is well recognized. The addition and removal of m^{6}A residues produces local

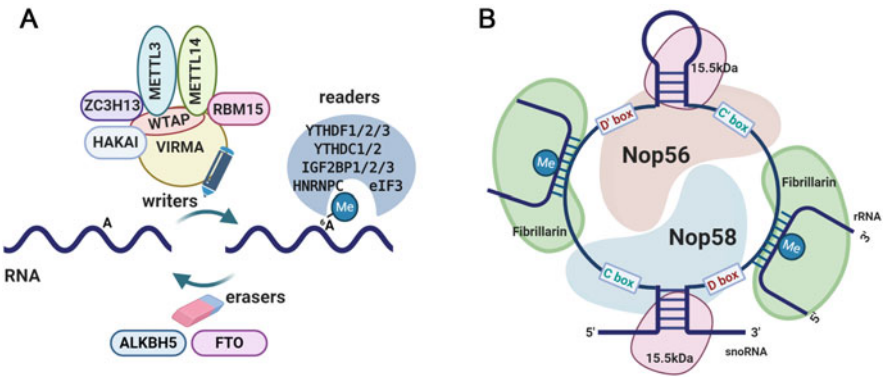


Fig. 2 The cellular machinery responsible for m^{6}A and Nm installation in RNA. (**a**) Me—methyl group; A—adenosine; writers—multicomponent complex that catalyzes the transformation of A to m^{6}A; erasers—enzymes responsible for m^{6}A demethylation; readers—proteins which recognize and bind to m^{6}A RNA modification. (**b**) The structure of the small nucleolar ribonucleoprotein (snoRNP) complex; Me—indicates the position of the nucleotide to be methylated in the target rRNA; Fibrillarin (methyltransferase), 15.5 kDa, Nop56, and Nop58 are the scaffold nucleolar proteins. Created by BioRender.com (https://biorender.com/)

changes in the RNA scaffold, which alters RNA folding and regulates RNA–RNA and RNA–protein interactions (Edupuganti et al. 2017).

2.2 2′-O-*Methylation (Nm)*

2′-*O*-methylation (Nm, where N can be any nucleoside A, C, G, or U) occurs when the hydrogen atom in the 2′-hydroxyl group (–OH) of the ribose moiety is replaced by a methyl group ($–CH_3$). Nm is a highly abundant and conserved modification found at multiple locations in ribosomal RNA (rRNA), transfer RNA (tRNA), and small nuclear RNA (snRNA) (Dimitrova et al. 2019). Regarding coding transcripts, Nm is present in the mRNA cap, but also internally within the coding sequences (CDS) (Dai et al. 2017).

Nm increases the hydrophobicity of RNA molecules, stabilizes their structure, protects the RNA backbone from enzymatic attack, and affects potential interactions with other RNAs, proteins, and molecules (Schwartz and Motorin 2017). Ribosomal RNAs (rRNAs) are acknowledged for carrying the highest number of Nm modifications that are fundamental for the biogenesis and function of ribosomes. In eukaryote ribosomes, 2′-*O*-methylation marks are added by the methyltransferase Fibrillarin. It is the catalytic unit of a small nucleolar ribonucleoprotein complex (snoRNP) guided to the specific ribosomal target sites by different C/D-box snoRNAs (SNORDs). Similarly to the interaction of microRNAs with their targets, C/D-box snoRNAs recognize their ribosomal target sites via base-pairing of complementary regions (Lin et al. 2011) (Fig. 2b). Nm modifications are clustered in functionally important parts of the ribosome, such as the tRNA binding sites or the peptidyltransferase center, influencing the intrinsic capabilities of ribosomes to translate mRNAs (Sloan et al. 2017). Remarkably, 2′-*O*-methylation sites are not always equally methylated and, depending on the position, residues may vary between fully and partially methylated. Fully modified sites (i.e., sites that are modified in 90–100% of all ribosomes) show little variation among different tissues, cell lines, and conditions, suggesting that these constitutively methylated positions are critical for ribosome biogenesis and operation. On the other hand, the portion of partially methylated sites (where only a proportion of the ribosomes would carry a methylation mark at the specific site) display much greater variability under different physiological or pathological conditions. This makes partially 2′-*O*-methylated ribosomal sites more likely to be involved in regulatory functions (Krogh et al. 2016).

Unlike m^6A, no "erasers" are known to act on the 2′*O*-methylated sites at rRNA, evidencing that RNA methylation reversibility is not a rule but rather an exception depending on the available machinery. Indeed, no apparent mechanisms of demethylation are reported for 2′*O*-methylated residues, perhaps also because of the relatively short half-life inherent to RNA (Baudrimont et al. 2017). Under such a scenario, the changes observed in the proportion of methylated positions at different conditions, for example in cancer cells (Sharma et al. 2017), or during embryonic

development (Hebras et al. 2020), should reflect the RNA-turnover-associated replacement of one set of rRNA by another (Wiener and Schwartz 2020), indicating the existence of a stringent regulatory apparatus. In the same line, the distinct methylation profiles mentioned above are more likely to reflect global alteration patterns of the enzymes or SNORDs responsible for the methylation establishment across different conditions (Marcel et al. 2013; Liang et al. 2019).

3 RNA Methylation Profiling Methods

RNA is known to hold methylated residues for more than half a century ago (Adler et al. 1958), and important discoveries have been made regarding the location and function of methylated bases. The characterization and measurement of RNA methylation depends on the position and abundance of the methyl residue and on the preexisting knowledge of consensus sequences and different methodologies are applied for its investigation (Fig. 3). The main approaches for RNA methylation detection are summarized in Table 1 and further discussed below.

Historically RNA modifications were first characterized from a structural point of view by resolving modified ribonucleotides based on changes in their migration properties via capillary electrophoresis (CE) or fluorescent labeling and one-dimensional or two-dimensional thin-layer chromatography (TLC). These analytical approaches allowed only the detection of global patterns of methylation, since they were mostly based on the chemical and physical properties of modified nucleotides in comparison to unmodified ones (Gupta and Randerath 1979). Additionally, as nearly all modifications result in an increase in the mass of canonical nucleosides, mass spectrometry (MS) became one of the most powerful methods to identify and quantify a broad range of methylated residues in a single sample, also enabling the detection of low abundant modifications. This technique was critical in the initial characterization of the majority of known modifications as it can detect any alteration that causes a mass change, both in the ribose and in the nucleotide base (Gaston and Limbach 2014). However, mass spectrometry experiments require relatively large RNA quantities as well as preexisting sequence information in order to generate consistent and reliable results (Wetzel and Limbach 2016).

Despite the insightful gains that traditional approaches brought to the epitranscriptomic field in the past, these techniques could not provide information about the exact position of individual modifications in the RNA sequence (Chen et al. 2019a). To overcome this major drawback, a plethora of strategies, including endonuclease digestion, antibody-based approaches, reverse transcription arrest, or chemical labeling/treatment followed by electrophoresis, PCR, or more recently, NGS, were developed to comprehensively map RNA methylation sites at the nucleotide resolution.

- *Endonuclease digestion-based approaches (ENDBA)* were developed on the principle that some enzymes may specifically recognize a particular RNA

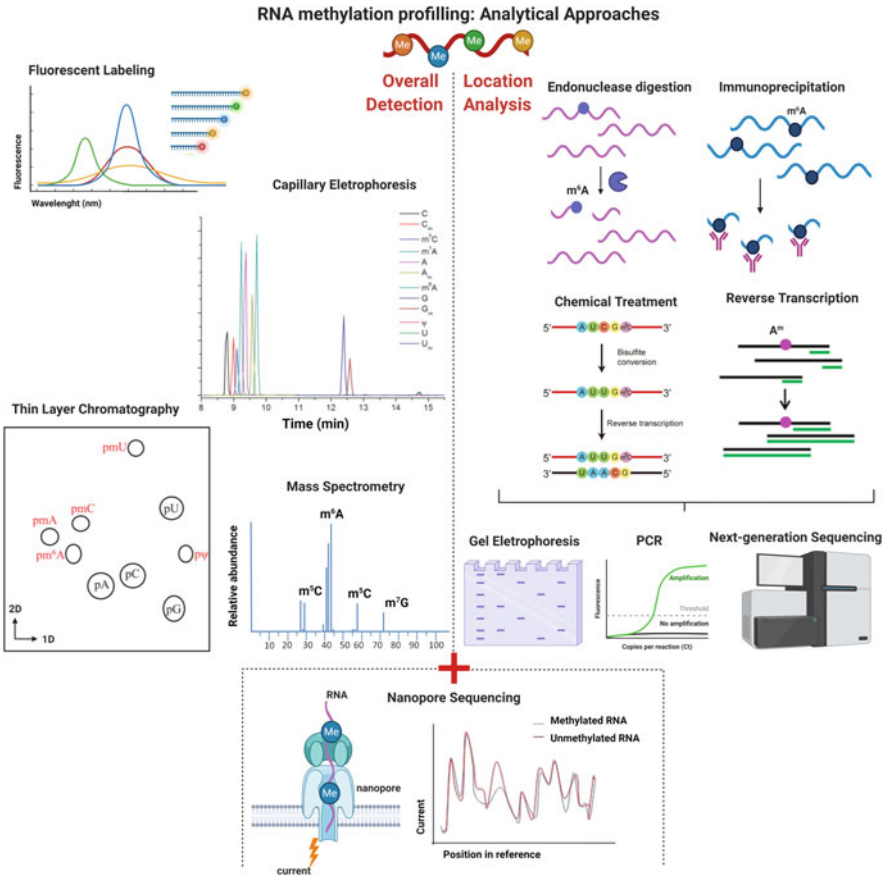


Fig. 3 Schematic illustration of the strategies for RNA methylation profiling. On the left the overall detection techniques and on the right the location analysis methods. Nanopore sequencing combines both approaches in a single analysis. Created by BioRender.com (https://biorender.com/)

modification. For instance, MazF endoribonuclease was recently identified as the first m^{6}A-sensitive RNA cleavage enzyme (Imanishi et al. 2017). RNase H is another example of a methyl-sensitive enzyme able to specifically cleave RNA at sites where the 2′-position of the sugar ring is not modified, being highly sensitive method for detecting 2′-*O*-methylated nucleotides in RNAs (Zhao and Yu 2004). However, the need of a large amount of purified RNA limits the use of these strategies, as individual mRNAs or lncRNAs are seldom efficiently purified.

- *Antibody-based approaches* (ABBA) rely on antibodies raised against modified ribonucleotide epitopes. Indeed, the wide usage of m^{6}A-specific antibodies fostered the development of antibodies against other much less abundant modifications. Nevertheless, the specificity of these antibodies is often insufficient and the results can be highly variable (Helm et al. 2019). Moreover, the antibodies are not always properly validated due to the low abundance of the modification;

Table 1 Main approaches for RNA methylation detection

Method	Principle	Main advantage	Limitations	Detected modification
Traditional methods				
Capillary electrophoresis, thin-layer chromatography, mass spectrometry	Based on physical and chemical properties of modified nucleotides	Currently function as powerful validation tools	Time-consuming techniques which require large RNA amounts and no information about the exact position of modifications in the RNA sequence	Any modification
Based on next-generation sequencing				
Endonuclease digestion	Based on enzymes that specifically recognize a particular RNA methylation mark	Methylation profiles are reproducible and quantitative	Require large amount of purified RNA	m^6A
Ab enrichment	Applied only for selected modified nucleotides when high-specificity antibodies against it are available	Allows enriching for RNAs harboring the modification	Limited Ab specificity, low resolution, and lack of stoichiometric quantification	m^1A, m^6A, m^5C
Reverse transcription	Applied when the modified nucleoside can interfere with base-pairing and mutated or abortive cDNA is generated	Suitable for mapping and relative quantification of unknown modification sites	The intensity of RT-stop varies for different sequence contexts, leading to uncertainties in quantification	m^1A, m^3C, m^1G, Nm
Chemical conversion	Chemical reagents are known to either selectively react with modified nucleosides or selectively affect non-modified nucleotides	Useful for silent RNA modifications which are otherwise undetected	Harsh chemical treatments can lead to RNA degradation	m^5C, m^7G, Nm
Based on direct sequencing				
Oxford Nanopore Technology (ONT)	Uses current signal from nanopore sequencing to detect the presence of RNA modifications	Direct measurement of modification states in the full-length RNA transcripts without need of amplification or reverse transcription	Currently, only a few research groups have the expertise for effective and reliable detection, measurement, and mapping of RNA modifications	Any modification

particularly on mRNAs, the generated datasets might resemble noise rather than specific modification patterns (Weichmann et al. 2020).

- *Reverse transcription-based approaches* (*RTBA*) represent an alternative to the techniques described above. RTBAs make use of the ability of RNA methylation to interfere with the interaction of RNA and the reverse transcriptase. Under specific conditions, reverse transcriptase enzymes may stall in the presence of particular modified nucleotides, leading to either base misincorporation or termination of reverse transcription (Brownlee and Cartwright 1977). RT-based mapping of RNA modifications can be considerably improved by means of radioactive or chemical labeling, allowing not only the identification of the modifications but also their location in RNA sequence. This phenomenon has enabled the development of methods such as the primer extension assay and (quantitative) RT-PCR-based techniques. The primer extension assay allows the mapping of modifications at single-nucleotide resolution. In contrast to other methods, it does not require purification steps and may be directly applied to total RNA. However, the addition of methylated groups to RNA nucleosides is often RT-silent and not always affects base-pair conformation, such as m7G (Motorin et al. 2007). For example, in the case of m^7G, a preliminary RNA chemical treatment with $NaBH_4$ is used as leverage to convert "silent" and otherwise invisible modified nucleotides into chemical adducts relatively stable and detected during the subsequent reverse transcription where they are permanently recorded as mutations in the cDNA sequence (Enroth et al. 2019).
- *Chemical conversion-based approaches* (CCBA), such as bisulfite treatment, have been widely used for mapping m^5C in DNA and the same principle may be applied to trace m5C in human mRNA and noncoding RNA (Squires et al. 2012). Alkaline treatment is another approach generally leading to RNA strand breaks when modifications are absent. Ribose 2′-*O*-methylation prevents alkaline RNA cleavage allowing for specific 2′-*O*-methylated nucleotides' detection (Krogh et al. 2016). However, in all cases the chemicals used are rather harsh and can lead to RNA damage, which may result in overall underrepresentation of the modified RNA pool.
- *NGS methodologies* for the analysis of RNA modifications developed in recent years revolutionized the epitranscriptomic field providing information at an unprecedented depth. Established methods for RNA modification mapping using deep sequencing were recently reviewed in detail (Motorin and Helm 2019). These new approaches enabled epitranscriptome-wide profiling at the single-nucleotide resolution. Importantly, major limitations of antibody-based technologies or methods involving polymerase pausing/arresting during reverse transcription were overcome by using these methodologies. Nonetheless, the in-depth epitranscriptome exploration is in its first steps, as just a few research groups have the expertise relevant for effective and reliable detection, measurement, and mapping of RNA modifications in general and in particular for m^6A and Nm RNA modifications. It is also important to have in mind that all sequencing approaches only provide inferences/models for the occurrence of a modification, but none directly exposes the true chemical nature of the modified

nucleic acid. Thus, among the epitranscriptomic community there is a high demand for new and more robust approaches capable of high-confidence and reproducible results. Nevertheless, technical progress and innovation are expected to significantly improve RNA methylation profiling in the near future.

- *Direct (third-generation) sequencing* methods, such as Oxford Nanopore Technologies (ONT), have great potential, although still in early stage of development. This new strategy preserves RNA integrity and might detect virtually any given RNA modification in native RNA molecules. Unlike NGS, direct sequencing techniques do not rely on light detection and labeling of nucleic acids. Instead, the nucleotide sequence and methylation profile are analyzed by measuring how different nucleotide sequences and the presence of modifications affect the electric charge of ribonucleotides (Garalde et al. 2018; Leger et al. 2019). Expectably, its upcoming applications can further facilitate studies on RNA modifications and allow for the assessment of their usefulness in clinical practice and disease management.

Summarizing, although the versatility of RNA methylation profiling methods is increasing, customized protocols must be individually set up and optimized for each RNA modification type, implying a prior knowledge of the RNA modification to be assessed before experimental design. This limitation precludes for now the ability to characterize the plasticity of the epitranscriptome in a systematic and unbiased manner. Besides the innovative approaches for site-specific detection of methylated residues, strategies for monitoring the presence and ratios of known methylated nucleotides are crucial when medical applications are envisioned. It is also important to consider that compared to transcriptome-wide analytical methods, monitoring approaches for clinical applications need to be cheaper, faster, and easier to perform to make them cost-effective and thus clinically useful.

3.1 Detection of m^6A

The earliest method used to identify m^6A in mRNA dates from 1974 and was carried out by incorporating radioactive isotopes into methylated purine rings (Desrosiers et al. 1974). The detection of m^6A residues using specific antibodies that recognize and precipitate m^6A-modified RNA was first reported in 1984 (Horowitz et al. 1984), but the coupling to NGS was only developed at a later stage in 2012 (Meyer et al. 2012), and it remains as one of the most commonly applied approaches despite its drawbacks.

Traditionally, dot blot is used to detect changes in the m^6A content of total RNA. Nevertheless, this method is only semi-quantitative as it does not separate RNA by size and therefore cannot assess individual target transcripts (Nagarajan et al. 2019). The distribution of m^6A among individual transcripts is determined by antibody-based techniques combined with NGS, the so-called MeRIP-Seq (Meyer et al. 2012). This method revealed that m^6A is particularly enriched in 3′-UTR regions and near

mRNA stop codons. The major pitfalls of this strategy are the low resolution of approximately 100 nucleotides and the requirement for a significant amount of input RNA. To overcome some of these issues, the m^6A individual-nucleotide-resolution cross-linking and immunoprecipitation (miCLIP) technique was developed in 2015 (Linder et al. 2015). The progress in high-throughput methodologies provided a resource for specific identification of m^6A consensus motifs, as well as knowledge of the potential m^6A-enriched regions. Therefore, these are fundamental strategies to profile the m6A landscape in normal and cancer cells and to understand how changes in m^6A position and distribution may be involved in carcinogenesis. Nevertheless, the mentioned NGS approaches do not quantify the m^6A stoichiometry, i.e., the fraction of m^6A-modified residues. This is an important bottleneck to overcome in elucidating the biological function of m^6A in disease as the ratio of modified to unmodified residues in the transcript body might also carry information on modification dynamics in response to a stimulus.

To address this challenge, new methods to unambiguously determine the m^6A status at single-nucleotide resolution emerged. When sequence information is already known, site-specific cleavage and radioactive labeling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET) can be used to assay the stoichiometry of m^6A methylation (Liu et al. 2013). More recently, ArrayStar Inc. released a m^6A single-nucleotide array platform that can be used for the detection of m^6A modifications at single-nucleotide resolution and for quantification of m^6A stoichiometry (ArrayStar Inc 2020). The microarray chemistry is based on the specific properties of the methyl-sensitive MazF RNase, a so far unique enzyme that specifically recognizes and cleaves m^6A containing RNA (Imanishi et al. 2017). The introduction of this technique overcomes also another downside of the antibody-based techniques, namely the cross-reactivity of m^6A antibodies with similar modifications, e.g., m^6Am. Remarkably, the ArrayStar technology quantifies the fraction of m^6A residues based on specific transcript abundancy providing more reliable and consistent results. This is of particular interest for mRNAs, since they represent a limited percentage of total RNA and noncoding RNAs are often heavily methylated. Lastly, techniques without immunoprecipitation require much lower amounts of RNA, being more suitable for samples of limited supply and quantity, such as clinical specimens, and particularly body fluids.

Despite these significant technology improvements in the last few years, m^6A profiling at the single-nucleotide resolution remains a challenge especially for highly structured transcripts. In many cases, the interference of secondary RNA structures near the modification site often reduces accuracy as the localization of m^6A sites in double-stranded stem regions may hinder the access of m^6A antibody and inhibits the MazF enzyme, which cannot cleave double-stranded RNA.

ONT direct RNA sequencing is a newly emerging technique that can overcome some of the major limitations in m6A profiling. Indeed, the ability of ONT to directly detect m^6A RNA modifications in endogenous transcripts was recently demonstrated (Lorenz et al. 2020). Another study showed that m^6A RNA modification patterns can be detected with high accuracy. The created *EpiNano* algorithm can calculate m^6A RNA methylation with ~90% accuracy (Liu et al. 2019). This third-generation

approach provides numerous benefits over traditional methodologies including isoform-specific context, single experimental pipeline, and simplified bioinformatics detection.

3.2 *Detection of 2′-O-methylation (Nm)*

Several classical methods, such as chromatography and mass spectrometry, although not specific for detecting particularly modified residues, but RNA modifications in general, were developed in the past for the detection of Nm. Nevertheless, the full spectrum of specific 2′-*O*-methylated RNA sites in the transcriptome is still scarcely described, in part because of the complexity of the existing approaches. Detection of Nm remains challenging due to its inert chemical nature and the lack of an antibody that would allow selective recognition, reason why material evidence about its impact in cancer development is largely unexplored yet.

In 1997, a specific detection strategy based on an endonuclease digestion such as RNase H was developed to locate 2′-*O*-methylation in RNA (Yu et al. 1997). Later, this approach was combined with high-performance liquid chromatography coupled with tandem mass spectrometry (HPLC-MS/MS) for quantitative assessment of 2′-*O*-methylation status of the target nucleotide (Yang et al. 2016).

Reverse transcription (RT)-dependent primer extension is a successful strategy for Nm detection based on the knowledge that RT enzymes pause at 2′-*O*-methylated sites when deoxynucleotide triphosphate (dNTP) concentration is very low (Maden 2001; Maden et al. 1995). The method is position specific as the RT reaction is performed with target-specific labeled primers and the obtained abortive cDNA sequences can be analyzed by urea-gel electrophoresis. Indeed, most ribosomal RNA Nm sites were successfully mapped using this technique (Rebane et al. 2002). Recently, reverse transcription at low dNTP concentration followed by PCR (RTL-P) demonstrated precise mapping and superior sensitivity without using radiolabeled or fluorescent primers compared with previous techniques (Dong et al. 2012).

During the past decade with the development of next-generation sequencing technologies, researchers gathered efforts to improve and fast-track 2′-*O*-methylation sites' identification and quantification. Several protocols using sequencing methods preceded by different types of treatment were developed for mapping Nm modifications. The first of these techniques called RiboMethSeq applies a chemical conversion-sequencing method, which is based on the protection of RNA by 2′-*O*-methylation against alkaline hydrolysis (Krogh et al. 2016). Since the protection signal is highly dependent on the methylation level, RiboMethSeq also enables quantification of the relative methylation levels at specific positions. This technique was originally developed for Ion Torrent sequencing (Krogh et al. 2016; Birkedal et al. 2015) and later adapted for Illumina sequencing (Marchand et al. 2016). Illumina-RiboMethSeq is currently a commonly used strategy to profile 2′-*O*-methylation sites in eukaryote rRNA because it requires only a low amount of input RNA

and can be performed without further adjustments using the full range of Illumina sequencers. As rRNA represents about 90% of total RNA, the analysis can be directly performed without the need of prior fractionation or enrichment.

Several other NGS-based approaches were developed in the last 5 years. In 2017, *Incarnato* and colleagues employed 2′OMe-Seq to map reverse transcription stops generated by primer extension at low dNTP concentration, allowing the identification and relative quantitation of 2′-*O*-methylated residues (Incarnato et al. 2017). However, regardless of being relatively sensitive, partial methylation is still difficult to quantify using such RT-based methods, especially in areas with multiple 2′-*O*-methylated residues positioned in close proximity. In the same year, Nm-seq also emerged as new 2′-*O*-methylation mapping strategy with single-base resolution. This method benefits from the resistance of 2′-*O*-methylated-3′-terminal riboses to periodate oxidation (Dai et al. 2017). Nm-seq appears to be more efficient for the mapping of 2′-*O*-methylation sites in mRNA. Insightful description and comparison of sequencing-based Nm detection methods was published recently (Motorin and Marchand 2018).

Currently, nanopore sequencing is not yet available for 2′-*O*-methylation RNA marks profiling. However, first encouraging results were recently reported (Begik et al. 2020), opening the road toward the development of third-generation techniques for future detection of Nm patterns in RNA.

4 Relevance of RNA Methylation Profiling Methods in Cancer

In the past decade, increasing evidence was gathered suggesting that RNA modification pathways are often dysregulated in human cancers (Barbieri and Kouzarides 2020; Huang et al. 2020). Thus, the constant development and improvement of RNA methylation profiling methods may provide novel insights into tumor biology and afford innovative tools for patient management, aiding in diagnosis and prognosis assessment, as well as the identification of new therapeutic targets.

Remarkably, m^6A RNA methylation plays a significant role in cancer onset and progression with particular influence in proliferation, invasion, metastasis, and immune response (Yang et al. 2020; Delaunay and Frye 2019). The functional relevance and alteration frequency of m^6A modification in multiple human cancers were recently reviewed (Zhou et al. 2020; Shen et al. 2020; Lobo et al. 2018). Additionally, the role of m^6A RNA methylation in clinical application has been broadly studied. The expression levels of both m^6A-related genes and proteins are likely to be potential cancer diagnostic and prognostic biomarkers (Zhao and Cui 2019; Cho et al. 2018). Furthermore, m^6A methylation is also involved in drug resistance and may provide new targets for the research and development of targeted cancer therapies (Zhu et al. 2019; Klinge et al. 2019).

The detection of Nm modifications has been a challenge for a long time and its role is still largely unexplored in disease. Similar to m^6A, Nm patterns at rRNA prove to be deregulated in cancer in order to ensure cancer cell survival and accommodate the enhanced protein needs associated with continuous proliferation (Erales et al. 2017). The rapid methodological advances in Nm profiling in recent years have opened the research field and first results describing the involvement of Nm in tumorigenesis are now emerging. For instance, dysregulated levels of C/D-box snoRNA correlated with alterations in Nm at rRNA, and these alterations can contribute to cancer progression and outcome (Belin et al. 2009). Notably, global loss of C/D-box snoRNAs with concomitant loss of Nm marks in rRNA resulted in decreased leukemia self-renewal potential (Zhou et al. 2017; Pauli et al. 2020). In colorectal cancer increased ribosomal Nm levels were associated with enhanced translational activity of downstream targets, thereby mediating proliferation (Wu et al. 2020).

In view of these new findings, understanding of the ribosomal Nm code and the mechanisms by which it influences cancer cell survival and proliferation becomes imperative for a better and more insightful understanding of tumor biology and the subsequent development of new strategies for cancer diagnosis and treatment.

5 Prospects for the Use of RNA Methylation for Drug Targeting

Despite the versatility of RNA modifications and RNA methylation, most of the scientific attention has been traditionally focused on chromatin and protein modifications. Scientific advances in the latter turned into substantial translational efforts aiming to adopt generated findings for cancer intervention and pharmacological targeting, ultimately leading to a considerable number of new therapies (Zhang et al. 2009). Likewise, the current excitement around epitranscriptome's profiling is driven not only by the demanding knowledge on the biology and function of RNA modifications but also by their very promising although unexploited biomarker and therapeutic potential. In recent years, RNA methylation is fueling industrial interest as a potential source of novel cancer treatment approaches as several companies already explore the possibilities to develop a new class of medicines targeting either RNA modification itself and/or RNA-modifying enzymes in cancer (Storm Therapeutics Ltd 2019; Epics Therapeutics 2020).

Traditionally, the removal or inhibition of RNA modification activity has been done by genetic techniques. However, to assess whether the RNA editing machinery makes a feasible cancer drug target, scientists need to explore several paths toward the successful development of new therapeutics. A key factor to achieve this goal is the upgrading and constant improvement of RNA profiling techniques able to track RNA modifications. The gold standard proof for a drug's target is the identification of modified RNA residues that confer resistance in a cellular context. Thus, new

high-throughput technologies hold promise for the informed design of RNA-based small-molecule therapeutics, and also for the better assessment of the potential drugs' efficacy (Sztuba-Solinska et al. 2019).

So far, m^6A demethylation's small-molecule inhibitors targeting FTO and ALKBH5 have been proposed. Nevertheless, their pharmaceutical effects in vivo or in clinical trials are yet to be fully verified, substantially because of the lack of specificity (Niu et al. 2018). Moreover, the pharmacological inhibition of the RNA m^6A methyltransferase METTL3 in vivo has shown strong antitumor effects in physiologically and clinically relevant models of acute myeloid leukemia (AML) (Tzelepis et al. 2019). Furthermore, the depletion of METTL3 has been shown to sensitize pancreatic cancer cells to other common anticancer treatments such as gemcitabine, 5-fluorouracil, cisplatin, and irradiation (Taketo et al. 2018). This suggests that the addition of RNA methylation inhibitors as a part of combination therapy might be valuable in fighting drug resistance in advanced cancer stages (Song et al. 2020).

The potential scope of epitranscriptomics' drug discovery expands beyond the targeting of m^6A machinery. For example, changes in the 2′-*O*-methylation patterns of rRNA may influence the performance of specialized ribosome populations in cancer cells that are associated with the increased translation of a subset of key oncoproteins (Dinman 2016). Indeed, rRNA synthesis and ribosome biogenesis are valid druggable pathways in cancer. Several compounds such as CX-5461, CX-3543, and BMH-21 are known to effectively prevent ribosome production by inhibiting the function of RNA Polymerase I (Pol I) and have entered clinical trials (Catez et al. 2019). The rationale for establishing ribosome biogenesis as a druggable pathway in cancer therapeutics further resides in the observation that cancer cells are "addicted" to high rates of ribosome and protein synthesis. Hence, drugs inhibiting rRNA transcription and/or ribosomal maturation would have stronger effect on cancer cells compared to non-malignant ones (Brighenti et al. 2015). Nevertheless, the approach is still associated with high off-target effects and, despite appealing, such drugs still have some limitations in terms of specificity and sensitivity. For instance, the tested compounds only have a limited selectivity toward rRNA genes, and, in theory, they can interact with any DNA sequence that is GC-rich or contains G-quadruplex (Catez et al. 2019). Furthermore, acquired resistance to CX-5461 inhibitory action was already reported, illustrating that the anticancer activity of the drug can somehow be bypassed (Bruno et al. 2020). Resistance to small-molecule inhibitors often relies on chemical modifications of the ribosomes as altered Nm pattern at a single rRNA residue may be sufficient to prevent efficient drug action (Lin et al. 2018). Under this scenario, targeting of specific ribosomal methylation marks may be a way to avoid or postpone the development of drug resistance or to increase the specificity of ribosomal biosynthesis inhibitors to cancer ribosomes (Zhang et al. 2019). Currently, such alternative therapies exist mainly as a concept and their effectiveness must be first established preclinically to enable safe extrapolation and development toward real-life solutions. Nevertheless, several arguments can support the suitability of ribosomal modifications as targets for cancer therapy. First of all, the presence of cell-type, tissue-specific, and cancer-associated

snoRNA expression patterns (Liang et al. 2019) and the possibility to target them without serious effect on essential cellular processes in human noncancer cell lines (Filippova et al. 2019) create opportunities to develop more precise therapeutic strategies targeting the cancer ribosome. In addition, the rapid developments in the field of antisense RNA-targeting and drug delivery systems can also anticipate the opportunity to target the snoRNA component in cancer (Quemener et al. 2020).

At present, epitranscriptomics-based medicine is still taking its first steps. The development of potent anticancer drugs targeting RNA methylation with higher specificity is still warranted. Moreover, a careful dissection of tissue- and cell-specific drug-induced biological changes using in vivo models is a prerequisite in rendering the RNA modification-based medicines to the near future. Upgraded RNA modification's detection tools will be critical for recruiting and monitoring volunteers into clinical studies testing inhibitors of epitranscriptomics machinery.

6 Conclusion

Overall, the molecular mechanisms that regulate RNA methylation in cancer biology remain largely elusive and require further systematic exploration. The latest technical advances in RNA methylation detection anticipate novel insights in RNA biology that can open new gateways for clinical application. We expect that this portrait of current insights and limitations on RNA methylation field will provide an opportunity for researchers to deepen the knowledge on epitranscriptomic regulation of cancer proliferation. Several specialized NGS-based protocols were developed to profile different RNA methylation marks, but no universal or single best way to identify the wide variety of RNA methylated residues is available so far. Indeed, the lack of a broad strategy for simultaneous mapping of different RNA modification types is still a major challenge in the field. The use of several of the abovementioned strategies in combination is currently the most suitable approach to gain insight in the quantity and in distribution of modified residues in situ. New methods are continually emerging, not only for transcriptome-wide detection of putative methylated sites, but also for fast and easy monitoring of known methylated residues at a given position. It is expected that the field of epitranscriptomics will undergo rapid development in the coming years, allowing the decoding of the dynamic RNA methylation code, elucidating its acknowledged impact on carcinogenesis, and generating a new layer of putative anticancer targets.

References

Accornero F, Ross RL, Alfonzo JD (2020) From canonical to modified nucleotides: balancing translation and metabolism. Crit Rev Biochem Mol Biol 55:525–540

Adler M, Weissmann B, Gutman AB (1958) Occurrence of methylated purine bases in yeast ribonucleic acid. J Biol Chem 230:717–723
ArrayStar Inc (2020) Limitations of m6A-seq and solutions of arraystar m6A single nucleotide arrays. https://www.arraystar.com/reviews/limitations-of-m6aseq-and-solutions-of-arraystar-m6a-single-nucleotide-arrays/. Accessed 11 Nov 2020
Barbieri I, Kouzarides T (2020) Role of RNA modifications in cancer. Nat Rev Cancer 20:303–322
Baudrimont A, Voegeli S, Viloria EC et al (2017) Multiplexed gene control reveals rapid mRNA turnover. Sci Adv 3:e1700006
Begik O, Lucas MC, Ramirez JM et al (2020) Decoding ribosomal RNA modification dynamics at single molecule resolution. bioRxiv. https://doi.org/10.1101/2020.07.06.189969
Belin S, Beghin A, Solano-Gonzàlez E et al (2009) Dysregulation of ribosome biogenesis and translational capacity is associated with tumor progression of human breast cancer cells. PLoS One 4:e7147
Birkedal U, Christensen-Dalsgaard M, Krogh N et al (2015) Profiling of ribose methylations in RNA by high-throughput sequencing. Angew Chem Int Ed Engl 54:451–455
Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46(D1):D303–d307
Brighenti E, Treré D, Derenzini M (2015) Targeted cancer therapy with ribosome biogenesis inhibitors: a real possibility? Oncotarget 6:38617–38627
Brownlee GG, Cartwright EM (1977) Rapid gel sequencing of RNA by primed synthesis with reverse transcriptase. J Mol Biol 114:93–117
Bruno PM, Lu M, Dennis KA et al (2020) The primary mechanism of cytotoxicity of the chemotherapeutic agent CX-5461 is topoisomerase II poisoning. Proc Natl Acad Sci USA 117:4053–4060
Catez F, Dalla Venezia N, Marcel V et al (2019) Ribosome biogenesis: an emerging druggable pathway for cancer therapeutics. Biochem Pharmacol 159:74–81
Chen B, Yuan B-F, Feng Y-Q (2019a) Analytical methods for deciphering RNA modifications. Anal Chem 91:743–756
Chen X-Y, Zhang J, Zhu J-S (2019b) The role of m6A RNA methylation in human cancer. Mol Cancer 18:103
Cho SH, Ha M, Cho YH et al (2018) ALKBH5 gene is a novel biomarker that predicts the prognosis of pancreatic cancer: a retrospective multicohort study. Ann Hepatobiliary Pancreat Surg 22 (4):305–309
Dai Q, Moshitch-Moshkovitz S, Han D et al (2017) Nm-seq maps 2′-O-methylation sites in human mRNA with base precision. Nat Methods 14:695–698
Delaunay S, Frye M (2019) RNA modifications regulating cell fate in cancer. Nat Cell Biol 21:552–559
Desrosiers R, Friderici K, Rottman F (1974) Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc Natl Acad Sci USA 71:3971–3975
Dimitrova DG, Teysset L, Carré C (2019) RNA 2′-O-methylation (Nm) modification in human diseases. Genes (Basel) 10:117
Dinman JD (2016) Pathways to specialized ribosomes: the Brussels lecture. J Mol Biol 428:2186–2194
Dong ZW, Shao P, Diao LT et al (2012) RTL-P: a sensitive approach for detecting sites of 2′-O-methylation in RNA molecules. Nucleic Acids Res 40:e157
Edupuganti RR, Geiger S, Lindeboom RGH et al (2017) N6-methyladenosine (m6A) recruits and repels proteins to regulate mRNA homeostasis. Nat Struct Mol Biol 24:870–878
Enroth C, Poulsen LD, Iversen S et al (2019) Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing. Nucleic Acids Res 47:e126–e126
Epics Therapeutics (2020) EPICS invents medicines. https://www.epicstherapeutics.com/. Accessed 8 Dec 2020

Erales J, Marchand V, Panthu B et al (2017) Evidence for rRNA 2′-O-methylation plasticity: control of intrinsic translational capabilities of human ribosomes. Proc Natl Acad Sci USA 114:12934–12939
Filippova JA, Matveeva AM, Zhuravlev ES et al (2019) Are small nucleolar RNAs "CRISPRable"? A report on box C/D small nucleolar RNA editing in human cells. Front Pharmacol 10:1246
Garalde DR, Snell EA, Jachimowicz D et al (2018) Highly parallel direct RNA sequencing on an array of nanopores. Nat Methods 15:201–206
Gaston KW, Limbach PA (2014) The identification and characterization of non-coding and coding RNAs and their modified nucleosides by mass spectrometry. RNA Biol 11:1568–1585
Gupta RC, Randerath K (1979) Rapid print-readout technique for sequencing of RNA's containing modified nucleotides. Nucleic Acids Res 6:3443–3458
Hebras J, Krogh N, Marty V et al (2020) Developmental changes of rRNA ribose methylations in the mouse. RNA Biol 17:150–164
Helm M, Lyko F, Motorin Y (2019) Limited antibody specificity compromises epitranscriptomic analyses. Nat Commun 10:5669
Horowitz S, Horowitz A, Nilsen TW et al (1984) Mapping of N6-methyladenosine residues in bovine prolactin mRNA. Proc Natl Acad Sci USA 81:5667–5671
Huang H, Weng H, Deng X et al (2020) RNA modifications in cancer: functions, mechanisms, and therapeutic implications. Ann Rev Cancer Biol 4:221–240
Imanishi M, Tsuji S, Suda A et al (2017) Detection of N 6-methyladenosine based on the methyl-sensitivity of MazF RNA endonuclease. Chem Commun 53:12930–12933
Incarnato D, Anselmi F, Morandi E et al (2017) High-throughput single-base resolution mapping of RNA 2′-O-methylated residues. Nucleic Acids Res 45:1433–1441
Kellner S, Burhenne J, Helm M (2010) Detection of RNA modifications. RNA Biol 7:237–247
Klinge CM, Piell KM, Tooley CS et al (2019) HNRNPA2/B1 is upregulated in endocrine-resistant LCC9 breast cancer cells and alters the miRNA transcriptome when overexpressed in MCF-7 cells. Sci Rep 9:1–22
Krogh N, Jansson MD, Häfner SJ et al (2016) Profiling of 2′-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. Nucleic Acids Res 44:7884–7895
Leger A, Amaral PP, Pandolfini L et al (2019) RNA modifications detection by comparative nanopore direct RNA sequencing. BioRxiv:843136
Li S, Mason CE (2014) The pivotal regulatory landscape of RNA modifications. Ann Rev Genomics Hum Genet 15:127–150
Liang J, Wen J, Huang Z et al (2019) Small nucleolar RNAs: insight into their function in Cancer. Front Oncol 9(587)
Lin J, Lai S, Jia R et al (2011) Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. Nature 469:559–563
Lin J, Zhou D, Steitz TA et al (2018) Ribosome-targeting antibiotics: modes of action, mechanisms of resistance, and implications for drug design. Annu Rev Biochem 87:451–478
Linder B, Grozhik AV, Olarerin-George AO et al (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods 12:767–772
Liu N, Parisien M, Dai Q et al (2013) Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA 19:1848–1856
Liu H, Begik O, Lucas MC et al (2019) Accurate detection of m6A RNA modifications in native RNA sequences. Nat Commun 10:4079
Lobo J, Barros-Silva D, Henrique R et al (2018) The emerging role of epitranscriptomics in cancer: focus on urological tumors. Genes (Basel) 9:552
Lorenz DA, Sathe S, Einstein JM et al (2020) Direct RNA sequencing enables m6A detection in endogenous transcript isoforms at base-specific resolution. RNA 26:19–28
Maden BE (2001) Mapping 2′-O-methyl groups in ribosomal RNA. Methods (San Diego, Calif) 25:374–382

Maden BE, Corbett ME, Heeney PA et al (1995) Classical and novel approaches to the detection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA. Biochimie 77:22–29
Marcel V, Ghayad SE, Belin S et al (2013) p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. Cancer Cell 24:318–330
Marchand V, Blanloeil-Oillo F, Helm M et al (2016) Illumina-based RiboMethSeq approach for mapping of 2′-O-Me residues in RNA. Nucleic Acids Res 44:e135
Meyer KD, Jaffrey SR (2017) Rethinking m(6)A readers, writers, and erasers. Ann Rev Cell Dev Biol 33:319–342
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149:1635–1646
Motorin Y, Helm M (2019) Methods for RNA modification mapping using deep sequencing: established and new emerging technologies. Genes (Basel) 10:35
Motorin Y, Marchand V (2018) Detection and analysis of RNA ribose 2′-O-methylations: challenges and solutions. Genes (Basel) 9:642
Motorin Y, Muller S, Behm-Ansmant I et al (2007) Identification of modified residues in RNAs by reverse transcription-based methods. Methods Enzymol 425:21–53
Nagarajan A, Janostiak R, Wajapeyee N (2019) Dot blot analysis for measuring global N 6-methyladenosine modification of RNA. In: Epitranscriptomics. Springer, pp 263–271
Niu Y, Wan A, Lin Z et al (2018) N6-Methyladenosine modification: a novel pharmacological target for anti-cancer drug development. Acta Pharm Sin B 8:833–843
Pauli C, Liu Y, Rohde C et al (2020) Site-specific methylation of 18S ribosomal RNA by SNORD42A is required for acute myeloid leukemia cell proliferation. Blood 135:2059–2070
Quemener AM, Bachelot L, Forestier A et al (2020) The powerful world of antisense oligonucleotides: from bench to bedside. Wiley Interdiscip Rev RNA e1594
Rebane A, Roomere H, Metspalu A (2002) Locations of several novel 2′-O-methylated nucleotides in human 28S rRNA. BMC Mol Biol 3:1–5
Romano G, Veneziano D, Nigita G et al (2018) RNA methylation in ncRNA: classes, detection, and molecular associations. Front Genet 9:243
Saletore Y, Meyer K, Korlach J et al (2012) The birth of the Epitranscriptome: deciphering the function of RNA modifications. Genome Biol 13:175
Schwartz S, Motorin Y (2017) Next-generation sequencing technologies for detection of modified nucleotides in RNAs. RNA Biol 14:1124–1137
Sharma S, Marchand V, Motorin Y et al (2017) Identification of sites of 2′-O-methylation vulnerability in human ribosomal RNAs by systematic mapping. Sci Rep 7:11490
Shen H, Lan Y, Zhao Y et al (2020) The emerging roles of N6-methyladenosine RNA methylation in human cancers. Biomark Res 8:24
Sloan KE, Warda AS, Sharma S et al (2017) Tuning the ribosome: the influence of rRNA modification on eukaryotic ribosome biogenesis and function. RNA Biol 14:1138–1152
Song H, Liu D, Dong S et al (2020) Epitranscriptomics and epiproteomics in cancer drug resistance: therapeutic implications. Signal Transduct Targeted Ther 5:1–16
Squires JE, Patel HR, Nousch M et al (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40:5023–5033
Storm Therapeutics Ltd (2019) RNA epigenetics is coming of age. https://www.stormtherapeutics.com/news-events/storm-in-the-news/rna-epigenetics-is-coming-of-age/. Accessed 1 Dec 2020
Sztuba-Solinska J, Chavez-Calvillo G, Cline SE (2019) Unveiling the druggable RNA targets and small molecule therapeutics. Bioorg Med Chem 27:2149–2165
Taketo K, Konno M, Asai A et al (2018) The epitranscriptome m6A writer METTL3 promotes chemo- and radioresistance in pancreatic cancer cells. Int J Oncol 52:621–629
Tzelepis K, De Braekeleer E, Yankova E et al (2019) Pharmacological inhibition of the RNA m6A writer METTL3 as a novel therapeutic strategy for acute myeloid leukemia. American Society of Hematology, Washington, DC

Weichmann F, Hett R, Schepers A et al (2020) Validation strategies for antibodies targeting modified ribonucleotides. RNA 26:1489–1506
Wetzel C, Limbach PA (2016) Mass spectrometry of modified RNAs: recent developments. Analyst 141:16–23
Wiener D, Schwartz S (2020) The epitranscriptome beyond m6A. Nat Rev Genet 1–13
Wu H, Qin W, Lu S et al (2020) Long noncoding RNA ZFAS1 promoting small nucleolar RNA-mediated 2′-O-methylation via NOP58 recruitment in colorectal cancer. Mol Cancer 19:95
Yang J, Sharma S, Watzinger P et al (2016) Mapping of complete set of ribose and base modifications of yeast rRNA by RP-HPLC and mung bean nuclease assay. PLoS One 11: e0168873
Yang G, Sun Z, Zhang N (2020) Reshaping the role of m6A modification in cancer transcriptome: a review. Cancer Cell Int 20:353
Yu YT, Shu MD, Steitz JA (1997) A new method for detecting sites of 2′-O-methylation in RNA molecules. RNA 3:324–331
Zhang J, Yang PL, Gray NS (2009) Targeting cancer with small molecule kinase inhibitors. Nat Rev Cancer 9:28–39
Zhang D, Zhou J, Gao J et al (2019) Targeting snoRNAs as an emerging method of therapeutic development for cancer. Am J Cancer Res 9:1504–1516
Zhao X, Cui L (2019) Development and validation of a m6A RNA methylation regulators-based signature for predicting the prognosis of head and neck squamous cell carcinoma. Am J Cancer Res 9:2156
Zhao X, Yu YT (2004) Detection and quantitation of RNA base modifications. RNA 10:996–1002
Zhou F, Liu Y, Rohde C et al (2017) AML1-ETO requires enhanced C/D box snoRNA/RNP formation to induce self-renewal and leukaemia. Nat Cell Biol 19:844–855
Zhou Z, Lv J, Yu H et al (2020) Mechanism of RNA modification N6-methyladenosine in human cancer. Mol Cancer 19:104
Zhu L, Zhu Y, Han S et al (2019) Impaired autophagic degradation of lncRNA ARHGAP5-AS1 promotes chemoresistance in gastric cancer. Cell Death Dis 10:1–16

Roles of m^{6}A RNA Modification in Normal Development and Disease

Mohammad Burhan Uddin, Zhishan Wang, and Chengfeng Yang

Contents

Abstract During normal developmental and disease processes gene expression plays a central role. Multiple layers of regulation play roles in expression of developmental regulatory genes in a timely manner to maintain the developmental process smoothly and in order. Epitranscriptomic gene expression regulation through post-

M. B. Uddin · Z. Wang · C. Yang (✉)
Department of Toxicology and Cancer Biology, College of Medicine, University of Kentucky, Lexington, KY, USA

Division of Cancer Biology, Department of Medicine, MetroHealth Medicine, Case Western Reserve University School of Medicine, Cleveand, OH, USA
e-mail: chengfeng.yang@case.edu

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_10

transcriptional m^6A RNA modification has emerged as a novel concept of gene expression regulation in recent years. The reversible m^6A modification which is imprinted in the RNA transcripts by methyltransferases called "writers" and removed by demethylases called "erasers" has been discovered in both protein coding mRNAs and noncoding RNAs. The recognition of this modification mark by RNA-binding proteins called "readers" determines the stability, translation, or translocation of the transcripts carrying the m^6A modification. By controlling the life cycle of the mRNA transcript of a particular gene m^6A thus up/downregulates its expression. This mode of gene expression regulation not only is important for normal developmental processes but also plays major roles in disease processes as evidenced by many recently published studies. Abnormal m^6A levels and aberrant expression of its regulatory proteins have been implicated in multiple disease processes particularly in different types of cancers. In this chapter, we compiled the literatures focusing on the involvement of m^6A modification in normal development as well as in pathogenesis of diseases. We also highlighted the emerging role of m^6A modification in diseases caused by environmental exposure to toxic chemical substances.

Keywords m^6A RNA modification · Epitranscriptomic gene regulation · Stemness and pluripotencey · Cancer · Environmental exposure

1 Introduction

The N6-methyladenosine (m^6A) is one of the most well-studied post-transcriptional RNA modifications among more than 160 RNA modifications that has been discovered so far (Boccaletto et al. 2018). The presence of m^6A has been reported in both coding and noncoding RNAs which include messenger RNAs (mRNAs), long noncoding RNAs (lncRNAs), micro RNAs (miRNAs), and small nuclear RNAs (snRNAs) (Alarcón et al. 2015a, b; Patil et al. 2016; Zhao et al. 2020a, b). m^6A methylation is the most prevalent internal modification in mammalian mRNA accounting for 0.1–0.4% of the total adenosine present in mRNAs. In mRNAs the m^6A modification most frequently occurs in the consensus sequence of RRACH (where R = A/G/U; R = A/G; A = m^6A; H = A/C/U) (Niu et al. 2013; Zhao et al. 2020a, b) and is enriched in 3′-UTRs, 5′-UTRs, and near the stop codons (Meyer et al. 2012). Presence of this modification mark regulates various aspects of RNA processing and metabolism ranging from alternative splicing, transport, and stability to subcellular localizations (Yang et al. 2018).

The presence of m^6A modification was first discovered in 1974 in a purified fraction of poly(A) RNA obtained from mammalian cells (Desrosiers et al. 1974). Despite the decades of its discovery, the significance of m^6A modification in cellular processes and its biological role came to light much later. With the discovery of m^6A modifying machineries (writers, erasers, and readers) and the development of methods for its precise mapping in the entire transcriptome, the role of m^6A modification in mRNA fate and function has now become widely accepted (Meyer et al. 2012;

Dominissini et al. 2012; Uddin et al. 2020). Since the pioneering works of Dominissini et al. and Meyer et al. to map the presence of m^{6}A modification in human transcripts, more than 500 mapping studies have become available in the public databases proving that m^{6}A is widely distributed in the entire human transcriptome (Zaccara et al. 2019). The studies also provided evidence of the relationship of m^{6}A modification with gene expression regulation related to developmental processes and disease-associated pathways (Roundtree and He 2016). In this chapter, we highlighted the role of m^{6}A methylation in epitranscriptomic regulation of normal developmental processes as well as how deregulation of this modification contributes to the pathogenesis of various human diseases. We also briefly discussed the effects of environmental pollutants on deregulation of m^{6}A modification as a cause of pathogenesis.

2 Regulators of m^{6}A Modification: Writers, Erasers, and Readers

The m^{6}A methylation is a dynamic and reversible process which is installed in the RNA molecules post-transcriptionally by methyltransferase enzymes (also called "writers") and removed by demethylase enzymes (also called "erasers"). The interplay between writers and erasers maintains the balance of m^{6}A level inside the cells. There is a third group of proteins that recognize the m^{6}A mark in the transcript and specifically bind to it called "readers" (Table 1). The recognition of the modification mark by the reader determines the fate of the transcript (Fig. 1) (Tong et al. 2018a, b; Shi et al. 2019; Zhao et al. 2020a, b).

The post-transcriptional modification of A to m^{6}A is introduced by a multicomponent complex of writers that site specifically recognize the modifiable target on the RNA molecule. In mammalian cells this complex consists of a catalytic core made of proteins named—methyltransferase-like 3 (METTL3), methyltransferase-like 14 (METTL14), and Wilms tumor 1-associated protein (WTAP) (Zhao et al. 2020a, b). In this complex, METTL3 is the main catalytic enzyme and METTL14 serves as an allosteric activator which also has binding capability with the target RNA. METTL3 catalyzes the transfer of a methyl group from *S*-adenosyl methionine (SAM) to adenosine on RNA transcripts by binding to SAM with a SAM-binding site present on its structure. METTL14 lacks SAM-binding capability but acts as an essential platform for binding of RNA molecules (Zhao et al. 2020a, b; Wang et al. 2016). METTL3 and METTL14 carry out most of the m^{6}A methylation in mRNA; loss of any of these components results in a reduction of more than 99% of m^{6}A site (Zaccara et al. 2019). Although individual methyltransferase activity of METTL3 and METTL14 is significantly low, as a complex they exhibit strong methyltransferase activity as they form a very stable complex (Yang et al. 2018; Liu et al. 2014). WTAP interacts with METTL3-METTL14 complex modulating their RNA binding and m^{6}A methylation capability although it lacks the catalytic potential (Liu et al. 2014). In contrast to METTL3 and

Table 1 m^6A RNA regulators—writers, erasers, and readers

Regulators	Effectors	Main functions
Writers	Methyltransferase-like 3 (METTL3)	m^6A catalysis
	Methyltransferase-like 14 (METTL14)	m^6A catalysis
	Methyltransferase-like 16 (METTL16)	m^6A catalysis
	Wilms'- tumor 1-associated protein (WTAP)	Catalytic enhancement
	RNA binding motifs protein 15 (RBM15)	Facilitating RNA binding
	RNA binding motifs protein 15B (RBM15B)	Facilitating RNA binding
	Vir-like m^6A methyltransferase-associated (VIRMA)/KIAA1429	Facilitating RNA binding
	HAKAI	Facilitating RNA binding
	Zinc finger CCCH domain-containing protein 13 (ZC3H13)	Facilitating RNA binding
Erasers	Fat mass and obesity-associated protein (FTO)	Demethylation
	AlkB homologue 5 (ALKBH5)	Demethylation
Readers	YTH domain family proteins 1–3 (YTHDF1–3)	m^6A recognition and RNA fate determination
	YTH domain-containing protein 1–2 (YTHDC1–2)	m^6A recognition and mRNA translation
	Insulin-like growth factor 2 mRNA-binding protein 1–3 (IGF2BP1–3)	m^6A recognition and RNA fate determination
	Heterogeneous ribonucleoprotein C (hnRNPC)	m^6A recognition and mRNA splicing
	Heterogeneous ribonucleoprotein G (hnRNPG)	m^6A recognition and mRNA splicing
	Heterogeneous ribonucleoprotein A2B1 (hnRNPA2B1)	m^6A recognition and mRNA splicing
	Eukaryotic initiation factor 3 (eIF3)	Cap-independent translation

METTL14, it lacks methyl transferase domain in its structure, but it is required for the localization of the complex in the nuclear speckles, the major site for pre-mRNA processing (Yang et al. 2018; Ping et al. 2014). Recently few other regulatory proteins have been identified which interact with the core catalytic complex via binding with WTAP. RBM15 and its paralogue RBM15B were identified as part of methyltransferase complex that facilitates the recruitment of the complex near the consensus sequence sites of the RNA transcript resulting in the methylation of adenosine residing in the motif (Patil et al. 2016). Vir-like m^6A methyltransferase-associated protein (VIRMA)/KIAA1429 was identified as another member of m^6A methyltransferase complex interacting with WTAP. Presence of VIRMA/KIAA1429 has been shown to be required for methylation, knockdown of which resulted in ~4-fold decrease in m^6A level (Schwartz et al. 2014). ZC3H13 and HAKAI have been demonstrated as other important components of m^6A methyltransferase complex through proteomic analysis and biochemical validation. In the multicomponent m^6A methyltransferase complex VIRMA/KIAA1429 serves as a scaffold physically interacting with WTAP to hold WTAP, HAKAI, and

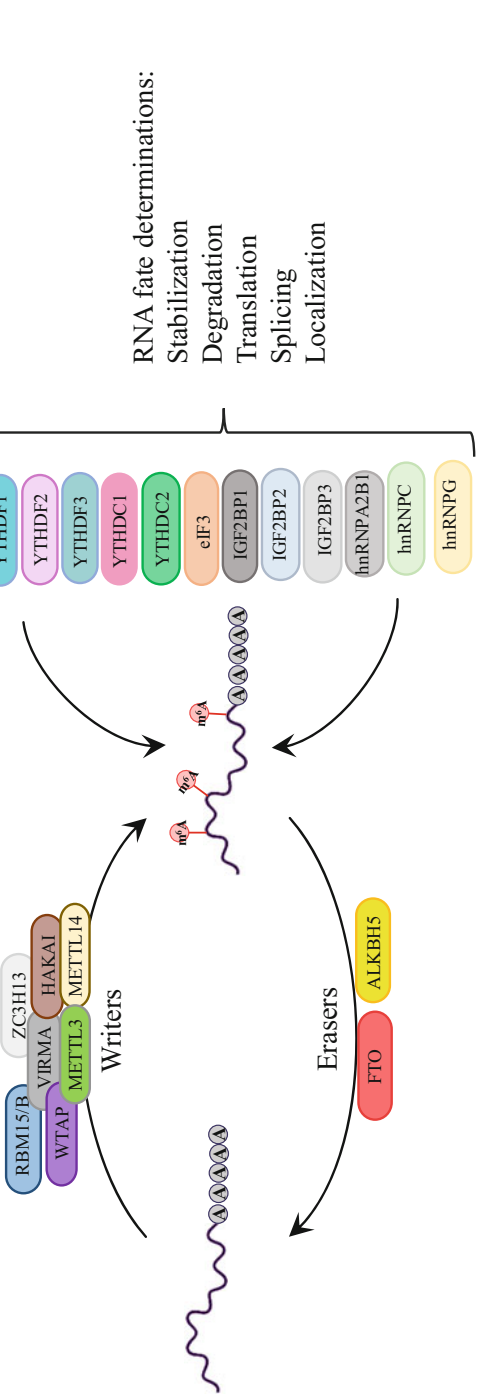


Fig. 1 m^{6}A RNA modification regulatory machinery—writers, erasers, and readers. The m^{6}A modification is installed in cellular RNA transcripts by the multicomponent complex of writer proteins containing methyltransferase-like 3 (METTL3), methyltransferase-like 14 (METTL14), Wilms tumor 1-associated protein (WTAP), Vir-like m^{6}A methyltransferase-associated protein (VIRMA), HAKAI, zinc finger CCCH domain-containing protein 13 (ZC3H13), and RNA-binding motif protein 15 (RBM15) and RBM15B. The modification is reversed by the erasers—fat mass and obesity-associated protein (FTO) and AlkB homologue 5 (ALKBH5). The m^{6}A modifications in the transcripts are recognized by the reader proteins—YTH domain family proteins 1–3 (YTHDF1-3), YTH domain-containing protein 1-2 (YTHDC1-2), eukaryotic initiation factor 3 (eIF3), insulin-like growth factor 2 mRNA-binding protein 1-3 (IGF2BP1-3), and heterogeneous ribonucleoprotein A2B1, C, and G (hnRNPA2B1, hnRNPC, and hnRNPG). Recognition of m^{6}A by readers contribute to RNA fate decisions

ZC3H13 together and creates a suitable binding pocket for METTL3/METTL14 to ensure optimal binding interaction (Wen et al. 2018; Yue et al. 2018). METTL16 has been identified as an independent m^6A methyltransferase which functions without forming complex with the other member of the writer proteins. METTL16 catalyzes m^6A modification mainly on U6 small nuclear RNAs (snRNAs) (Pendleton et al. 2017).

The m^6A modification is reverted back to adenosine by another group of proteins called m^6A demethylases or "erasers." Until now only two demethylating enzymes have been discovered named fat mass and obesity-associated protein (FTO) and α-ketoglutarate-dependent dioxygenase alkB homolog 5 (ALKBH5) (Jia et al. 2011; Zheng et al. 2013). FTO belongs to the Fe(II)- and 2OG-dependent dioxygenase enzyme family which has a distant homology with the ALKB family. FTO catalyzes the demethylation of m^3T and m^3U in single-stranded DNA as well as single-stranded RNA substrates. In the process of m^6A demethylation, FTO generates two intermediates—N6-hydroxymethyladenosine (hm^6A) and N6-formyladenosine (f^6A) the functions of which are unknown (Fu et al. 2013). FTO is localized mainly in the nuclear speckles along with the members of methyltransferase complex (Jia et al. 2011). ALKBH5 is another m^6A demethylase that belongs to ALKB family proteins. The demethylation effect of ALKBH5 is comparable to FTO, and it is also localized in the nuclear speckles playing a critical role in RNA metabolism, nuclear RNA export, and gene expression (Zheng et al. 2013).

The m^6A "readers" are the group of proteins that recognize the presence of m^6A modification in the cellular RNA transcripts and play vital roles in their fate determination. The YTH domain-containing proteins are the first identified reader proteins that directly interact with the RNA transcripts. Members of YTH domain-containing family YTHDF1, YTHDF2, and YTHDF3 were identified in a pull-down experiment where these proteins were separated along with m^6A using an m^6A-specific antibody (Dominissini et al. 2012). Other YTH domain-containing reader proteins were subsequently identified, namely YTHDC1 and YTHDC2 (Xu et al. 2014; Hsu et al. 2017). Although the YTH domain-containing proteins possess structural similarity, members of this family exhibit different subcellular localization and therefore diverse functionality. YTHDF2 decreases mRNA stability and enhances degradation, while YTHDF1 enhances mRNA stability and translation (Wang et al. 2015; 2014a, b). YTHDF3 on the other hand enhances the mRNA translation or degradation in a cellular context-dependent manner (Shi et al. 2017). YTHDC1 has been found to regulate mRNA splicing, and YTHDC2 was reported to enhance translational efficiency of its target transcripts (Xiao et al. 2016; Mao et al. 2019). Another group of reader proteins containing RNA-binding motifs (RBMs) in their structure were identified. The unique feature of these proteins is that the RNA-binding motifs are normally buried in their "inactive" state which are exposed upon m^6A recognition. The RBM-containing protein hnRNPC has been found to recognize the m^6A modification in the mRNAs and lncRNAs by structural alteration facilitating mRNA splicing and gene expression regulation (Liu et al. 2015). Other proteins of the same family hnRNPG and hnRNPA2B1 were found directly recognizing m^6A in mRNA and microRNA and contributing to alternative mRNA

splicing and microRNA processing (Alarcón et al. 2015a, b; Liu et al. 2017). Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs) were identified as novel reader proteins in recent years which contain two RNA recognition motifs (RRM) and four K homology (KH) domains in their structures. Three members of this group have been identified so far—IGF2BP1, IGF2BP2, and IGF2BP3 which affect RNA localization, translation, and stability (Degrauwe et al. 2016). Eukaryotic initiation factor 3 (eIF3), a component of the translation pre-initiation complex, is a distinct reader protein that can recognize the m^{6}A in the 5′-UTR of mRNAs and initiate translation in a cap-independent manner (Meyer et al. 2015).

3 m^{6}A RNA Modification in Developmental Processes

RNA m^{6}A modification is a dynamic process which not only increases diversity in RNA molecules but also poses added complexity to their functions. Some recent studies have shown that many of these modifications are involved in the process of post-transcriptional gene regulation. The location, distribution, and abundance of the modification depend on the type of cells, location of the organelle where the modification occurs, or even the type of RNA molecules undergoing modification (Yue et al. 2015). By controlling the gene expression in different organs or cell types, m^{6}A modification regulates the normal developmental processes in eukaryotes (Table 2).

3.1 Brain Development

m^{6}A methylation mediates the epitranscriptomic regulation of the lineage reprogramming of induced neuronal cells (iNs). During the reprogramming, elevated m^{6}A level was observed in the mRNA transcripts actively regulating the neuronal conversion. The increased m^{6}A modification in these transcripts is caused by METTL3 which was upregulated during the reprogramming of the cells. Knockdown of METTL3 resulted in failed reprogramming. Btg2, a powerful transcription factor and a potent inducer of neuronal differentiation, was found as the key target of METTL3. m^{6}A modification in Btg2 transcript by METTL3 was further recognized by YTHDF1 which increased its translation contributing to iNs differentiation (Choi et al. 2020). m^{6}A methyltransferase METTL14 is required for maintaining the self-renewal capability of neuronal stem cells (NSCs). Knockdown of METTL14 in NSCs showed a decrease in proliferative capability which was accompanied by premature differentiation in cortical regions of the mice brain. METTL14-knockdown mediated depletion of neuronal progenitor pool is brought about by alteration of expression of genes that regulate histone modification. Increased acetylation of histone H3 at lysine 27 (H3K27ac), trimethylation of histone H3 at lysine 4 (H3K4me3), and trimethylation of histone H3 at lysine 27 (H3K27me3)

Table 2 m^6A modification regulators in normal development

Effector organ/Tissue	Regulators	Target	Function	References
Brain development	↑ METTL3 YTHDF1	↑ Btg2	Lineage reprogramming of induced neuronal cells (iNs)	Choi et al. (2020)
	↑ METTL14	↓ Histone methyltransferase CBP ↓ Histone acetyltransferases p300	Proliferation and self-renewal of neuronal stem cells (NSCs)	Wang et al. (2018a, b)
	↑ METTL3 ↑ YTHDF1	Unspecified	Learning and memory formation	Shi et al. (2018)
	↓ METTL3	↑ Apoptosis-associated genes Dapk1 and Fadd Alternative splicing of Grin1	Developmental retardation in the hippocampus and cerebellum resulting in reduced motor activity and ataxia	Wang et al. (2018a, b)
	↑ METTL3 (early) ↑ ALKBH5 (late)	Unspecified	Development of the mouse cerebellum (early and late stage) during brain development	Ma et al. (2018)
	↑ METTL3 ↑ METTL14	↓ Neural stem cell maintenance genes ↑ Cell-cycle progression genes	Neurogenesis during cortical development	Yoon et al. (2017)
	↑ YTHDF2	Unspecified	Cortical neurogenesis	Li et al. (2018)
	↑ ALKBH5	Unspecified	Cerebellum and olfactory bulb development	Du et al. (2020)
Cardiac development	↑ METTL3	MAP3K6, MAP4K5, MAPK14	Cardiac hypertrophy	Dorn et al. (2019)
	↑ FTO	Unspecified	Autonomic regulation of cardiac function	Carnevali et al. (2014)
Liver development	↑ RBM15	Inhibition of mTORC1 pathway	Hepatocyte maturation	Hu et al. (2020)
Pancreatic β-cell maturation	↑ METTL14	Unspecified	β-cell differentiation, survival, and function	Liu et al. (2019a, b)
	↑ METTL3 ↑ METTL14	↑ MafA	β-cell maturation and functional	Wang et al. (2020)
Immune function	↑ METTL3	↓ IFNA and IFNB transcripts	Suppressed type I interferon-mediated antiviral response	Winkler et al. (2019),

(continued)

Table 2 (continued)

Effector organ/Tissue	Regulators	Target	Function	References
	↑ YTHDF2			Gao et al. (2020a, b)
	↓ METTL14	Unspecified	Impaired B-cell development and maturation	Zheng et al. (2020)
	↑ METTL3 ↑ METTL14	↓ Socs1, Socs3, and Cish IL-7/ JAK1/STAT5 pathway activation	T cell for differentiation and proliferation	Li et al. (2017a, b, c)
	↑ METTL3	↓ SOCS expression IL-2/STAT5 pathway activation	Increase of regulatory T cells (Tregs) function and stability	Tong et al. (2018a, b)
	↑ METTL3	↑ CD40, CD80, and Tirap	Promotes dendritic cell (DC) activation and function	Wang et al. (2019)
Stemness and pluripotency	↓ METTL3	Unspecified	Defective differentiation of naïve embryonic stem cells (ESCs), increased differentiation of primed pluripotent stem cells	Geula et al. (2015)
	↓ METTL3 ↓ METTL14	↓ Expression of pluripotency factors	ESC differentiation	Wang et al. (2014a, b)
	↑ METTL3	↑ Expression of pluripotency factors	Mouse embryonic fibroblasts (MEFs) reprogramming toward induced pluripotent stem cell (iPSC) formation	Chen et al. (2015)
	↑ METTL3	↑ JAK2, ↓ SOCS3	Maintaining pluripotency in porcine-induced pluripotent stem cells (piPSC)	Wu et al. (2019b)
	↑ FTO	↑ JAK2, ↑ p-STAT3	Adipogenesis	Wu et al. (2019a, b)
	↑ METTL3	↓ JAK1, ↓ p-STAT5	Decreased differentiation of bone marrow stem cells (BMSCs) to adipocytes	Yao et al. (2019)
	↓ METTL3	Unspecified	Increased differentiation of hematopoietic stem/progenitor cells (HSPCs), inhibition of cell growth	Vu et al. (2017)
	↓ METTL3	↓ MYC	Decreased differentiation of hematopoietic stem/progenitor cells (HSPCs)	Lee et al. (2019)
	↓ METTL3	↓ Parathyroid hormone receptor-1 (Pth1r)	Decreased osteogenic and increased adipogenic differentiation of mesenchymal stem cells (MSCs)	Wu et al. (2018)

(continued)

Table 2 (continued)

Effector organ/Tissue	Regulators	Target	Function	References
	↑ METTL3	↑ Runx2, ↑ Osterix	Increased bone mesenchymal stem cells (BMSCs) differentiation	Tian et al. (2019)
	↓ METTL3	↑ Atp6v0d2, nuclear retention of TRAF6	Reduced osteoclast differentiation	Li et al. (2020a, b)
Skeletal muscle development	↑ FTO	Activation of mTORC1 pathway	Increased myogenic differentiation	Wang et al. (2017)
	↑ METTL3	↑ MyoD	Myogenic differentiation to skeletal muscle	Kudou et al. (2017)

were observed in METTL14-KO NSCs. Increased histone modifications are due to increased stability of the mRNA transcripts of histone methyltransferase and acetyltransferases CBP and p300 by decreased m^6A methylation. METTL14 knockdown was found to inhibit the expression of neuritogenesis regulatory genes Kif26a, Gas7, and Pdgfrb by repressing H3K27ac and increase the expression of proliferation promoting genes Egr2 and Egr3 by repressing H3K27me3. m^6A modification thus regulates the proliferation and self-renewal ability of NSCs by direct regulation of histone modifications (Wang et al. 2018a, b).

m^6A modifications in mRNA transcripts of the genes regulating neuronal biological processes also contribute to learning and memory formation in mice. These functions are regulated by METTL3-mediated m^6A mRNA methylation and YTHDF1-mediated promotion of mRNA translation in the hippocampus. Knockdown of either METTL3 or YTHDF1 impaired learning process and memory formation (Shi et al. 2018). m^6A modification is also essential for brain development in the embryonic stage. Loss of m^6A due to METTL3 knockdown resulted in severe developmental retardation and hypoplasia of the cerebellum region in mice which is manifested by shrinking, reduced cerebellar weight, and foliation. Increased cerebellar hypoplasia resulted from increased apoptotic death of the cerebellar granular cells. METTL3 knockout (KO) decreased m^6A methylation in the apoptosis-associated genes Dapk1 and Fadd facilitating their stabilization and increased expression. Furthermore, m^6A depletion contributed to alternative splicing of Grin1 gene with increased NMDA splice variant which facilitates Ca^{2+} influx inside the neuronal cells. Both increase in expression of proapoptotic genes and intracellular Ca^{2+} contributed to apoptotic cell death of the newborn granular cells in the cerebellum. Cerebellar hypoplasia due to METTL3 KO exhibited features of cerebellar ataxia manifested by tremor and curled legs. Overall, METTL3-KO mice showed significant reduction in motor activity such as intermittent and sluggish movement, loss of speed and agility, as well as loss of body weight (Wang et al. 2018a, b). m^6A mRNA methylation also plays an essential role in postnatal development of mouse cerebellum by controlling developmental-related gene expression. The methylation pattern of these regulatory genes is specific to various

developmental stages. The global m^{6}A methylation decreases with increasing brain development which is controlled in the initial stages by METTL3 and later stages by ALKBH5. Defective cerebellar development was detected in the early stages of brain development by METTL3 knockdown and later stages by ALKBH5 knockdown. Therefore, proper regulation of m^{6}A modification is important in mouse brain development from postnatal stage to adulthood. ALKBH5 deficiency also causes deleterious effect on cerebellar development in hypoxic condition (Ma et al. 2018). METTL3 and METTL14 are also essential for maintaining the neurogenesis during cortical development in mice. Decreased m^{6}A due to deletion of METTL3 or METTL14 results in a delay in cell-cycle progression and exit in neural progenitor cells (NPCs) as well as reduced differentiation of radial glial cells (RGCs) during embryonic cortical neurogenesis. m^{6}A modification of the transcripts exhibits rapid turnovers during cortical neurogenesis, thus playing a critical role in developmental transitions. During the embryonic development process, the transcription factors related to neural stem cell maintenance are m^{6}A modified which accelerates their decay. Similarly, the cell-cycle regulation-related genes are also m^{6}A methylated which accelerates cell-cycle progression (Yoon et al. 2017). The m^{6}A reader YTHDF2 also plays key roles in cortical neurogenesis by recognizing the m^{6}A-modified transcripts and promoting their degradation by mRNA decay machinery. Deletion of YTHDF2 leads to failure in generation of functional neurons and other supporting cell types in the embryonic neocortex. Loss of YTHDF2 during embryonic cortical development compromises the neural development leading to embryonic lethality (Li et al. 2018). Widespread distribution of the m^{6}A demethylase ALKBH5 was observed particularly in the cerebellum and olfactory bulb in the mouse brain during embryonic stage which subsequently decreased gradually in the postnatal stage and adulthood. The distribution of ALKBH5 was observed mostly in the nuclear region of the neuron than any other cell type indicating potential role of ALKBH5 in brain functions (Du et al. 2020).

3.2 Cardiac Development

Nearly one-fourth of the transcripts in the healthy human heart contain m^{6}A modification implying its vital role in normal cardiac function (Berulava et al. 2020). METTL3 and m^{6}A modification plays a critical role in maintaining cardiomyocyte homeostasis, normal function, and response to stress. In normal cardiomyocytes METTL3 overexpression induces cardiac hypertrophy by inserting m^{6}A modification to specific mRNA transcripts that belong to protein kinase family including the MAPK proteins such as MAP3K6, MAP4K5, and MAPK14. METTL3 is also upregulated in response to hypertrophic stimuli and increases with development from neonatal state to adulthood indicating its involvement in cardiac homeostasis. METTL3 also changes the geometry of cardiomyocytes enhancing both their length and width. Knockdown of METTL3 on the other hand distorts the normal geometry of the adult myocytes although it does not affect the normal cardiac development.

METTL3 deleted myocytes show increased length-to-width ratio. Altered myocyte geometry causes enhanced left ventricular chamber dimension and ventricular dilation leading to progression toward cardiac failure. METTL3 and m^6A modification is therefore critical in maintaining normal cardiac function as well as cardiac remodeling in response to injury and stress (Dorn et al. 2019). m^6A modification regulation by FTO also plays an important role in autonomic regulation of cardiac function. FTO knockout mice showed increased heart rate and body temperature in resting as well as stress conditions. FTO-deficient mice were more vulnerable to stress-induced tachyarrythmias indicating modulation of sympathetic nerve activity by FTO knockdown. Alterations in ventricular repolarization and hypertrophy of both right and left ventricles due to proarrhythmic remodeling of myocardium were also observed in FTO-deficient mice (Carnevali et al. 2014).

3.3 Liver Development

The generation of liver is initiated in the endoderm from the hepatic precursors which are differentiated to hepatoblasts. Upon maturation, hepatoblasts give rise to functional hepatocytes. RBM15, an m^6A methylation reader which is encoded by the gene *cq96*, has been reported to play a vital role in hepatocyte maturation. Using zebrafish as a model organism it has been shown that RBM15 is specifically expressed in hepatocytes during differentiation. Mutation in the *cq95* results in maturation defect of the hepatocytes while hepatoblast specification, differentiation, and proliferation are unaffected. *cq95* mutation-induced RBM15 deficiency causes hyperactivation of mTORC1 signaling causing the hepatocyte maturation defects which is reversed partially by the inhibition of mTORC1 pathway. However, which genes are targeted by RBM15 in the process of hepatocyte maturation is still unknown (Hu et al. 2020).

3.4 Pancreatic β-Cell Maturation

The insulin secreting β-cell mass in pancreatic islets is maintained at an adequate level by differentiation, proliferation, and functional maturation. The process of functional β-cell number establishment is accomplished during the early postnatal stage. The immature β-cells generated during the neonatal period are unable to secrete insulin but gain maturity to acquire glucose sensitivity for insulin secretion in the subsequent stages of postnatal life (Wang et al. 2020; Liu et al. 2019a, b). β-cell deficiency due to defects in development and expansion results in impaired glucose metabolism leading to diabetes. METTL14 has been shown to play pivotal roles in β-cell differentiation, survival, and function by m^6A modification of the regulatory genes. METTL14 knockout mice showed embryonic lethality, glucose intolerance, and reduced insulin secretion upon glucose stimulation due to decreased

insulin expression in the β-cells. Knockout of METTL14 decreases the β-cell population in pancreatic islets due to increased apoptosis as well as defective β-cell differentiation resulting in lower insulin secretion. Lack of METTL14 also decreases the compensatory islet expansion by high fat diet (HFD) feeding. In addition, METTL14 knockout mice exhibited increased insulin sensitivity due to Akt activation resulting in decreased lipogenesis and gluconeogenesis and increased lipolysis in the liver. All these effects were due to METTL14 depletion-associated alterations in the gene expression regulation in the β-cells (Liu et al. 2019a, b). Increased expression of METTL3 and METTL14 was also reported in mice during early postnatal period which was maintained during the subsequent stages throughout the adulthood. Knockout of both METTL3 and METTL14 resulted in decreased postnatal β-cell function manifested by an increase in blood glucose level as well as reduced β-cell number by decreasing proliferation and increasing apoptosis. β-cells in these mice showed decreased insulin release due to impaired maturation of insulin secretory vesicles. Lack of β-cell maturation and functional impairment in METTL3 and METTL14 loss is attributed to downregulation of MafA, a potent regulator of β-cell maturation and function. In neonatal β-cells, METTL3 and MELL14 maintain the stability of MafA mRNA by direct m^{6}A modification. METTL3 and METTL14 knockdown decreases the MafA mRNA stability and expression leading to impaired β-cell functional maturation. m^{6}A modification carried out by METTL3 and METTL14 thus plays central roles in governing the β-cell development during neonatal period (Wang et al. 2020).

3.5 Immune Function

Production of type I interferons (IFN-α, IFN-β) plays a vital role in immunity against viral infection by producing a rapid and effective response. Regulation of the type I interferon-mediated immune response is essential to protect the host cells. m^{6}A modification has been reported as a novel regulatory mechanism modulating the type I interferon-mediated antiviral response. m^{6}A methylations of the IFNA and IFNB transcripts by m^{6}A writers are destabilized by the reader proteins restraining the extent of the response. Depletion of m^{6}A by knocking down METTL3 and the reader of m^{6}A methylation YTHDF2 has been found to induce the expression of interferon stimulating genes (ISGs). Particularly, IFNB which encodes for IFN-β, the central cytokine for type I interferon response, has been shown to be the target of METTL3-mediated m^{6}A modification. m^{6}A-modified IFNB transcript is further recognized by YTHDF2 which accelerates its decay. Depletion of METTL3 or YTHDF2 leads to elevation of IFNB mRNA stability and IFN-β expression resulting in restriction of viral proliferation (Winkler et al. 2019). Increase in the ISG expression and upregulation of antiviral immune response genes due to decreased m^{6}A modification was also observed in METTL3 knocked down fetal liver HSCs (Gao et al. 2020a, b).

m^6A modification is a key player in B-cell development and maturation. In developing B cells widespread m^6A modification was observed in the transcripts. Altered m^6A methylation due to loss of METTL14 shows impairment in B-cell development (Zheng et al. 2020). m^6A modification is also essential for T-cell homeostasis and differentiation. In normal naïve T cells IL-7 stimulation causes m^6A modification of immediate-early inducible genes Socs1, Socs3, and Cish causing their rapid degradation which activates the IL-7/JAK1/STAT5 signaling pathway. This activation initiates the reprogramming of naïve T cell for differentiation and proliferation. Silencing METTL3 or METTL14 results in an increased expression of Socs1, Socs3, and Cish and suppression of the IL-7/JAK1/STAT5 pathway. Knockdown of METTL3 in naïve T cells diminishes its proliferation as well as differentiation maintaining it in naïve state (Li et al. 2017a, b, c). m^6A modification also regulates the function of regulatory T cells (Tregs). Depletion of METTL3 in Tregs increases the SOCS protein expression by increasing their mRNA stability. Upregulation of SOCS proteins inhibits IL-2/STAT5-mediated Treg function and stability (Tong et al. 2018a, b). METTL3-mediated m^6A modification has been reported to promote dendritic cell (DC) activation and function. Elevated m^6A methylation in the transcripts of CD40, CD80, and Tirap in DCs is recognized by YTHDF1 which increases their translation. Upregulated CD40 and CD80 enhance the T-cell stimulation and antigen presentation, whereas increased Tirap expression activates the TLR4/NF-κB signaling and proinflammatory cytokine secretion. METTL3 knockdown on the other hand decreases the expression of CD40, CD80, and Tirap resulting in impaired DC maturation and activation and DC-mediated T-cell function (Wang et al. 2019).

3.6 Stemness and Pluripotency

The m^6A writer protein METTL3 plays a vital role in regulating pluripotency of the stem cells. Depletion of METTL3 in naïve embryonic stem cells (ESCs) resulted in defective differentiation of embryonic bodies and neuronal stem cell maturation. Thus, lack of METTL3 maintains the pluripotency in naïve ESCs. The opposite effect of METTL3 depletion was observed in primed pluripotent stem cells. In this case, METTL3 depletion diminished self-renewal capability and accelerated the stem cell differentiation (Geula et al. 2015). Depletion of METTL3 and METTL14 in mESCs and impaired pluripotency have also been reported in another study. METTL3 and METTL14 KD cells showed reduced stem cell features than the control cells due to decreases in the expression of pluripotency factors and increased expression of developmental regulatory genes. Loss of methyltransferases and m^6A modification in the developmental-associated genes increased their stability in a HuR-dependent manner, thereby moving ESCs toward differentiation (Wang et al. 2014a, b). m^6A methylation catalyzed by METTL3 and METTL14 has been found to be involved in embryonic stem cell formation in mouse (Yue et al. 2015). Overexpression of METTL3 in mouse embryonic fibroblasts (MEFs) increased

m^{6}A abundance in the transcripts of pluripotency factors upregulating their expression. Increased expression of the pluripotency factors in MEFs induced their reprogramming toward induced pluripotent stem cell (iPSC) formation (Chen et al. 2015). m^{6}A RNA modification and its regulators also play crucial roles in maintaining pluripotency in porcine-induced pluripotent stem cells (piPSC). piPSCs maintain pluripotent attributes through METTL3-regulated activation of JAK2-STAT3 signaling pathway. METTL3 mediates the m^{6}A modification in both JAK2 and the negative regulator of the JAK2-STAT3 pathway, SOCS3. m^{6}A modification decreases SOCS3 expression diminishing the repressive effect of SOCS3 on this pathway. In this case, m^{6}A-modified JAK2 mRNA is recognized and stabilized by YTHDF1, whereas m^{6}A-modified SOCS3 mRNA is degraded by YTHDF2 (Wu et al. 2019a). The same signaling pathway was reported to be associated with adipogenesis in an FTO-dependent manner where it is shown that FTO demethylates the JAK2 mRNA which activates STAT3 for adipocyte differentiation by C/EBPβ. In the absence of FTO, the m^{6}A-modified JAK2 mRNA is recognized and degraded by YTHDF2 (Wu et al. 2019a, b). In contrast to the previous findings, METTL3-inhibited differentiation of bone marrow stem cells (BMSCs) to adipocytes was found to be associated with inactivation of JAK1/STAT5/C/EBPβ pathway. METTL3 incorporates m^{6}A modification to JAK1 mRNA which is recognized by YTHDF2, leading to JAK1 degradation and inhibition of STAT5 phosphorylation. Failure of STAT5 phosphorylation inhibits STAT5-mediated activation of C/EBPβ mRNA transcription resulting in failure of adipocyte differentiation to adipocytes (Yao et al. 2019).

Altered m^{6}A modification due to alteration of methyltransferases or demethylases has been reported to alter the differentiation of hematopoietic stem/progenitor cells (HSPCs). Knockdown of METTL3 in HSPCs resulted in a decreased global m^{6}A level leading to increased differentiation of these progenitor cells. Increased differentiation indicated by loss of stem cell markers was accompanied with inhibition of cell growth (Vu et al. 2017). An opposite effect of METTL3 knockdown on hematopoietic stem cell (HSC) differentiation was reported in another study. In this study, it was shown that METTL3 deletion led to accumulation of HSC in the bone marrow due to blocked HSC differentiation. The failure of HSC differentiation was attributed to the lack of MYC expression regulated by METTL3 (Lee et al. 2019). m^{6}A modification also affects the osteogenic and adipogenic differentiation of mesenchymal stem cells (MSCs) in mice. Conditional knockout of METTL3 in mice resulted in a decreased translation of parathyroid hormone receptor-1 (Pth1r) causing interruption in parathyroid hormone (PTH)-mediated osteogenic differentiation and promotion of preferential adipogenic differentiation. Loss of METTL3 in MSCs thus results in severe bone loss and excessive adipose tissue accumulation in the bone marrow leading to osteoporosis (Wu et al. 2018). METTL3 has also been reported to be required for osteogenic differentiation of bone mesenchymal stem cells (BMSCs) in mice. During osteogenic differentiation, METTL3 is upregulated in BMSCs, knockdown of which results in the decreased expression of genes Runx2 and Osterix that are required for bone formation (Tian et al. 2019). Elevated m^{6}A modification due to increased METTL3 is also observed during osteoclast

differentiation. Knockdown of METTL3 results in increased stability of Atp6v0d2 gene that causes fusion of osteoclast precursor cells and regulates the osteoclast size. METTL3 depletion also causes nuclear retention of TRAF6, an adaptor protein required for MAPK, NFκB, and PI3K-AKT activation-induced osteoclast differentiation. These together lead to reduced osteoclast differentiation by METTL3 KO-associated decrease in m^6A levels (Li et al. 2020a, b).

3.7 Other Developmental Functions

m^6A modification is also reported to be involved in skeletal muscle cell differentiation mediated by FTO. During myoblast differentiation, FTO is upregulated with a subsequent decrease in m^6A levels. FTO contributes to myogenic differentiation by activating the downstream mTORC1 pathway which upregulates peroxisome proliferator-activated receptor gamma coactivator 1 alpha (PGC-1α), the master regulator of mitochondrial biogenesis. Silencing of FTO on the other hand suppresses the myogenic differentiation by decreasing the mitochondrial biogenesis and intracellular energy production. FTO-mediated m^6A demethylation of the target transcripts is therefore essential for myogenic differentiation (Wang et al. 2017). Expression of METTL3 is another important determinant of skeletal muscle development which is required for differentiation of myogenic stem/progenitor cells to mature skeletal muscle cells. METTL3 regulates the expression of transcription factor MyoD which is critical for myogenic stem cell differentiation. Knockdown of METTL3 decreases the mRNA processing of MyoD which downregulates its expression. Therefore, METTL3 is required for expression of MyoD in proliferating myoblasts (Kudou et al. 2017).

4 Aberrant m^6A RNA Modification in Human Diseases

The RNA post-transcriptional m^6A modification not only plays a role in normal cellular processes but is also implicated in the pathogenesis of different diseases. Abnormally high/low m^6A levels due to aberrant expression of its regulators have been observed in different types of diseases (Fig. 2; Table 3). The dysregulated m^6A modification machineries target specific RNA transcripts leading to their abnormal expressions, many of which are associated with disease initiation or progression.

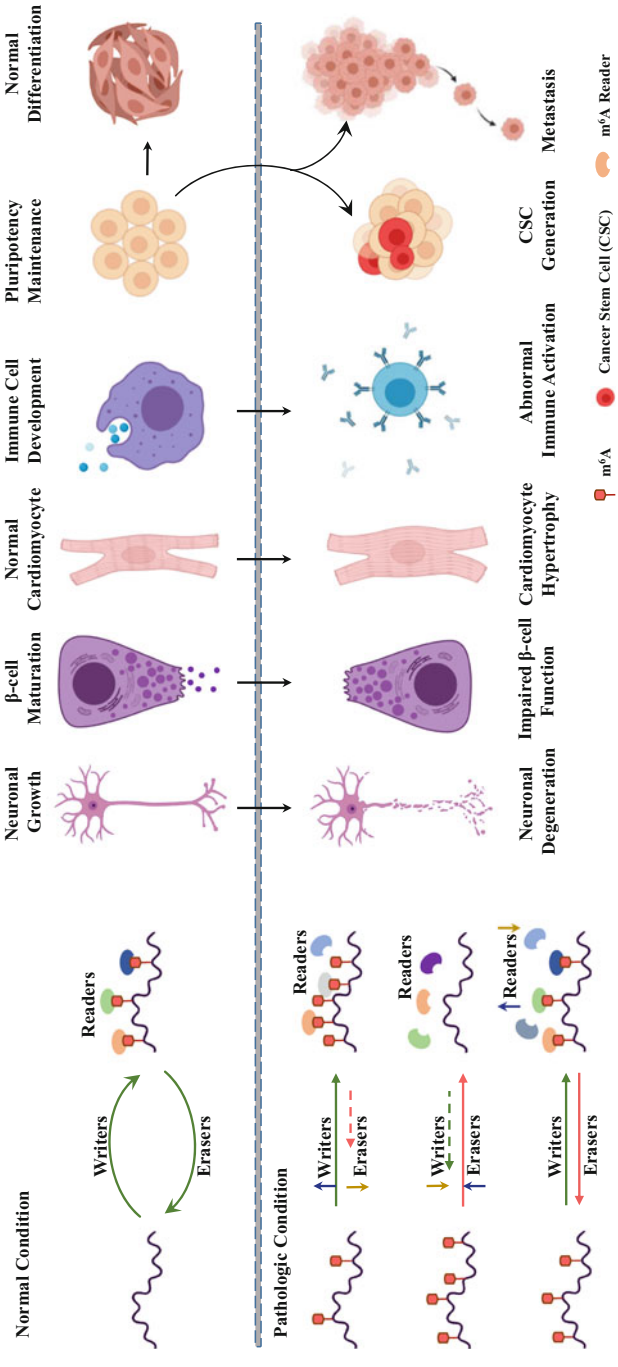


Fig. 2 RNA m^{6}A modification regulates normal developmental process and contributes to disease pathogenesis when deregulated. In normal healthy cells, m^{6}A modification is maintained at an adequate level which is essential for maintaining the pluripotency of the stem cells or differentiation to specific cell lineage as well as proper cell functioning. However, m^{6}A modification is deregulated in pathologic conditions giving rise to aberrantly expressed genes that are implicated in the pathogenesis of various diseases

Table 3 m^6A modification regulators in diseases

CNS Disorders				
Alzheimer's disease	↓ METTL3 ↑ RBM15B	Unspecified	Increased Tau protein in the hippocampus	Huang et al. (2020a, b, c)
Parkinson's disease (PD)	↑ FTO	↑ Grin1 (encodes for N-methyl-D-aspartate receptor 1 (NMDAR1))	Apoptotic death of dopaminergic neurons in substantia nigra	Chen et al. (2019)
Autosomal-recessive lethal syndrome	FTO-Mut (R316Q)	Unspecified	Severe postnatal growth retardation, microcephaly, impaired psychomotor activity, functional brain deficits, and facial deformities	Boissel et al. (2009)
	FTO-Mut (S319P)	Unspecified	Developmental defect	Daoud et al. (2016)
Major depressive disorder	↑ ALKBH5	Unspecified	Mental disorders	Du et al. (2015)
Cardiovascular diseases				
Ischemic heart disease	↓ FTO	↓ SERCA2A, MYH6–7, and RYR2 ↓ lncRNAs Chast and Mhrt	Decreased contractility of myocytes Increased cardiac fibrosis and hypertrophy	Mathiyalagan et al. (2019)
Heart failure	↓ FTO	Unspecified	Severe reduction in the ejection function, increased cardiac dilation	Berulava et al. (2020)
Ischemic heart disease	↑ METTL3	↓ TFEB	Autophagy inhibition-induced apoptotic cell death	Song et al. (2019)
Cardiomyopathy	↑ METTL3	↑ Arhgef3 ↓ Myl2	Cardiac hypertrophy and heart failure Loss of contractile function	Kmietczyk et al. (2019)
Atherosclerosis	↑ METTL14	↑ Forkhead box O1 (FOXO1)	Monocyte adhesion to the blood vessel endothelium to initiate the plaque formation	Jian et al. (2020)
Pulmonary disease				
Chronic obstructive pulmonary disease (COPD)	METTL3, FTO, YTHDC2, IGF2BP3	Unspecified	Development of COPD	(Huang et al.)
Metabolic disorders				
Liver fibrosis	↑ METTL3	↑ miR-350	PI3K/AKT and ERK signaling pathway-mediated liver fibrosis	(Zhu et al.)

(continued)

Table 3 (continued)

CNS Disorders				
Type 2 diabetes	↓ METTL3 ↓ METTL14	↓ p-Akt ↓ PDX1	Decreased β-cell mass due to suppressed proliferation and enhanced apoptosis	De Jesus et al. (2019)
	↓ METTL14	↓ p-Akt ↓ PDX1	Decreased proliferation and increased apoptosis of β-cells	Liu et al. (2019a, b)
	↓ METTL3 ↓ METTL14	Unspecified	Pathogenesis of diabetes	Wang et al. (2020)
Aberrant immune response	↓ METTL3	Unspecified	Enhanced inflammatory response and embryonic lethality	Gao et al. (2020a, b)
Cancer				
Breast cancer	↑ METTL14	↑ has-miR-146a-5p	Enhanced migration and invasiveness	Yi et al. (2020)
	↑ METTL14	↑ CXCR4 and CYP1B1	Breast cancer progression	Sun et al. (2020a, b, c)
	↑ METTL14	↑ DROSHA (RNase III) ↑ STC1	Increased breast cancer stem-like cells	Peng et al. (2020)
	↑ KIAA1429	↑ Cyclin-dependent kinase 1 (CDK1)	Increased breast cancer cell proliferation and metastasis	Qian et al. (2019)
	↓ METTL3	↑ Collagen type III alpha 1 chain (COL3A1)	Enhanced invasiveness, adhesion, and migration of triple-negative breast cancer (TNBC) cells	Shi et al. (2020)
	↑ FTO	↓ miR-181b-3p ↑ ARL5B	Increased invasiveness and migration of HER2-positive breast cancer cells	Xu et al. (2020a, b)
Colorectal cancer (CRC)	↑ METTL3	↑ CCNE1	CRC progression	Zhu et al. (2020a, b)
	↑ METTL3	↑ MYC	CRC progression	Xiang et al. (2020)
	↑ METTL3	↓ SOCS2	CRC cell proliferation	Xu et al. (2020a, b)
	↑ METTL3	Mut-p53 (R273H)	Chemoresistance	Uddin et al. (2019)
	↑ METTL3	↑ miR-1246 ↓ SPRED2	CRC metastasis	Peng et al. (2019)
	↓ METTL14	↑ SOX4	CRC progression	Chen et al. (2020a, b)
	↑ FTO ↑ ALKBH5	↑ Programmed cell death ligand (PD-1)	CRC progression	Tsuruta et al. (2020)

(continued)

Table 3 (continued)

CNS Disorders				
	↑ ALKBH5	↑ lncRNA NEAT1	Increased tumorigenesis	Guo et al. (2020)
Liver cancer	↑ METTL3	↑ CTNNB1	Hepatoblastoma development	Liu et al. (2019a, b)
	↑ KIAA1429	↑ GATA3	Generation of hepatocellular carcinoma (HCC)	Lan et al. (2019)
	↓ FTO ↓ ALKBH5	↑ LY6/PLAUR domain-containing 1 (LYPD1)	Hepatocarcinogenesis	Chen et al. (2020a, b)
	↑ YTHDF2	↑ OCT4	Increased stemness of liver cancer cells	Zhang et al. (2020a, b, c)
Lung cancer	↑ METTL3	↑ lncRNA ABHD11-AS1	Promotes proliferation and Warburg effect in NSCLC cells	Xue et al. (2020)
	↓ YTHDC2	Unspecified	Decreased differentiation, increased metastasis, and tumor progression in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC)	Sun et al. (2020a, b, c)
Ovarian cancer	↑ METTL3	EIF3C, AXL, CSF-1, and FZD10	Enhanced endometrial epithelial ovarian cancer (EEOC) cell proliferation and migration	Ma et al. (2020)
	↑ METTL3	↑ miR-126-5p	Promotes ovarian cancer cell proliferation, migration, invasion	Bi et al. (2020)
	↑ YTHDF1	↑ EIF3C	Increased tumorigenicity and metastasis of ovarian cancer cells	Liu et al. (2020a, b, c)
	↑ YTHDF1	↑ Tripartite motif protein 29 (TRIM29)	Increased cancer stem-like properties and chemoresistance	Hao et al. (2020)
	↑ YTHDF2	↓ miR-145	Increased proliferation and migration of epithelial ovarian cancer (EOC) cells	Li et al. (2020a, b)
	↓ FTO	↑ PDE1C and PDE4B	Decreased ovarian cancer cells and cancer stem cells (CSCs)	Huang et al. (2020a, b, c)

(continued)

Table 3 (continued)

CNS Disorders				
Gastric cancer	↑ METTL3	↑ miR-17-92 cluster	Increased gastric cancer cell proliferation and tumor growth	Sun et al. (2020a, b, c)
	↑ METTL3	↑ MYC	Increased gastric cancer cell proliferation, migration, and invasion	(Yang et al.)
	↑ METTL3	↓ SOCS2	Increased gastric cancer cell proliferation	Jiang et al. (2020)
	↑ FTO	↑ MYC	Gastric cancer progression	Yang et al. (2020a, b)
	↑ YTHDF1	↑ Frizzled7 (FZD7)	Increased gastric cancer cell proliferation and aggressiveness	Pi et al. (2020)
Bladder cancer	↑ METTL3	↓ SETD7 and KLF4	Promotes bladder cancer cell proliferation and metastasis	Xie et al. (2020)
	↑ METTL3	↑ AF4/FMR2 family member 4 (AFF4)	Increased bladder cancer stem cells and tumorigenicity	Gao et al. (2020a, b)
	↑ METTL3	↑ AFF4, IKBKB, RELA, and MYC	Bladder cancer progression	Cheng et al. (2019)
	↑ METTL3	↑ ITGA6	Increased growth and progression of bladder cancer cells	Jin et al. (2019)
	↑ METTL3	↑ miR221/222	Increased bladder cancer cell proliferation	Han et al. (2019)
Renal cell carcinoma (RCC)	↓ METTL3	PI3K/Akt/mTOR pathway activation	Increased cancerous cell proliferation, migration, and invasiveness	Li et al. (2017a, b, c)
	↑ FTO	↑ PGC-1α	Growth suppression of clear cell renal cell carcinoma	Zhuang et al. (2019)
Acute myeloid leukemia (AML)	↑ METTL3	↑ AML-associated genes	AML progression	(Barbieri et al.)
	↑ WTAP	↓ MYC	AML cell proliferation, tumorigenesis, and resistance to chemotherapy	Naren et al. (2020)
	↑ METTL14	↑ MYB and MYC	Promotes leukemogenesis	Weng et al. (2018)
	↑ FTO ↑ ALKBH5	↑ ATRA ↓ ASB2 and RARA	Increased AML cell differentiation and tumorigenesis	Li et al. (2017a, b, c)

(continued)

Table 3 (continued)

CNS Disorders				
	↑ ALKBH5	↑ TACC3	Increased tumorigenesis	Shen et al. (2020)
	↑ YTHDF2	↓ Tumor necrosis factor receptor TNFRFSF2	Promotes initiation and propagation of cancerous cells	Paris et al. (2019)
Diffuse large B-cell lymphoma (DLBCL)	↑ WTAP	↑ HK2	Initiation and progression of DLBCL	Han et al. (2021)
	↑ METTL3	↑ Pigment epithelium-derived factor (PEDF)	Increased proliferation of DLBCL cells	Cheng et al. (2020a, b)
Glioblastoma	↑ METTL3	↑ SOX2	Glioma stem-like cell (GSC) generation GSC radioresistance, tumor progression	Visvanathan et al. (2018)
	↓ METTL3 ↓ METTL14	↑ ADAM19	Increased GSC growth, self-renewal, and tumorigenesis	Cui et al. (2017)
	↑ FTO	↑ MYC	Glioblastoma generation	Xiao et al. (2020)
	↑ ALKBH5	↑ FOXM1	Promoted GSC tumorigenesis	Zhang et al. (2017)
	↑ YTHDF2	↑ MYC and VEGF	Glioblastoma generation	Dixit et al. (2020)
Neuroblastoma (NB)	↑ m^6A	↓ MYCN	Decreased neuroblastoma progression	Cheng et al. (2020a, b)
Cervical cancer	↑ IGF2BP3	↑ KCNMB2-AS1	Cervical cancer progression	Zhang et al. (2020a, b, c)
Endometrial cancer	↑ ALKBH5	↑ IGF1R	Enhanced proliferation and invasiveness of cancer cells	Pu et al. (2020)
Oral squamous cell carcinoma (OSCC)	↑ METTL3	↑ c-Myc	OSCC cell proliferation, invasion, and migration	Zhao et al. (2020a, b)
	↑ METTL3	↑ BMI1	Enhanced tumor growth and metastasis	Liu et al. (2020a, b, c)
Nasopharyngeal carcinoma	↑ METTL3	↑ Tankyrase	Increased cancer cell motility	Liu et al. (2020a, b, c)
	↑ YTHDF2	↑ IGF1R	Radioresistance	He et al. (2020)
Environmental exposure-associated diseases				
Transformed bronchial epithelial cells	↑ METTL3, METTL14, WTAP ↓ FTO, ALKBH5	miRNAs	Arsenic (As)-induced transformed phenotype	Gu et al. (2018)

(continued)

Table 3 (continued)

CNS Disorders				
Oncogenic transformation of kidney, prostate, and bronchial epithelial cells	↑ METTL3	↑ CDCP1	Cadmium (Cd), 3-methylcholanthrene, and nickel (Ni)-induced carcinogenesis	Yang et al. (2019)
Alzheimer's disease	↑ m^{6}A	Transcripts regulating synaptic transmission and neuronal development	Cobalt (Co) exposure-mediated impaired brain function	Tang et al. (2020)
Dopaminergic transmission deficit	↑ FTO	Unspecified	Arsenic (As)-induced impairment of learning and memory formation	Bai et al. (2018)
Keratinocyte damage	↑ METTL14 ↑ WTAP	Unspecified	Arsenic (As)-induced oxidative stress	Zhao et al. (2019)
Developmental defects	↑ ALKBH5	Unspecified	Carbon black nanoparticle-mediated damage to cerebral cortex	Zhang et al. (2020a, b, c)

4.1 CNS Disorders

Alzheimer's Disease

Decreased METTL3 and increased RBM15B expression were observed in the hippocampus region of the brain in Alzheimer's disease patients. However, increased METTL3 level was found along with the neurotoxic Tau proteins in some particular regions of the hippocampus of these patients although no direct correlation between Tau protein expression and METTL3 expression could be established in that study (Huang et al. 2020a, b, c).

Parkinson's Disease (PD)

Decreased m^{6}A methylation is also associated with PD initiation via increased dopaminergic neuron apoptosis in the substantia nigra region of the brain. In these neurons, decreased m^{6}A methylation is due to increased FTO expression. FTO overexpression causes increased m^{6}A methylation in *Grin1* gene which encodes for N-methyl-D-aspartate receptor 1 (NMDAR1) and stabilizes it. Increased NMDAR1 receptor increases the Ca^{2+} influx inside the neurons leading to impaired mitochondrial function. FTO overexpression further increases the oxidative stress inside the cells. Both increased oxidative stress and decreased mitochondrial function activate the mitochondrial apoptotic pathway leading to the apoptotic death of dopaminergic neurons (Chen et al. 2019).

Other CNS Disorders

FTO is essential for normal development of cardiovascular and nervous systems in humans. A loss-of-function mutation R316Q within the FTO gene which inactivates the FTO enzymatic activity has been found associated with autosomal-recessive lethal syndrome characterized by severe postnatal growth retardation, microcephaly, impaired psychomotor activity, functional brain deficits, and facial deformities (Boissel et al. 2009). A similar developmental defect was observed in another individual carrying a homozygous S319P missense mutation in FTO confirming the role of FTO in normal developmental processes (Daoud et al. 2016). m^6A modification in genes regulating the behavior has been implicated in mental disorders. An association between the m^6A eraser ALKBH5 and development of major depressive disorder has been reported in one study (Du et al. 2015).

4.2 Heart Disease

Ischemia and Heart Failure

Dysregulated m^6A modification is an early indicator of heart failure. Changes in the m^6A methylation have been observed during heart failure progression which affects the translation of the modified transcripts (Mathiyalagan et al. 2019; Berulava et al. 2020). The m^6A demethylase FTO plays a protective role in cardiomyocytes during hypoxia and ischemic heart disease. FTO exerts its protective role on myocytes via selectively demethylating SERCA2A, MYH6–7, and RYR2 transcripts which regulate cardiac contractility by stabilizing them and increasing their translation. FTO also regulates transcripts related to sarcomere organization, myofibril assembly, and Ca^{2+} transport. FTO maintains low m^6A level in lncRNAs Chast and Mhrt, thereby regulating the fibrosis and hypertrophy. FTO level decreases immediately after ischemia with a subsequent increase in the m^6A level leading to a decreased contractility of myocytes. Decreased FTO expression and a transcriptome-wide increase in m^6A modification have been observed in the pre-infarct and infarct regions of the heart during chronic heart failure in humans (Mathiyalagan et al. 2019). FTO knockout mice also showed severe reduction in the ejection functionality and increased cardiac dilation leading to heart failure (Berulava et al. 2020). Increased expression of FTO during myocardial infarction has been shown to decrease m^6A level and improve cardiac dysfunction in mice (Mathiyalagan et al. 2019). Increased m^6A methylation in hypoxia/reoxygenation-treated cardiomyocytes is also reported in another study. Increased METTL3 expression following hypoxia/reoxygenation has been shown to increase the m^6A level which inhibits autophagy, thereby inducing apoptotic cell death. METTL3 inhibits autophagy by downregulation of TFEB which is a positive regulator of autophagy-related gene expression and lysosomal biogenesis. METTL3-mediated m^6A methylation of TFEB mRNA is recognized by the RNA-binding protein HNRNPD which decreases the TFEB mRNA stability, thereby decreasing its expression. The m^6A demethylation of TFAB mRNA shows the opposite effect. TFAB on the other hand

downregulates METTL3 expression by decreasing its mRNA stability which creates a negative feedback regulatory loop in cardiomyocytes. TFEB also upregulates ALKBH5 by increasing its transcriptional activity. This study shows an important role of m^{6}A modification in apoptotic death of myocytes in ischemic heart disease (Song et al. 2019). Increased m^{6}A methylation was also observed in human cardiomyopathy which is regulated mainly by METTL3. m^{6}A methylation has profound effect on myocyte growth; increased m^{6}A level decreases the cell size, while decreased level shows the opposite effect which is also correlated with METTL3 level. METTL3 affects the translational efficiency of target transcripts Arhgef3 and Myl2 during cardiomyopathy increasing Arhgef3 expression while decreasing Myl2. Increased Arhgef3 activates the mTORC1 leading to cardiac hypertrophy and heart failure, whereas decreased Myl2 leads to loss of contractile function (Kmietczyk et al. 2019).

Atherosclerosis

m^{6}A RNA methylation has been found to contribute in the development of atherosclerotic plaque following endothelial inflammation. In TNF-α-induced endothelial cell inflammation, the m^{6}A writer METTL14 is upregulated which binds to and increases the m^{6}A methylation of forkhead box O1 (FOXO1), a powerful transcription factor in atherosclerosis formation. m^{6}A methylated FOXO1 mRNA is recognized by YTHDF1 which increases its stability and translation. Increased FOXO1 in turn increases the expression of adhesion molecules VCAM-1 and ICAM-1 which mediate the adhesion of monocytes to the endothelial cells in the blood vessel wall to initiate the plaque formation. METTL14 thus mediates the TNF-α-induced inflammatory response by regulation of FOXO1 mRNA m^{6}A methylation and stability, knockdown of which decreases the development of atherosclerosis (Jian et al. 2020).

4.3 Lung Diseases

Chronic Obstructive Pulmonary Disease

Using bioinformatics analysis, Huang et al. demonstrated that several key m^{6}A methylation regulators (METTL3, FTO, YTHDC2, IGF2BP3) can directly interact with and regulate the expression of genes associated with the pathogenesis of chronic obstructive pulmonary disease (COPD) (Huang et al.).

4.4 Liver Diseases

Liver Fibrosis

Liver fibrosis is a chronic pathologic condition culminating in liver cirrhosis which is initiated by the activation of hepatic stellate cells (HSCs). m^{6}A methylation has been shown to play a role in the activated HSC-associated liver fibrosis.

Upregulation of the acid-sensitive ion channel 1a (ASIC1a) in HSCs induces expression of METTL3 which promotes m^6A methylation and processing of pri-miR-350. Mature miR-350 targets SPRY2 and activates PI3K/AKT and ERK signaling pathway promoting liver fibrosis (Zhu et al. 2020a, b).

4.5 Diabetes

In type 2 diabetes patients, several m^6A regulators (METTL3, METTL14, ALKBH5, YTHDF1, etc.) have been found downregulated resulting in decreased m^6A methylation of the genes involved in cell-cycle regulation, signaling pathways, insulin secretion, and development of the pancreas. With bioinformatics analysis and in vitro cell model as well as in vivo mouse models, it has been shown that the insulin/IGF1-Akt-PDX1 pathway is affected due to hypomethylation. Similar to type 2 diabetes patients, knockdown of METTL3 or METTL14 in EndoC-βH1 cells (used as a model of human β-cell) resulted in decreased Akt phosphorylation and PDX1 hypomethylation decreasing its expression. Knockdown of METTL14 in mice also exhibited reduced β-cell mass, suppressed proliferation, and enhanced apoptosis (De Jesus et al. 2019). METTL14 knockout mice also showed reduced Akt activation and PDX1 expression which was in agreement with the observations by Liu et al. Similar to the study of Liu et al., this study showed downregulation of genes associated with cell-cycle progression, upregulation of genes involved in autophagy, and negative regulation of Akt signaling pathway. m^6A deregulation thus significantly contributes to the pathogenesis of type 2 diabetes (De Jesus et al. 2019; Liu et al. 2019a, b). Decreased METTL3 and METTL14 expression in type 2 diabetes patients is also reported in another study further corroborating the role of m^6A methylation in diabetes pathogenesis (Wang et al. 2020).

4.6 Aberrant Immune Response

m^6A modification is essential for normal hematopoiesis in the fetal liver. Deletion of METTL3 and consequent reduction of m^6A modification in fetal liver hematopoietic stem cells (HSCs) resulted in bone marrow failure and embryonic lethality. Defects in proliferative capacity, lineage commitment, and maturity in HSCs were also observed in METTL3-deleted fetal liver. The deleterious effects of METTL3 knockdown on HSCs and resulting bone marrow failure and embryonic lethality are due to aberrant activation of the innate immune response. METTL3 deletion causes decreased m^6A modification in the mRNA transcripts leading to the formation of abnormal double-stranded RNAs (dsRNAs) endogenously. Detection of the dsRNA activates the protein kinase R (PKR)-eIF2α, melanoma-differentiation-associated gene 5 (MDA5), RNA: retinoic-acid-inducible gene I (RIG-I), and OAS-RNase L

pathways leading to inflammatory response and embryonic death (Gao et al. 2020a, b).

4.7 Cancer

Breast Cancer

Abnormal m^{6}A RNA methylation has been shown to play an oncogenic role in breast cancer progression. Increased METTL14 and m^{6}A level was observed in breast cancer cells which enhanced migration and invasive capabilities of cancerous cells by m^{6}A modification-mediated upregulation of microRNA has-miR-146a-5p (Yi et al. 2020). Increased METTL14 expression in breast cancer tissues derived from patients has also been reported in another study. In this study, an oncogenic lncRNA, lnc942, has been reported to recruit METTL14 causing m^{6}A modification to its downstream targets CXCR4 and CYP1B1, thus increasing their stability and expression triggering oncogenic effect (Sun et al. 2020a, b, c). METTL14-mediated increase of m^{6}A modification also contributes to increased breast cancer stem-like cell population. In breast cancer cells, the Aurora kinase A (AURKA) oncogene increases the stability of METTL14 by inhibiting its ubiquitination. Upregulated METTL14 causes the m^{6}A modification of the RNase III DROSHA mRNA which is recognized by IGF2BP2, thus stabilizing it. DROSHA further interacts with β-catenin to activate the transcription factor STC1 which contributes to cancer stem-like cell phenotype (Peng et al. 2020). The m^{6}A writer KIAA1429 is associated with increased breast cancer cell proliferation and metastasis. The oncogenic role of KIAA1429 is exerted by m^{6}A-mediated upregulation of the cyclin-dependent kinase 1 (CDK1) in breast cancer cells (Qian et al. 2019). m^{6}A modification has also been shown to negatively regulate breast cancer progression. Knockdown of METTL3 was found to enhance the invasiveness, adhesion, and migration potential of triple-negative breast cancer (TNBC) cells. The metastatic inhibitory effect of METTL3 is due to m^{6}A-mediated downregulation of Collagen type III alpha 1 chain (COL3A1), expression of which is increased due to METTL3 depletion (Shi et al. 2020). Increased FTO expression contributes to invasiveness and migration of HER2-positive breast cancer cells. FTO upregulates the oncogene ARL5B by inhibiting microRNA miR-181b-3p (Xu et al. 2020a, b).

Colon Cancer

Increased METTL3 expression can promote colorectal cancer (CRC) by increasing the m^{6}A modification in its target molecules as shown in several studies. It was reported that METTL3 contributes to CRC progression by increasing the m^{6}A modification at the 3′-UTR region of CCNE1 mRNA, thus stabilizing it (Zhu et al. 2020a, b). In another similar study, METTL3 has been shown to upregulate MYC expression by increasing m^{6}A modification which is recognized by IGF2BP1, thereby increasing its stability and CRC progression (Xiang et al. 2020). m^{6}A modification and decreased SOCS2 mRNA stability and expression by METTL3

have also been shown to contribute to CRC progression. Knockdown of METTL3 increased SOCS2 expression which downregulated the expression of leucine-rich repeat-containing G protein-coupled receptor5 (LGR5), thus inhibiting CRC cell proliferation (Xu et al. 2020a, b). METTL3 also contributes to increased colon cancer cell chemoresistance by increasing the m^6A methylation in the p53 pre-mRNA directing it toward mutant p53 (R273H) protein expression (Uddin et al. 2019). METTL3 overexpression increases CRC progression by a miRNA-mediated manner. METTL3-mediated m^6A modification promotes the maturation of pri-miR-1246 which suppresses the antioncogene SPRED2 and diminishes its suppressive effect on MAPK pathway, thereby promoting CRC metastasis (Peng et al. 2019). In contrast, decreased m^6A modification and CRC progression was also reported. Chen et al. observed a decreased expression of METTL14 in CRC due to lysine-specific demethylase5C (KDM5C)-mediated inhibition of METTL14 transcription. Absence of METTL14 abolished the m^6A methylation of SOX4 mRNA and elevated its expression which otherwise would be degraded by YTHDF2. Increased SOX4 expression promotes CRC progression through PI3K/Akt signaling (Cheng et al. 2020a, b). Increased expression of m^6A demethylase FTO and ALKBH5 and increased CRC progression were also reported. FTO contributed to carcinogenesis by demethylation of programmed cell death ligand (PD-1) mRNA which increases its expression (Tsuruta et al. 2020). ALKBH5 has also been found to increase colorectal tumorigenicity by increasing the expression of lncRNA NEAT1 (Guo et al. 2020).

Liver Cancer

Aberrant expressions of m^6A regulators contribute to liver cancer development by m^6A modification to the molecules contributing to carcinogenesis. Increased METTL3 expression has been reported in hepatoblastoma tumor tissues. METTL3 increases the m^6A methylation to its downstream target CTNNB1 and activates the Wnt/β-catenin signaling pathway and contributes to hepatoblastoma development (Liu et al. 2019a, b). Another m^6A regulator KIAA1429 was also found upregulated in hepatocellular carcinoma (HCC). KIAA1429 regulated hepatocarcinogenesis via m^6A-mediated upregulation of GATA3 expression (Lan et al. 2019). Consistently, increased m^6A methylation due to decreased demethylase FTO and ALKBH5 expression contributing to increased HCC has been reported in two independent studies (Mittenbühler et al. 2020; Chen et al. 2020a, b). ALKBH5 downregulation increased m^6A methylation in the LY6/PLAUR domain-containing 1 (LYPD1) transcript which was recognized and stabilized by IGF2BP1 inducing the oncogenicity of the liver cancer cells (Chen et al. 2020a, b). The m^6A reader YTHDF2 has also been correlated to the increased stemness of liver cancer cells which regulates the expression of pluripotency factor OCT4 (Zhang et al. 2020a, b, c).

Lung Cancer

m^6A modification is a critical regulator of lung carcinogenesis. Increased METTL3 has been observed in tissues of non-small cell lung cancer (NSCLC) patients and cultured cells. METTL3 enhances the NSCLC carcinogenesis by increasing the expression of lncRNA, ABHD11-AS1. Increased m^6A in the ABHD11-AS1

transcript enhances its stability and expression which promotes the proliferation and Warburg effect in NSCLC cells (Xue et al. 2020). In addition to NSCLC, reduced expression of the m^{6}A reader YTHDC2 has been also been reported in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) patients which is associated with poor differentiation, increased metastasis, and tumor progression leading to poor overall patient survival (Sun et al. 2020a, b, c).

Ovarian Cancer

In ovarian carcinogenesis and advancement, abnormal m^{6}A modification has been identified as a critical determinant regulating the expression of key oncogenic molecules. Expression of METTL3 and m^{6}A level has been upregulated in patients with endometrial epithelial ovarian cancer (EEOC). In EEOC cells, METTL3 overexpression enhances cell proliferation and migration as well as decrease apoptosis by regulating expression of EIF3C, AXL, CSF-1, and FZD10. Therefore, increased METTL3 in EOCC serves as an indicator of malignancy and poor patient survival (Ma et al. 2020). METTL3 also promotes ovarian cancer progression in miRNA regulation-dependent manner. METTL3 has been shown to enhance the m^{6}A modification in pri-miR-126-5p inducing its maturation, thereby increasing expression of miR-126-5p. Increased miR-126-5p inhibits expression of PTEN, thereby activating the PI3K/Akt/mTOR pathway in promoting ovarian cancer cell proliferation, migration, invasion, and decreasing apoptosis (Bi et al. 2020). The m^{6}A reader YTHDF1 has been reported to be upregulated in ovarian cancer patients which is associated with poor disease prognosis and survival. EIF3C, a subunit of the translation initiation factor EIF3, is the target of YTHDF1. The m^{6}A modification in the EIF3C mRNA is recognized by YTHDF1 which promotes its translation, thus facilitating the tumorigenicity and metastasis of ovarian cancer cells (Liu et al. 2020a, b, c). YTHDF1 has also been reported to increase the translation of m^{6}A-modified transcripts of tripartite motif protein 29 (TRIM29), which is associated with cancer development and progression. Increased TRIM29 is also associated with increased cancer stem-like properties and cisplatin resistance of the ovarian cancer cells (Hao et al. 2020). Another m^{6}A reader YTHDF2 was found significantly upregulated in epithelial ovarian cancer (EOC) which promotes the proliferation and migration of EOC cells. YTHDF2 promotes tumorigenesis by downregulation of miRNA, miR-145, which is a negative regulator of cell proliferation and migration (Li et al. 2020a, b). FTO expression is decreased in ovarian cancer cells and cancer stem cells (CSCs). Overexpression of FTO suppresses the tumorigenesis by reducing the m^{6}A modification in 3′-UTR of phosphodiesterases PDE1C and PDE4B transcripts reducing their stability. FTO also decreases the stemness features of ovarian cancer cells by activating cAMP-mediated signaling (Huang et al. 2020a, b, c).

Gastric Cancer

Changes in m^{6}A modification due to altered expression of its regulatory proteins are observed in gastric cancer cells. METTL3 upregulation and elevated m^{6}A methylation in gastric cancer patients have been shown to indicate increased malignancy and poor treatment outcome. METTL3 increased the gastric cancer cell proliferation and tumor growth by promoting the biogenesis of a cluster of miRNAs, miR-17-92,

through m^6A modification. The miR-17-92 cluster activated the Akt/mTOR signaling pathway by inhibiting PTEN (Sun et al. 2020a, b, c). Increased METTL3 due to overexpression of hepatitis B X-interacting protein (HBXIP) was also reported in another study. METTL3 increased gastric cancer cell proliferation, migration, and invasion by m^6A modification-mediated increase of MYC mRNA translation and thereby cancer progression (Yang et al. 2020a, b). Knockdown of METTL3 showed decreased cell proliferation in gastric cancer cells. Suppressor of cytokine signaling (SOCS) family proteins, particularly SOCS2, have been shown as the target of METTL3. Knockdown of METTL3 increased expression of SOCS2 which otherwise would be downregulated due to m^6A modification and accelerated decay (Jiang et al. 2020). In contrast to the above findings, the m^6A demethylase FTO was also found to upregulate and contribute to gastric cancer progression. FTO is upregulated by HDAC3-mediated degradation of FOXA2 resulting in FTO overexpression. The oncogenic role of FTO was due to upregulation of MYC expression by decreased m^6A modification in MYC mRNA, thus stabilizing it (Yang et al. 2020a, b). Elevated expression of YTHDF1 was reported in aggressive gastric tumors and associated with poor overall patient survival. YTHDF1 promoted the gastric cancer cell proliferation by m^6A-mediated enhancement of translation of frizzled 7 (FZD7) protein leading to Wnt/β-catenin signaling hyperactivation (Pi et al. 2020).

Urological Cancer

m^6A modification and its regulators are closely associated with urological carcinogenesis. In bladder cancer, METTL3 has been shown to play tumor promoting function in a recent study. METTL3 increased the m^6A methylation in the tumor suppressor proteins SETD7 and KLF4 which were recognized by YTHDF2 leading to their degradation. Combined effect of METTL3 and YTHDF2 thus promotes bladder cancer cell proliferation and metastasis (Xie et al. 2020). METTL3 has also been shown to promote self-renewal capability of bladder cancer stem cells (BCSCs). Increased METTL3 caused upregulation of the AF4/FMR2 family member 4 (AFF4) protein by m^6A methylation of its mRNA. Increased AFF4 further enhanced the transcriptional activation of SOX2 and MYC by binding to their promoter regions conferring the BCSCs self-renewal ability and tumorigenicity (Gao et al. 2020a, b). Increased expression of AFF4, two key regulators of NF-κB pathway—IKBKB and RELA—as well as MYC due to METTL3 overexpression was also reported in another study. METTL3-mediated m^6A modification increased bladder cancer progression through AFF4/NF-κB/MYC signaling axis (Cheng et al. 2019). Increased METTL3 expression also promotes the bladder cancer development and progression by upregulating ITGA6. m^6A modification in the 3′-UTR of ITGA6 mRNA is recognized by YTHDF1 and YTHDF3 promoting its translation. Increased ITGA6 accelerates the growth and progression of bladder cancer cells (Jin et al. 2019). METTL3-mediated m^6A modification also regulates the maturation of pri-miR221/222. Increased miR221/222 reduce the PTEN expression enhancing bladder cancer cell proliferation (Han et al. 2019). An opposite effect of METTL3 was observed in renal cell carcinoma (RCC). METTL3 downregulation played a protumorigenic role by increasing cancerous cell proliferation, migration, and

invasiveness by modulating PI3K/Akt/mTOR signaling pathway (Li et al. 2017a, b, c). In contrast, the m^{6}A demethylase FTO played an anti-tumorigenic role in clear cell renal cell carcinoma. FTO increased the expression of PGC-1α by decreasing m^{6}A methylation to its mRNA transcript and increased oxidative stress and ROS production in cancer cells, thus suppressing the tumor growth (Zhuang et al. 2019).

Hematologic Malignancies

Dysregulated m^{6}A modification has been implicated in different hematological malignancies. In a recent study, METTL3 has been indicated as an essential gene for acute myeloid leukemia (AML) growth. In this study, METTL3 has been shown to attach with the transcription start site of AML-associated genes through CAATT-box binding protein CEBPZ inducing m^{6}A modification within the coding region and thereby enhance their translation (Barbieri et al. 2017). Another m^{6}A writer WTAP has been shown to play an oncogenic role in AML. WTAP downregulated c-Myc expression by m^{6}A methylation of the MYC mRNA causing AML cell proliferation, tumorigenesis, and resistance to chemotherapeutic agent daunorubicin (Naren et al. 2020). METTL14 was also reported to promote leukemogenesis by inhibiting the differentiation of AML cells. The oncogenic role of METTL14 is carried out by m^{6}A modification of MYB and MYC mRNA transcripts (Weng et al. 2018). The m^{6}A demethylases FTO and ALKBH5 were also shown to play an oncogenic role in AML. FTO enhances the expression of oncogenes in AML cells and inhibits all-trans-retinoic acid (ATRA)-induced differentiation of cancerous cells through downregulation of ASB2 and RARA by m^{6}A demethylation (Li et al. 2017a, b, c). ALKBH5 on the other hand exerts its oncogenic effect by upregulation of an oncogene TACC3 in a post-transcriptional m^{6}A demethylation manner (Shen et al. 2020). The m^{6}A reader YTHDF2 has been shown to be overexpressed in AML which promotes the initiation and propagation of cancerous cells. YTHDF2 decreases the half-life of m^{6}A-modified transcripts of tumor necrosis factor receptor TNFRFSF2 (Paris et al. 2019). WTAP upregulation is associated with the initiation and progression of diffuse large B-cell lymphoma (DLBCL). High WTAP expression is regulated by PIWI-interacting RNA-30473 (piRNA-30473) which is highly expressed in DLBCL. WTAP increases the m^{6}A level of its target HK2 mRNA and increases its expression which further increases the DLBCL progression (Han et al. 2021). Increased METTL3 expression was also observed in DLBCL cells. METTL3 upregulation causes DLBCL progression by enhancing the m^{6}A methylation of pigment epithelium-derived factor (PEDF) mRNA and increasing its expression. PEDF overexpression in turn increases the proliferation of cancer cells (Cheng et al. 2020a, b).

CNS Malignancies

The m^{6}A modification is highly abundant in human brain tissue playing pivotal roles in brain development. Dysregulation of m^{6}A modification may lead to malignant transformation giving rise to different cancers types. METTL3 is overexpressed in tumor initiating glioma stem-like cells (GSCs) leading to development of glioblastoma. In GSCs, METTL3 overexpression increased m^{6}A modification to the SOX2

mRNA transcripts and stabilizes it through human antigen R (HuR) binding. Increased SOX2 expression renders GSCs resistance to γ-irradiation and enhances DNA repair capacity leading to glioblastoma development (Visvanathan et al. 2018). However, an opposite effect of METTL3 or METTL14 in GSCs has been reported in another study. In this study, a decreased expression of METTL3 or METTL14 has been shown to correlate with increased GSC growth, self-renewal, and tumorigenesis. Consistently, an increased FTO and decreased m^6A modification were associated with tumor progression. Decreased methyltransferases or increased demethylases caused an alteration in the m^6A abundance and expression of oncogenic transcripts such as ADAM19 promoting glioblastoma progression (Cui et al. 2017). FTO also increased tumorigenesis in glioma cells by targeting MYC expression. Enhanced MYC expression suppressed MXI1 by increasing expression of miR-155 and miR-23a cluster leading to tumor generation (Xiao et al. 2020). m^6A demethylase ALKBH5 was also found highly expressed in GSCs. ALKBH5 demethylated the transcription factor FOXM1 increasing its expression and promoted GSC tumorigenesis (Zhang et al. 2017). The m^6A reader YTHDF2 has been found to preferentially express in GSCs over normal neuronal stem cells (NSCs). YTHDF2 promoted glioblastoma generation by increasing the stability of MYC and VEGF transcripts in an m^6A-dependent manner (Dixit et al. 2020). Increased m^6A modification has been shown to play a role in decreasing neuroblastoma (NB) progression. m^6A modification in the 3′-UTR of the protooncogene MYCN decreased its expression via promoting its binding with tumor suppressive microRNA, miR-98 (Cheng et al. 2020a, b).

Other Cancer Types

Abnormal expression of m^6A methylation and its regulatory proteins is also reported in other types of cancers. The m^6A reader IGF2BP3 and the lncRNA KCNMB2-AS1 are upregulated in the cervical cancer which promote cancer growth by mutual upregulation. KCNMB2-AS1 upregulates IGF2BP3 in a miRNA-dependent manner while IGF2BP3 upregulates KCNMB2-AS1 by binding to the m^6A methylation sites on it. These together form a positive regulatory feedback loop to enhance cervical cancer progression (Zhang et al. 2020a, b, c). In endometrial cancer, ALKBH5 has been found to be upregulated which enhances the proliferation and invasiveness of cancer cells. ALKBH5 demethylates IGF1R mRNA which enhances its stability and translation, thereby activating IGF-mediated signaling, and promotes tumorigenicity (Pu et al. 2020). METTL3 upregulation in oral squamous cell carcinoma (OSCC) has been linked to poor cancer prognosis. METTL3 promotes OSCC cell proliferation, invasion, and migration through upregulation of c-Myc. m^6A modification of the 3′-UTR of c-Myc transcript increases its stability mediated by YTHDF1 resulting in OSCC tumorigenesis (Zhao et al. 2020a, b). Increased METTL3 expression with poor prognostic signature of OSCC was also reported in two patient cohorts in another study. METTL3-mediated tumor growth and metastasis was due to increased m^6A modification in the 3′-UTR of BMI1 mRNA which increased its translation with the cooperation of m^6A reader IGF2BP1 (Liu et al. 2020a, b, c). METTL3 has also been found to play an oncogenic role in nasopharyngeal

carcinoma by m^{6}A methylation and enhancement of Tankyrase level resulting in increased cancer cell motility (Liu et al. 2020a, b, c). YTHDF2 confers the nasopharyngeal carcinoma cells resistance to radiotherapy by direct binding to m^{6}A-modified IGF1R mRNA promoting its translation. Increased IGF1R activates the IGF-mediated Akt/S6 signaling pathway contributing to radioresistance (He et al. 2020).

5 m^{6}A RNA Modification in Environmental Exposure-Associated Diseases

Increased m^{6}A modification has been shown to play roles in chemical exposure-induced oncogenic cellular transformation. Increased m^{6}A modification was observed in human bronchial epithelial cells upon chronic exposure to arsenic (As) which caused malignant transformation of the cells. As exposure increased the expression of methyltransferases METTL3, METTL14, and WTAP while downregulating the demethylases FTO and ALKBH5. Abnormal expressions of these m^{6}A regulatory proteins caused malignant transformation by targeting several microRNA regulatory pathways as shown by the bioinformatics analysis (Gu et al. 2018). Epithelial cells from the human kidney, prostate, and bronchi were exposed to chemical carcinogens like cadmium (Cd), 3-methylcholanthrene, and nickel (Ni) which caused upregulation of METTL3 and increased m^{6}A-methylated transcripts regulating RNA processing, metabolism, translation, and cell proliferation. Among other mRNA targets, METTL3 increased m^{6}A modification in the mRNA of CDCP1, a transmembrane glycoprotein playing oncogenic role. m^{6}A-modified CDCP1 mRNA was recognized by YTHDF1 which induced its translational upregulation. This shows the contribution of METTL3 and m^{6}A in environmental exposure-associated chemical carcinogenesis (Yang et al. 2019). Environmental exposure to cobalt (Co) causes impaired brain function including learning and memory by altering the m^{6}A modification in transcripts regulating synaptic transmission and neuronal development. Mice exposed to CoCl2 resulted in differentially m^{6}A-modified mRNA transcripts in the brain due to changes in expression of m^{6}A methyltransferase and demethylase proteins. Co exposure-mediated m^{6}A methylation also upregulated the neurodegenerative disease (e.g., Alzheimer's Disease)-associated gene expression (Tang et al. 2020). Chronic exposure of mice to arsenic (As) in drinking water had a similar effect as Co exposure causing impaired spatial learning and memory formation. In addition, the As-exposed mice showed deficit in condition avoidance and escape response as well as increased anxiety-like behavior. All these neurodegenerative disorders caused by As exposure are due to altered m^{6}A modification-associated dopaminergic transmission deficit and decreased dopamine content in the brain. Increased m^{6}A level in the dopaminergic neurons due to As exposure-induced downregulation of FTO contributes to the defective dopaminergic transmission (Bai et al. 2018). As exposure also enhances m^{6}A level by causing

oxidative stress in human keratinocytes. As induces the transcriptional repression of m^6A methyltransferase and demethylase regulatory genes HECTD4, ABCA5, SLC22A17, and KCNQ5, thus increasing the expression of METTL14 and WTAP and m^6A modification in the keratinocytes (Zhao et al. 2019). Mice exposed to carbon black nanoparticles during pregnancy were found to exhibit disoriented maternal behavior to the offspring. The offspring also exhibited developmental defects and behavioral alterations which are caused mainly by altered m^6A modification in the related genes of their cerebral cortex. The overall m^6A modification levels in the cerebral cortex tissue decreased sharply in the early stages of their life due to enhanced expression of ALKBH5 and downregulation of METTL3 and WTAP (Zhang et al. 2020a, b, c).

6 Conclusion and Future Perspectives

The process of cellular differentiation and development largely depends on the expression pattern of developmental regulatory genes. In addition to epigenetic regulation of gene expression by DNA and histone modification, RNA m^6A modification has opened a new avenue to understand the complex process of gene expression regulation in eukaryotic cells. In diverse biological processes of eukaryotic cells, correct deposition of m^6A modification in mRNA transcripts of critical genes is essential. For example, the differentiation of embryonic stem cells to a specific cell lineage requires the expression of specific genes to exit from pluripotent stage to differentiated stage. Post-transcriptional m^6A modification plays a critical regulatory role in the expression of these genes by determining their pattern of splicing, rate of translation, or stability (Geula et al. 2015; Wang et al. 2014a, b; Yue et al. 2015). In addition, m^6A modification also has an important regulatory role in the maintenance of proper functioning of adult differentiated cells such as learning and memory functions by neurons, maintaining cardiac contractility of cardiomyocytes, insulin secretory capacity of pancreatic β-cells, or proper immune functions by the cells of innate or adaptive immunity. The key players of these m^6A modification-mediated functions are its regulatory proteins—writers, erasers, and readers, expression patterns of which shape the m^6A landscape inside the cells.

The m^6A modification not only regulates the normal developmental process, but dysregulation of it is also implicated in the pathogenesis of various diseases. Recent studies have established the RNA m^6A modification as a major regulator of disease processes. Abnormal expression of m^6A and its regulators has been observed in diseases like-heart failure, diabetes, CNS disorders, and different types of malignancies (Table 3). The underlying mechanisms of the association of m^6A and its regulators are gradually being discovered. The main features of the regulations of developmental and disease processes by m^6A are all related to the extent of modifications in the transcripts of key molecules that are associated with the process. But many questions remain to be answered regarding the regulations and functions of m^6A modification regulation itself. The same m^6A regulatory molecule is acting as a

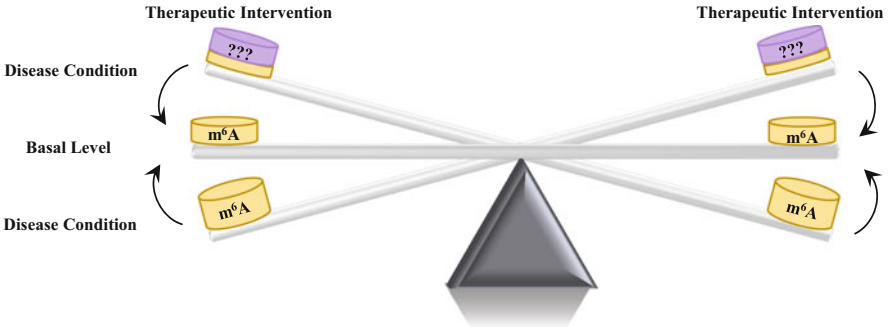


Fig. 3 m^{6}A modification may serve as a potential therapeutic target. The homeostatic balance of m^{6}A is perturbed in disease conditions. Reinstating the balance by targeting the m^{6}A modification regulators can lead to development of novel therapeutic agents for different diseases

player for developmental process as well as disease generation. Which cellular context is transforming this key player of development to a disease initiator needs further insight. The functions of m^{6}A regulators may be influenced under different environmental cues activating or repressing multiple cellular signaling pathways. Which factor determines the regulation of a signaling process targeted by m^{6}A modifiers is a matter to contemplate. The m^{6}A RNA modification has also been shown to play roles in abnormalities associated with environmental exposure to toxic chemical substances in several recent studies. But research in the environmental exposure-induced disease pathogenesis and the m^{6}A regulation is still in its early stages (Yang 2020). Future research work in the mechanistic insights of m^{6}A-mediated chemical exposure-induced disease processes is warranted.

Implications of m^{6}A and aberrant expression of its regulators in the pathogenesis of diseases show promises as well. As the m^{6}A modification and its regulators are abnormally expressed in various disease processes, targeting these regulatory pathways can be promising approaches in therapeutic interventions in these pathologic conditions. Similarly, since the m^{6}A regulators are involved in the pathogenesis of different types of cancers and aberrant expression of some of them is signature of particular cancer types, targeting them with specific inhibitors may facilitate in developing novel therapeutics against cancer (Fig. 3).

Acknowledgment This work was supported by the National Institutes of Health (R01ES026151; R01ES028256; R01ES029496; R01ES029942).

References

Alarcón CR, Goodarzi H, Lee H et al (2015a) HNRNPA2B1 is a mediator of m^{6}A-dependent nuclear RNA processing events. Cell 162:1299–1308

Alarcón CR, Lee H, Goodarzi H et al (2015b) N6-methyladenosine marks primary microRNAs for processing. Nature 519:482–485

Bai LL, Tang Q, Zou Z et al (2018) m^6A demethylase FTO regulates dopaminergic neurotransmission deficits caused by arsenite. Toxicol Sci 165:431–446
Barbieri I, Tzelepis K, Pandolfini L et al (2017) Promoter-bound METTL3 maintains myeloid leukaemia via m^6A-dependent translation control. Nature 552(7683):126–131
Berulava T, Buchholz E, Elerdashvili V et al (2020) Changes in m^6A RNA methylation contribute to heart failure progression by modulating translation. Eur J Heart Fail 22:54–66
Bi X, Lv X, Liu D et al (2020) METTL3-mediated maturation of miR-126-5p promotes ovarian cancer progression via PTEN-mediated PI3K/Akt/mTOR pathway. Cancer Gene Ther 1–15
Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307
Boissel S, Reish O, Proulx K et al (2009) Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. Am J Hum Genet 85:106–111
Carnevali L, Graiani G, Rossi S et al (2014) Signs of cardiac autonomic imbalance and proarrhythmic remodeling in FTO deficient mice. PLoS One 9:e95499
Chen T, Hao Y-J, Zhang Y et al (2015) m^6A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. Cell Stem Cell 16:289–301
Chen X, Yu C, Guo M et al (2019) Down-regulation of m^6A mRNA methylation is involved in dopaminergic neuronal death. ACS Chem Neurosci 10:2355–2363
Chen X, Xu M, Xu X et al (2020a) METTL14-mediated N6-methyladenosine modification of SOX4 mRNA inhibits tumor metastasis in colorectal cancer. Mol Cancer 19:1–16
Chen Y, Zhao Y, Chen J et al (2020b) ALKBH5 suppresses malignancy of hepatocellular carcinoma via m^6A-guided epigenetic inhibition of LYPD1. Mol Cancer 19:1–19
Cheng M, Sheng L, Gao Q et al (2019) The m^6A methyltransferase METTL3 promotes bladder cancer progression via AFF4/NF-κB/MYC signaling network. Oncogene 38:3667–3680
Cheng Y, Fu Y, Wang Y et al (2020a) The m^6A methyltransferase METTL3 is functionally implicated in DLBCL development by regulating m^6A modification in PEDF. Front Genet 11
Cheng J, Xu L, Deng L et al (2020b) RNA N6-methyladenosine modification is required for miR-98/MYCN axis-mediated inhibition of neuroblastoma progression. Sci Rep 10:1–9
Choi H, Baek S, Cho BY et al (2020) Epitranscriptomic N6-methyladenosine modification is required for direct lineage reprogramming into neurons. ACS Chem Biol 15:2087–2097
Cui Q, Shi H, Ye P et al (2017) m^6A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. Cell Rep 18:2622–2634
Daoud H, Zhang D, McMurray F et al (2016) Identification of a pathogenic FTO mutation by next-generation sequencing in a newborn with growth retardation and developmental delay. J Med Genet 53:200–207
De Jesus DF, Zhang Z, Kahraman S et al (2019) m^6A mRNA methylation regulates human β-cell biology in physiological states and in type 2 diabetes. Nat Metab 1:765–774
Degrauwe N, Suvà M-L, Janiszewska M et al (2016) IMPs: an RNA-binding protein family that provides a link between stem cell maintenance in normal development and cancer. Genes Dev 30:2459–2474
Desrosiers R, Friderici K, Rottman F (1974) Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc Natl Acad Sci USA 71:3971–3975
Dixit D, Prager BC, Gimple RC et al (2020) The RNA m^6A reader YTHDF2 maintains oncogene expression and is a targetable dependency in glioblastoma stem cells. Cancer Discov 11 (2):480–499
Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. Nature 485:201–206
Dorn LE, Lasman L, Chen J et al (2019) The N6-methyladenosine mRNA methylase METTL3 controls cardiac homeostasis and hypertrophy. Circulation 139:533–545
Du T, Rao S, Wu L et al (2015) An association study of the m^6A genes with major depressive disorder in Chinese Han population. J Affect Disord 183:279–286

Du T, Li G, Yang J et al (2020) RNA demethylase Alkbh5 is widely expressed in neurons and decreased during brain development. Brain Res Bull 163:150–159

Fu Y, Jia G, Pang X et al (2013) FTO-mediated formation of N6-hydroxymethyladenosine and N6-formyladenosine in mammalian RNA. Nat Commun 4:1–8

Gao Y, Vasic R, Song Y et al (2020a) m^{6}A modification prevents formation of endogenous double-stranded RNAs and deleterious innate immune responses during hematopoietic development. Immunity 52(6):1007–1021.e8

Gao Q, Zheng J, Ni Z et al (2020b) The m^{6}A methylation-regulated AFF4 promotes self-renewal of bladder cancer stem cells. Stem Cells Int 2020:1–12

Geula S, Moshitch-Moshkovitz S, Dominissini D et al (2015) m^{6}A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science 347:1002–1006

Gu S, Sun D, Dai H et al (2018) N6-methyladenosine mediates the cellular proliferation and apoptosis via microRNAs in arsenite-transformed cells. Toxicol Lett 292:1–11

Guo T, Liu D-F, Peng S-H et al (2020) ALKBH5 promotes colon cancer progression by decreasing methylation of the lncRNA NEAT1. Am J Transl Res 12:4542

Han J, Wang J-Z, Yang X et al (2019) METTL3 promote tumor proliferation of bladder cancer by accelerating pri-miR221/222 maturation in m^{6}A-dependent manner. Mol Cancer 18:110

Han H, Fan G, Song S et al (2021) piRNA-30473 contributes to tumorigenesis and poor prognosis by regulating m^{6}A RNA methylation in DLBCL. Blood 137(12):1603–1614

Hao L, Wang J-M, Liu B-Q et al (2020) m^{6}A-YTHDF1-mediated TRIM29 upregulation facilitates the stem cell-like phenotype of cisplatin-resistant ovarian cancer cells. Biochim Biophys Acta, Mol Cell Res 1868(2021):118878

He J-J, Li Z, Rong Z-X et al (2020) m^{6}A reader YTHDC2 promotes radiotherapy resistance of nasopharyngeal carcinoma via activating IGF1R/AKT/S6 signaling axis. Front Oncol 10:1166

Hsu PJ, Zhu Y, Ma H et al (2017) Ythdc2 is an N6-methyladenosine binding protein that regulates mammalian spermatogenesis. Cell Res 27:1115–1127

Hu L, Li H, Chi Z et al (2020) Loss of the RNA-binding protein Rbm15 disrupts liver maturation in zebrafish. J Biol Chem 295(33):11466–11472

Huang H, Camats-Perna J, Medeiros R et al (2020a) Altered expression of the m^{6}A methyltransferase METTL3 in Alzheimer's disease. eNeuro 7(5):1–10

Huang X, Lv D, Yang X et al (2020b) m^{6}A RNA methylation regulators could contribute to the occurrence of chronic obstructive pulmonary disease. J Cell Mol Med 00:1–10

Huang H, Wang Y, Kandpal M et al (2020c) FTO-dependent N6-methyladenosine modifications inhibit ovarian cancer stem cell self-renewal by blocking cAMP signaling. Cancer Res 80:3200–3214

Jia G, Fu Y, Zhao X et al (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887

Jian D, Wang Y, Jian L et al (2020) METTL14 aggravates endothelial inflammation and atherosclerosis by increasing FOXO1 N6-methyladeosine modifications. Theranostics 10:8939

Jiang L, Chen T, Xiong L et al (2020) Knockdown of m^{6}A methyltransferase METTL3 in gastric cancer cells results in suppression of cell proliferation. Oncol Lett 20:2191–2198

Jin H, Ying X, Que B et al (2019) N6-methyladenosine modification of ITGA6 mRNA promotes the development and progression of bladder cancer. EBioMedicine 47:195–207

Kmietczyk V, Riechert E, Kalinski L et al (2019) m^{6}A-mRNA methylation regulates cardiac gene expression and cellular growth. Life Sci Alliance 2(2):1–12

Kudou K, Komatsu T, Nogami J et al (2017) The requirement of Mettl3-promoted MyoD mRNA maintenance in proliferative myoblasts for skeletal muscle differentiation. Open Biol 7:170119

Lan T, Li H, Zhang D et al (2019) KIAA1429 contributes to liver cancer progression through N6-methyladenosine-dependent post-transcriptional modification of GATA3. Mol Cancer 18:1–19

Lee H, Bao S, Qian Y et al (2019) Stage-specific requirement for Mettl3-dependent m^{6}A mRNA methylation during haematopoietic stem cell differentiation. Nat Cell Biol 21:700–709

Li X, Tang J, Huang W et al (2017a) The M^{6}A methyltransferase METTL3: acting as a tumor suppressor in renal cell carcinoma. Oncotarget 8:96103
Li H-B, Tong J, Zhu S et al (2017b) m^{6}A mRNA methylation controls T cell homeostasis by targeting the IL-7/STAT5/SOCS pathways. Nature 548:338–342
Li Z, Weng H, Su R et al (2017c) FTO plays an oncogenic role in acute myeloid leukemia as a N6-methyladenosine RNA demethylase. Cancer Cell 31:127–141
Li M, Zhao X, Wang W et al (2018) Ythdf2-mediated m^{6}A mRNA clearance modulates neural development in mice. Genome Biol 19:1–16
Li D, Cai L, Meng R et al (2020a) METTL3 modulates osteoclast differentiation and function by controlling RNA stability and nuclear export. Int J Mol Sci 21:1660
Li J, Wu L, Pei M et al (2020b) YTHDF2, a protein repressed by miR-145, regulates proliferation, apoptosis and migration in ovarian cancer cells. J Ovarian Res 13:111
Liu J, Yue Y, Han D et al (2014) A METTL3–METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol 10:93–95
Liu N, Dai Q, Zheng G et al (2015) N6-methyladenosine-dependent RNA structural switches regulate RNA–protein interactions. Nature 518:560–564
Liu N, Zhou KI, Parisien M et al (2017) N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. Nucleic Acids Res 45:6051–6063
Liu J, Luo G, Sun J et al (2019a) METTL14 is essential for β-cell survival and insulin secretion. Biochim Biophys Acta Mol basis Dis 1865:2138–2148
Liu L, Wang J, Sun G et al (2019b) m^{6}A mRNA methylation regulates CTNNB1 to promote the proliferation of hepatoblastoma. Mol Cancer 18:1–13
Liu T, Wei Q, Jin J et al (2020a) The m^{6}A reader YTHDF1 promotes ovarian cancer progression via augmenting EIF3C translation. Nucleic Acids Res 48:3816–3831
Liu L, Wu Y, Li Q et al (2020b) METTL3 promotes tumorigenesis and metastasis through BMI1 m6A methylation in oral squamous cell carcinoma. Mol Ther 28(10):2177–2190
Liu Z-F, Yang J, Wei S-P et al (2020c) Upregulated METTL3 in nasopharyngeal carcinoma enhances the motility of cancer cells. Kaohsiung J Med Sci 36:895–903
Ma C, Chang M, Lv H et al (2018) RNA m^{6}A methylation participates in regulation of postnatal development of the mouse cerebellum. Genome Biol 19:68
Ma Z, Li Q, Liu P et al (2020) METTL3 regulates m^{6}A in endometrioid epithelial ovarian cancer independently of METTl14 and WTAP. Cell Biol Int 44:2524–2531
Mao Y, Dong L, Liu X-M et al (2019) m^{6}A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. Nat Commun 10:1–11
Mathiyalagan P, Adamiak M, Mayourian J et al (2019) FTO-dependent N6-methyladenosine regulates cardiac function during remodeling and repair. Circulation 139:518–5132
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149:1635–1646
Meyer KD, Patil DP, Zhou J et al (2015) 5′ UTR m^{6}A promotes cap-independent translation. Cell 163:999–1010
Mittenbühler MJ, Saedler K, Nolte H et al (2020) Hepatic FTO is dispensable for the regulation of metabolism but counteracts HCC development in vivo. Mol Metab 42:101085
Naren D, Yan T, Gong Y et al (2020) High Wilms' tumor 1 associating protein expression predicts poor prognosis in acute myeloid leukemia and regulates m^{6}A methylation of MYC mRNA. J Cancer Res Clin Oncol 1–15
Niu Y, Xu Zhao X, Wu Y-S et al (2013) N6-methyl-adenosine (m^{6}A) in RNA: an old modification with a novel epigenetic function. Genomics Proteomics Bioinformatics 11:8–17
Paris J, Morgan M, Campos J et al (2019) Targeting the RNA m6A reader YTHDF2 selectively compromises cancer stem cells in acute myeloid leukemia. Cell Stem Cell 25:137–148.e6
Patil DP, Chen C-K, Pickering BF et al (2016) m^{6}A RNA methylation promotes XIST-mediated transcriptional repression. Nature 537:369–373
Pendleton KE, Chen B, Liu K et al (2017) The U6 snRNA m^{6}A methyltransferase METTL16 regulates SAM synthetase intron retention. Cell 169:824–835.e14

Peng W, Li J, Chen R et al (2019) Upregulated METTL3 promotes metastasis of colorectal Cancer via miR-1246/SPRED2/MAPK signaling pathway. J Exp Clin Cancer Res 38:393

Peng F, Xu J, Cui B et al (2020) Oncogenic AURKA-enhanced N6-methyladenosine modification increases DROSHA mRNA stability to transactivate STC1 in breast cancer stem-like cells. Cell Res 1–17

Pi J, Wang W, Ji M et al (2020) YTHDF1 promotes gastric carcinogenesis by controlling translation of FZD7. Cancer Res

Ping X-L, Sun B-F, Wang L et al (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res 24:177–189

Pu X, Gu Z, Gu Z (2020) ALKBH5 regulates IGF1R expression to promote the proliferation and tumorigenicity of endometrial cancer. J Cancer 11:5612

Qian J-Y, Gao J, Sun X et al (2019) KIAA1429 acts as an oncogenic factor in breast cancer by regulating CDK1 in an N6-methyladenosine-independent manner. Oncogene 38:6123–6141

Roundtree IA, He C (2016) RNA epigenetics-chemical messages for posttranscriptional gene regulation. Curr Opin Chem Biol 30:46–51

Schwartz S, Mumbach MR, Jovanovic M et al (2014) Perturbation of m^{6}A writers reveals two distinct classes of mRNA methylation at internal and 5′ sites. Cell Rep 8:284–296

Shen C, Sheng Y, Zhu AC et al (2020) RNA demethylase ALKBH5 selectively promotes tumorigenesis and cancer stem cell self-renewal in acute myeloid leukemia. Cell Stem Cell 27 (1):64–80.e9

Shi H, Wang X, Lu Z et al (2017) YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. Cell Res 27:315–328

Shi H, Zhang X, Weng Y-L et al (2018) m^{6}A facilitates hippocampus-dependent learning and memory through YTHDF1. Nature 563:249–253

Shi H, Wei J, He C (2019) Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers. Mol Cell 74:640–650

Shi Y, Zheng C, Jin Y et al (2020) Reduced expression of METTL3 promotes metastasis of triple-negative breast cancer by m^{6}A methylation-mediated COL3A1 up-regulation. Front Oncol 10:1126

Song H, Feng X, Zhang H et al (2019) METTL3 and ALKBH5 oppositely regulate m^{6}A modification of TFEB mRNA, which dictates the fate of hypoxia/reoxygenation-treated cardiomyocytes. Autophagy 15:1419–1437

Sun S, Han Q, Liang M et al (2020a) Downregulation of m^{6}A reader YTHDC2 promotes tumor progression and predicts poor prognosis in non-small cell lung cancer. Thorac Cancer 11:3269–3279

Sun Y, Li S, Yu W et al (2020b) N6-methyladenosine-dependent pri-miR-17-92 maturation suppresses PTEN/TMEM127 and promotes sensitivity to everolimus in gastric cancer. Cell Death Dis 11:1–16

Sun T, Wu Z, Wang X et al (2020c) LNC942 promoting METTL14-mediated m^{6}A methylation in breast cancer cell proliferation and progression. Oncogene 39:5358–5372

Tang J, Zheng C, Zheng F et al (2020) Global N6-methyladenosine profiling of cobalt-exposed cortex and human neuroblastoma H4 cells presents epitranscriptomics alterations in neurodegenerative disease-associated genes. Environ Pollut 266:115326

Tian C, Huang Y, Li Q et al (2019) Mettl3 regulates osteogenic differentiation and alternative splicing of Vegfa in bone marrow mesenchymal stem cells. Int J Mol Sci 20:551

Tong J, Cao G, Zhang T et al (2018a) m^{6}A mRNA methylation sustains Treg suppressive functions. Cell Res 28:253–256

Tong J, Flavell RA, Li H-B (2018b) RNA m^{6}A modification and its function in diseases. Front Med 12:481–489

Tsuruta N, Tsuchihashi K, Ohmura H et al (2020) RNA N6-methyladenosine demethylase FTO regulates PD-L1 expression in colon cancer cells. Biochem Biophys Res Commun 530:235–239

Uddin MB, Roy KR, Hosain SB et al (2019) An N6-methyladenosine at the transited codon 273 of p53 pre-mRNA promotes the expression of R273H mutant protein and drug resistance of cancer cells. Biochem Pharmacol 160:134–145

Uddin MB, Wang Z, Yang C (2020) Dysregulations of functional RNA modifications in cancer, cancer stemness and cancer therapeutics. Theranostics 10:3164

Visvanathan A, Patil V, Arora A et al (2018) Essential role of METTL3-mediated m^6A modification in glioma stem-like cells maintenance and radioresistance. Oncogene 37:522–533

Vu LP, Pickering BF, Cheng Y et al (2017) The N6-methyladenosine (m^6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. Nat Med 23:1369

Wang Y, Li Y, Toth JI et al (2014a) N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nat Cell Biol 16:191–198

Wang X, Lu Z, Gomez A et al (2014b) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505:117–120

Wang X, Zhao BS, Roundtree IA et al (2015) N6-methyladenosine modulates messenger RNA translation efficiency. Cell 161:1388–1399

Wang X, Feng J, Xue Y et al (2016) Structural basis of N6-adenosine methylation by the METTL3-METTL14 complex. Nature 534:575–578

Wang X, Huang N, Yang M et al (2017) FTO is required for myogenesis by positively regulating mTOR-PGC-1 α pathway-mediated mitochondria biogenesis. Cell Death Dis 8:e2702–e2e02

Wang C-X, Cui G-S, Liu X et al (2018a) METTL3-mediated m^6A modification is required for cerebellar development. PLoS Biol 16:e2004880

Wang Y, Li Y, Yue M et al (2018b) N6-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications. Nat Neurosci 21:195–206

Wang H, Hu X, Huang M et al (2019) Mettl3-mediated mRNA m^6A methylation promotes dendritic cell activation. Nat Commun 10:1–12

Wang Y, Sun J, Lin Z et al (2020) m^6A mRNA methylation controls functional maturation in neonatal murine β cells. Diabetes 69(8):1708–1722

Wen J, Lv R, Ma H et al (2018) Zc3h13 regulates nuclear RNA m^6A methylation and mouse embryonic stem cell self-renewal. Mol Cell 69:1028–1038.e6

Weng H, Huang H, Wu H et al (2018) METTL14 inhibits hematopoietic stem/progenitor differentiation and promotes leukemogenesis via mRNA m^6A modification. Cell Stem Cell 22:191–205.e9

Winkler R, Gillis E, Lasman L et al (2019) m^6A modification controls the innate immune response to infection by targeting type I interferons. Nat Immunol 20:173–182

Wu Y, Xie L, Wang M et al (2018) Mettl3-mediated m^6A RNA methylation regulates the fate of bone marrow mesenchymal stem cells and osteoporosis. Nat Commun 9:1–12

Wu R, Guo G, Bi Z et al (2019a) m^6A methylation modulates adipogenesis through JAK2-STAT3-C/EBPβ signaling. Biochim Biophys Acta Gene Regul Mech 1862:796–806

Wu R, Liu Y, Zhao Y et al (2019b) m^6A methylation controls pluripotency of porcine induced pluripotent stem cells by targeting SOCS3/JAK2/STAT3 pathway in a YTHDF1/YTHDF2-orchestrated manner. Cell Death Dis 10:171

Xiang S, Liang X, Yin S et al (2020) N6-methyladenosine methyltransferase METTL3 promotes colorectal cancer cell proliferation through enhancing MYC expression. Am J Transl Res 12:1789

Xiao W, Adhikari S, Dahal U et al (2016) Nuclear m^6A reader YTHDC1 regulates mRNA splicing. Mol Cell 61:507–519

Xiao L, Li X, Mu Z et al (2020) FTO inhibition enhances the antitumor effect of temozolomide by targeting MYC-miR-155/23a cluster-MXI1 feedback circuit in glioma. Cancer Res 80:3945–3958

Xie H, Li J, Ying Y et al (2020) METTL3/YTHDF2 m^6A axis promotes tumorigenesis by degrading SETD7 and KLF4 mRNAs in bladder cancer. J Cell Mol Med 24:4092–4104

Xu C, Wang X, Liu K et al (2014) Structural basis for selective binding of m^{6}A RNA by the YTHDC1 YTH domain. Nat Chem Biol 10:927–929

Xu J, Chen Q, Tian K et al (2020a) m^{6}A methyltransferase METTL3 maintains colon cancer tumorigenicity by suppressing SOCS2 to promote cell proliferation. Oncol Rep 44:973–986

Xu Y, Ye S, Zhang N et al (2020b) The FTO/miR-181b-3p/ARL5B signaling pathway regulates cell migration and invasion in breast cancer. Cancer Commun 40(10):484–500

Xue L, Li J, Lin Y et al (2020) m^{6}A transferase METTL3-induced lncRNA ABHD11-AS1 promotes the Warburg effect of non-small-cell lung cancer. J Cell Physiol

Yang C (2020) ToxPoint: dissecting functional RNA modifications in responses to environmental exposure-mechanistic toxicology research enters a new era. Toxicol Sci 174:1–2

Yang Y, Hsu PJ, Chen Y-S et al (2018) Dynamic transcriptomic m^{6}A decoration: writers, erasers, readers and functions in RNA metabolism. Cell Res 28:616–624

Yang F, Jin H, Que B et al (2019) Dynamic m^{6}A mRNA methylation reveals the role of METTL3-m^{6}A-CDCP1 signaling axis in chemical carcinogenesis. Oncogene 38:4755–4772

Yang Z, Jiang X, Li D et al (2020a) HBXIP promotes gastric cancer via METTL3-mediated MYC mRNA m^{6}A modification. Aging 12

Yang Z, Jiang X, Zhang Z et al (2020b) HDAC3-dependent transcriptional repression of FOXA2 regulates FTO/m^{6}A/MYC signaling to contribute to the development of gastric cancer. Cancer Gene Ther:1–15

Yao Y, Bi Z, Wu R et al (2019) METTL3 inhibits BMSC adipogenic differentiation by targeting the JAK1/STAT5/C/EBPβ pathway via an m^{6}A-YTHDF2-dependent manner. FASEB J 33:7529–7544

Yi D, Wang R, Shi X et al (2020) METTL14 promotes the migration and invasion of breast cancer cells by modulating N6-methyladenosine and hsa-miR-146a-5p expression. Oncol Rep 43:1375–1386

Yoon K-J, Ringeling FR, Vissers C et al (2017) Temporal control of mammalian cortical neurogenesis by m^{6}A methylation. Cell 171:877–889.e17

Yue Y, Liu J, He C (2015) RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. Genes Dev 29:1343–1355

Yue Y, Liu J, Cui X et al (2018) VIRMA mediates preferential m^{6}A mRNA methylation in 3′ UTR and near stop codon and associates with alternative polyadenylation. Cell Discov 4:1–17

Zaccara S, Ries RJ, Jaffrey SR (2019) Reading, writing and erasing mRNA methylation. Nat Rev Mol Cell Biol 1–17

Zhang S, Zhao BS, Zhou A et al (2017) m^{6}A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. Cancer Cell 31:591–606.e6

Zhang C, Huang S, Zhuang H et al (2020a) YTHDF2 promotes the liver cancer stem cell phenotype and cancer metastasis by regulating OCT4 expression via m^{6}A RNA methylation. Oncogene 39:4507–4518

Zhang S, Meng P, Cheng S et al (2020b) Pregnancy exposure to carbon black nanoparticles induced neurobehavioral deficits that are associated with altered m^{6}A modification in offspring. Neurotoxicology 81:40–50

Zhang Y, Wang D, Wu D et al (2020c) Long noncoding RNA KCNMB2-AS1 stabilized by N6-methyladenosine modification promotes cervical cancer growth through acting as a competing endogenous RNA. Cell Transplant 29:1–11

Zhao T, Li X, Sun D et al (2019) Oxidative stress: one potential factor for arsenite-induced increase of N6-methyladenosine in human keratinocytes. Environ Toxicol Pharmacol 69:95–103

Zhao W, Cui Y, Liu L et al (2020a) METTL3 facilitates oral squamous cell carcinoma tumorigenesis by enhancing c-Myc stability via YTHDF1-mediated m^{6}A modification. Mol Ther Nucleic Acids 20:1–12

Zhao Y, Shi Y, Shen H et al (2020b) m^{6}A-binding proteins: the emerging crucial performers in epigenetics. J Hematol Oncol 13:1–14

Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29

Zheng Z, Zhang L, Cui X-L et al (2020) Control of early B cell development by the RNA N6-methyladenosine methylation. Cell Rep 31:107819

Zhu Y, Pan X, Du N et al (2020a) ASIC1a regulates miR-350/SPRY2 by N6-methyladenosine to promote liver fibrosis. FASEB J 34(11):14371–14388

Zhu W, Si Y, Xu J et al (2020b) Methyltransferase like 3 promotes colorectal cancer proliferation by stabilizing CCNE1 mRNA in an m^6A-dependent manner. J Cell Mol Med 24:3521–3533

Zhuang C, Zhuang C, Luo X et al (2019) N6-methyladenosine demethylase FTO suppresses clear cell renal cell carcinoma through a novel FTO-PGC-1α signalling axis. J Cell Mol Med 23:2163–2173

N^6-Methyladenosine in the Heart

Vivien Kmietczyk, Ellen Malovrh, and Mirko Völkers

Contents

Abstract Similar to DNA modifications, an epitranscriptome of mRNA modifications emerged over the last few years in the cardiovascular system. Specifically, the discovery of the N6-methyladenosine (m^6A) as reversible post-transcriptional modification in mRNAs raised recent interest in those modifications in cardiomyocytes and their role in diseased hearts. As mRNA levels of different genes have been shown to not necessarily correspond to respective protein levels, post-transcriptional mRNA modifications have been recognized as a relevant mechanism of gene expression in cardiomyocytes. Consequently, m^6A-mRNA methylomes of human cardiac tissue and of animal model heart samples were reported and point toward an important role of m^6A modifications in cardiomyocyte cell function. Clearly, additional studies are needed to fully understand this novel stress–response mechanism in the heart for maintaining normal cardiac function. In this chapter, we review the current state of the art of epitranscriptomic research specifically in cardiomyocytes.

V. Kmietczyk · E. Malovrh · M. Völkers (✉)
German Center for Cardiovascular Research (DZHK) – Partner site Heidelberg/Mannheim, Heidelberg, Germany

Department of Internal Medicine III, University Hospital Heidelberg, Heidelberg, Germany
e-mail: mirko.voelkers@med.uni-heidelberg.de

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_11

Keywords Epitranscriptome · m^6A · Heart failure · Growth · Ischemia · Cardiomyocytes

1 Introduction

The heart is a muscular organ providing constant blood flow through the body. Main characteristics of the heart are contractility and rhythmicity, which are autoregulated by a specialized conducting system. Contracting cells of the heart are cardiomyocytes which only account for one-third of the total cell count but for 70–80% of the heart's mass (Nag 1980). Worldwide, diseases of the cardiovascular system are the leading cause of death according to the World Health Organization (Roth et al. 2018; WHO 1992).

Despite the progress in understanding the pathophysiology of cardiovascular diseases, identifying risk factors, and developing novel treatment options, one of the main disease subclasses associated with very poor prognosis is heart failure (HF), a condition in which the heart is not capable of maintaining the blood flow to meet the body's needs.

Especially heart failure is a major public health issue with a considerable burden for our healthcare system. Despite recent progress in understanding the pathophysiology, heart failure still carries a 5-year mortality that rivals most cancers. Multiple risk factors predict outcome and severity with the most important predisposing factor being ischemic heart disease, but the greater prevalence of hypertension contributes more to the incidence of heart failure in the Western world.

HF is typically caused by diseases such as hypertension and myocardial infarction or caused by hereditary forms such as genetic cardiomyopathies. HF is a detrimental clinical syndrome, characterized by insufficient cardiac blood output for the maintenance of the metabolic demands of the body during exercise and ultimately at rest. Among the abovementioned risk factors for developing HF, myocardial infarctions (MIs) are the most commonly known adverse cardiac events. The obstruction of a coronary artery leads to the reduction or stopping of blood flow to a part of the cardiac muscle, leading to a state of hypoxia or ischemia, thereby damaging the tissue. Available treatment options have reduced its acute mortality in the last few years. Nonetheless, the risk of developing HF after myocardial infarctions is still increasing dramatically (Mill et al. 2011).

Independent of the etiology, HF involves changes in cardiac structure, contractile dysfunction, as well as multiple genetic and molecular alterations that impact heart function and patient survival. These changes have been referred to as "cardiac remodeling." Changes in gene expression with up- or downregulation of specific sets of genes play a fundamental role during the pathogenesis of heart failure. Cardiac remodeling is clinically manifested by changes in size, shape, and function of the heart. The heart grows in size and mass to compensate the additional load or the injury and to regain normal cardiovascular function in the resting state. In

contrast to neoplastic tissue growth, the described volume and mass expansion of the heart is caused by hypertrophic growth due to increases in length or width of cardiomyocytes. In pathological hypertrophic hearts, oxygen diffusion distances from the capillaries are exceeded due to enlarged cardiomyocytes, which causes an ischemic environment inducing cardiomyocyte apoptosis (Cohn et al. 2000).

2 Gene Expression Control in Cardiac Myocytes

Dysfunctional cardiomyocytes that accumulate following pathological damage are characterized by altered growth kinetics, metabolic changes, and increased susceptibility to cell death. In addition, remodeling events include temporal changes in collagen deposition, capillary density, and the cell populations responsible for these changes. Collectively, these pathophysiological changes are the logical target for therapeutic intervention to inhibit heart failure.

Certainly, differential protein expression patterns are the basis of dysfunctional cardiomyocytes, regulating growth and modifying the structural and functional properties of cardiomyocytes. Myosin heavy chain (MHC) is the major constituent of the contractile protein complex. Its α-isoform is usually expressed in adult cardiomyocytes, while the β-isoform is present in the fetal but also in the hypertrophying heart. Also, during cardiac remodeling the natriuretic precursor peptides a and b (Nppa and Nppb) are highly expressed (Dirkx et al. 2013) and the upregulation of extracellular matrix and cytoskeletal proteins is often associated with the development of heart failure (Tan et al. 2002).

Mitogenic stimuli, cardiac agonists, and mechanical stress induce signaling cascades such as the mitogen-associated protein kinase (MAPK), the calcineurin/calmodulin, or the G protein-coupled receptor (GPCR) pathways, resulting in the stimulation of various transcription and epigenetic factors and finally in the specific changes in gene expression changes. In human heart failure prototypical transcription factors are highly active and contribute to the progression of the disease: Myocyte enhancer factor 2 (Mef2), the homeobox protein NKX, the nuclear factor of activated T-cells (Nfat), and the GATA binding protein 4 (Gata4). These transcription factors act in concert to regulate pathological cardiac growth by induction of several fetal and cardiac genes involved in contractility and growth control. In turn, selectively induced chromatin condensation by histone deacetylases and other epigenetic factors represses the Mef2 stimulated gene expression. Furthermore, the transcription factor serum response factor (Srf) regulates the expression of numerous muscle-specific genes, including cardiac hypertrophic genes such as *Nppa* and α-skeletal actin (Dirkx et al. 2013; Frey and Olson 2003).

In addition, overall mRNA translation and protein synthesis is highly increased in hypertrophying cardiomyocytes with specific mRNAs that are selectively translationally upregulated (Nagai et al. 1988; Spruill et al. 2008). Hence, not only transcriptional but also post-transcriptional regulation of gene expression significantly contributes to cardiac disease initiation and progression.

Regulating small RNAs such as microRNAs (miRNAs) or long ncRNAs (lncRNAs), as well as short open reading frames (sORFs) coding for micropeptides (Van Heesch et al. 2019), add to the complex layer of gene expression control in the heart. miRNAs, small noncoding nucleotides, silence gene expression post-transcriptionally by binding to complementary sequences, usually located in the 3′ UTR of the target transcript. Examples such as miR-1, miR-133, miR-208, and miR-499 are miRNAs that are enriched in the heart and play important roles in the pathogenesis of heart failure but also in the maintenance of cardiomyocyte physiology. For instance, miR-1 inhibits cardiac hypertrophy by targeting transcripts like *CALM1* and *CALM2*, *MEF2A*, and *GATA4*. Another interesting target transcript of miR-1 is the Ras homolog enriched in brain (Rheb) (Tang et al. 2020), which activates the mammalian target of rapamycin (mTOR), an important player in hypertrophic growth (Völkers et al. 2013). Also, lncRNA Myheart (Mhrt) prevents cardiac hypertrophy by blocking chromatin remodeling through Brg1 induced in stressed cardiomyocytes (Han et al. 2014). Other lncRNAs, like Chrf, rather act as miRNA sponges and thereby inhibit the repressing action of the miRNA, which in the case of Chrf is *MYD88* (Wang et al. 2014a).

3 m^6A mRNA Methylation

Although the knowledge about m^6A-dependent post-transcriptional regulation of gene expression in other tissues and contexts has been studied extensively, the cardiac research has not considered this additional path of gene expression to the same extent yet.

To date, more than 150 post-transcriptional base modifications have been identified in different RNA classes. Similar to DNA and protein modifications, RNA modifications are implemented through writer enzymes and may be removed through eraser proteins. The presence or absence of RNA modifications is typically interpreted through reader proteins. Currently, they are only five studies found on PubMed, all published in 2019, in comparison to about 2000 on m^6A in other research areas and around 580 in the cancer field (pubmed.ncbi. nlm.nih.gov search m6A+/− heart/cardiomyocyte or + cancer). In the following sections, we will focus on the role of m^6A mRNA methylation and its role in cardiovascular disease.

Methylation of rRNA was already found in 1966, and methylation of mRNA was reported in 1974 in eukaryotic cells. N^6-methyladenosine (m^6A) is the primary base methylation of mRNA in eukaryotic cells, viruses, and yeast. m^6A appears in evolutionary conserved motifs across species with a consensus sequence containing the m^6A site in the center (Fu et al. 2014). m^6A sites are enriched at the 3′-end of mRNA, clustering around the stop codon (Meyer et al. 2012), either before or within the last exon (Ke et al. 2015).

m^6A writers or methylases transfer the methyl group from *S*-adenosylmethionine (SAM) to the adenosine. The m^6A methylase is an enzymatic complex comprising N6-adenosine-methyltransferase subunit (Mettl3), the N6-adenosine-

methyltransferase non-catalytic subunit (Mettl14), and the Wilms tumor-associated protein (Wtap). Only Mettl3 possesses a SAM-binding site and demonstrates catalytic activity. However, binding to Mettl14, as the allosteric regulator, is essential for the full activity (Zaccara and Jaffrey 2020).

The methylation of mRNA is reversible and can be removed by Fat mass and obesity-associated protein (Fto) and AlkB family member 5 (Alkbh5). Both exhibit a demethylation activity toward mRNA containing m^6A (Fu et al. 2014).

m^6A is involved in several aspects of cellular functions, exerted by adjusting the mRNA life cycle. m^6A can enhance translation efficiency and reduces mRNA stability (Wang et al. 2014b) mainly through the binding by so-called m^6A readers like the YTH domain family (Ythdf1–Ythdf3). A recent study indicates that all three Ytdhf paralogs have similar functions and can compensate the loss of one Ythdf paralog; however, many earlier studies report distinct functions at least for Ytdhf1 and Ytdhf2. Ythdf2 stimulates mRNA degradation by transporting it to decay sites. Ythdf1 induces translation by binding to m^6A near stop codons and recruiting eIF3, a multiprotein complex involved in canonical translation initiation. Ythdf3 was shown to work in cooperation with Ythdf1 regulating mRNA translation, but it was also reported to withhold the mRNAs from the translatable pool, thereby preventing translation (Zaccara and Jaffrey 2020).

4 Differential m^6A Methylation in Diseased Hearts

In the healthy mouse heart, most m^6A carrying transcripts possess multiple marks in 5′ and 3′ UTR and the coding sequence (Berulava et al. 2020). While in healthy mouse heart tissue m^6A is present at lower levels in comparison to liver, kidney, and brain (Meyer et al. 2012), a rise of m^6A has been found in human failing hearts as well as animal models of heart failure (Kmietczyk et al. 2019; Dorn et al. 2019; Mathiyalagan et al. 2018). Moreover, the number of differentially methylated transcripts intensively exceeds the number of differential expressed genes in murine hypertrophying and failing hearts (Berulava et al. 2020). Also, cultured hypertrophic cardiomyocytes possess increased levels of m^6A-mRNA (Dorn et al. 2019). Similarly, increased m^6A levels are found in ischemia reperfused mouse hearts and cardiomyocytes stressed by hypoxia and reoxygenation, modeling conditions of myocardial infarctions followed the opening of the occluded artery, increased m^6A levels are found (Song et al. 2019). Interestingly, Berulava reports no changes in m^6A distribution on the transcriptome in hypertrophying murine hearts but finds hypomethylated start sites in murine failing hearts (Berulava et al. 2020).

According to the study published by Mathiyalagan et al. the decrease of FTO expression in failing heart may cause the augmented levels of m^6A. However, in hearts from myocardial infarction patients, METTL3 was found to be upregulated as well (Mathiyalagan et al. 2018). In contrast, Song et al. showed a parallel increase of Mettl3 with augmented m^6A levels in ischemia/reoxygenated cardiomyocytes. Also, additional data indicate that in a subset of dilative cardiomyopathy patients the m^6A

writer METLL3 is higher expressed compared to healthy human heart samples (Kmietczyk et al. 2019).

Similar to the epitransciptome in other tissue and cells, the m^6A sites on mRNAs in healthy and diseased cardiomyocytes are not distributed equally on the global transcriptome but rather on specific functional groups of transcripts. In DCM hearts genes from m^6A transcripts are involved in processes of actin binding, calmodulin binding, and delayed potassium channel activity (Kmietczyk et al. 2019). Actin is a part of the cytoskeleton providing structural support, scaffolding the myofibrils, and its abnormalities can cause cardiomyopathies (Sequeira et al. 2014). Calmodulin is important for calcium sensory processes in cardiomyocytes and regulation of cardiac gene expression (Dewenter et al. 2017), and the latter plays essential roles in cardiac repolarization, with its functional disturbance leading to arrhythmia (Chen et al. 2016). In induced murine failing hearts m^6A is also increased on mRNAs coding for calcium handling and contractile proteins ischemia induced murine heart failure (Mathiyalagan et al. 2018). In contrast, Berulava et al. specifically reported the absence of differential methylation on transcripts for contractile processes. In hypertrophic hearts rather transcripts coding for proteins involved in metabolic processes, gene expression regulation, and signaling were found, while in failing hearts modified mRNAs are specifically involved in mitochondrial functions, cardiac muscle development, and again metabolic processes (Berulava et al. 2020). Figure 1 summarizes our current knowledge about m^6A in human and experimental models of heart failure.

5 Manipulations of the m^6A Writers and Erasers Affect Cardiomyocyte Growth, Function, and Survival

The post-transcriptional regulation of gene expression by m^6A is intriguingly relevant for pathological cardiomyocyte growth and survival, shown by studies using manipulations of m^6A-mRNA levels either by loss- or gain-of-function experiments of m^6A writers or erasers in vitro and in vivo. However, contrary results were obtained from different research groups, which might be explained by differences in cell purification, culturing, and treatment procedures.

Cardiac growth (increase in cell mass) is a highly regulated and controlled process. This includes initiation of gene transcription, protein synthesis by controlling ribosomal functions, as well as nutrient and mitochondrial metabolism. Cardiomyocytes have mostly lost the capacity to divide but retain the ability to increase their cell mass. Cardiomyocytes withdraw from the cell cycle in neonates and continue to grow to obtain the correct adult heart size. During pathological stress, although fibroblast proliferation or other resident proliferating cells can increase cardiac mass, mass increase primarily occurs through the hypertrophy of individual cardiomyocytes. This form of pathological growth is an independent predictor of morbidity and mortality in human patients. However, after physiological

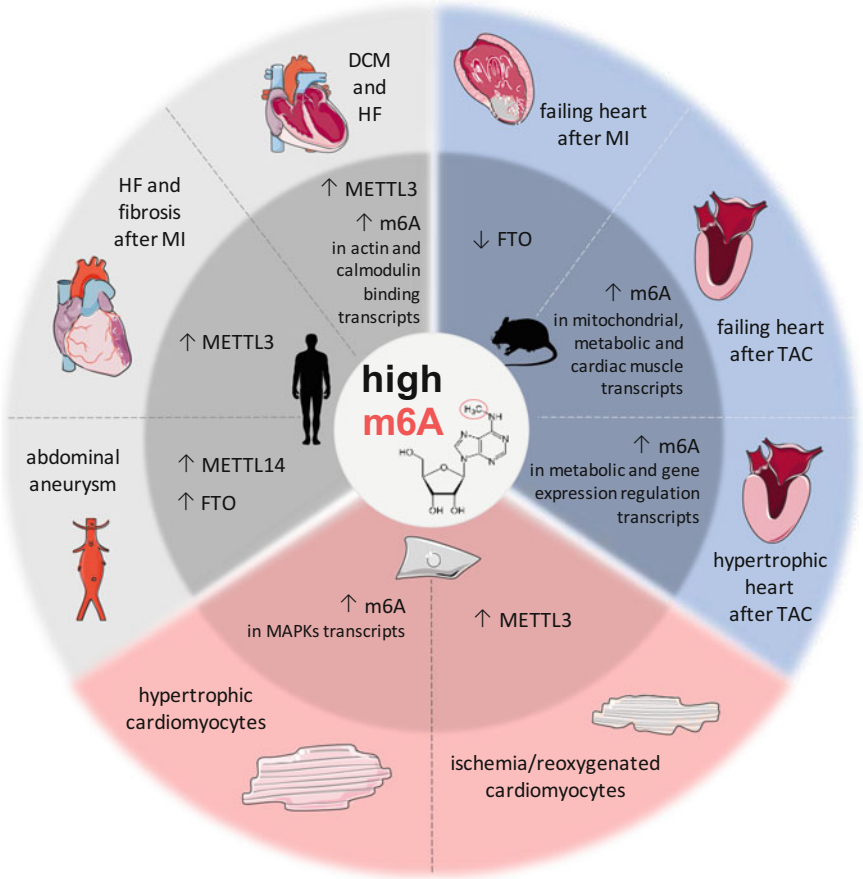


Fig. 1 Human diseased hearts and experimental models of heart failure exhibit increased levels of m^{6}A-mRNA. This is mostly due to the upregulation of m^{6}A-methylase subunit METTL3 and decrease of demethylase FTO in a variety of cardiac diseases. The modified m6A sites are not equally distributed throughout the transcriptome, but rather accumulate on specific functional groups of transcripts

stress (such as exercise or pregnancy) the myocardium undergoes an adaptive form of hypertrophy, which does not associate with disease and is fully reversible.

In our own experiments, Mettl3 overexpression, which increases the levels of m^{6}A in neonatal rat cardiomyocytes, leads to reduction of pathological hypertrophy in isolated cardiomyocytes after stimulation with phenylephrine, a α_1-adrenergic receptor agonist. In line, silencing of Mettl3 in order to reduce levels of m^{6}A-mRNA abundance leads to spontaneous hypertrophy of isolated cardiomyocytes even without neurohumoral stimulation (Kmietczyk et al. 2019).

As mentioned above, myocardial infarction (MI) is the most frequent cause of heart failure. Prolonged myocardial ischemia results in cellular loss by necrosis and apoptosis of cardiac cells. The adult mammalian heart has negligible regenerative

capacity; thus, the infarcted myocardium heals through formation of a scar. This myocardial healing again is associated with progressive cardiac remodeling and dysfunction. Progressive cell death is also a hallmark of chronic heart failure, and autophagic cell death, apoptosis, and necrosis have been observed during progression of heart disease. The ability to respond rapidly to changes in the environment (hypoxia, oxidative stress, or nutrient deprivations) is essential for survival for cardiac cells. Especially cardiac myocytes need remarkable mechanisms of cytoprotection against cellular stress, due to the absence of cell division and lack of regeneration capacity after cellular loss. Since most cellular processes are mediated by proteins, it is crucial for specialized cells to effect rapid and dynamic changes in protein levels in any stress response. It is still unknown how translational control of gene expression in response to hypoxia and how regulated changes in synthesis of subsets of proteins from a pool of mRNAs affect stress response in cardiomyocytes. However, Mettl3 silencing reduces apoptosis after 4 h of hypoxia followed by 3 h of reoxygenation (Song et al. 2019), while Mettl3 overexpression inhibits autophagy and induces cell death. The knockdown of Fto sensitizes primary cardiomyocytes to arrhythmic events. On the contrary, cardiomyocyte dysfunction induced by hypoxia is mitigated by the overexpression of Fto (Mathiyalagan et al. 2018).

In vivo data suggest that dysfunctional manipulations of m6A levels in the heart affect pathological growth, survival, and the contractility caused by deregulated expression of contractile components in cardiomyocytes (Kmietczyk et al. 2019; Mathiyalagan et al. 2018). In contrast to in vitro data, Mettl3 overexpression in adult mice does not affect heart growth or function, gene expression, or remodeling. Nevertheless, the additional induction of pressure overload to the heart by transverse aortic constriction, a well-established and accepted model for the induction of hypertrophy and heart failure (Dealmeida et al. 2010), causes diminished growth response characterized by reduced heart weight/body weight ratio, left ventricular wall thickening, and decreased cell surface areas in comparison to control mice. Although cardiac hypertrophy was reduced, the function was not restored in Mettl3-TAC mice (Kmietczyk et al. 2019).

Dorn et al. in turn reported that Mettl3 overexpression increased cellular growth while knockout of Mettl3 blocked the hypertrophic response in neonatal rat cardiomyocytes. Long-term cardiomyocyte-specific overexpression of Mettl3 in a transgenic mouse model, in which the overexpression starts during embryogenesis, caused mild physiological hypertrophy with preserved cardiac function at 8 months of age. The additional performance of transverse aortic constriction surgery on the mice leads to a similar outcome in cardiac function and growth similarly in control and Mettl3-overexpressing mice. Interestingly, Mettl3 is not required for the embryonic development of the heart during embryogenesis and postnatal growth. However, aged Mettl3 knockout mice develop heart failure with a phenotype and dysfunction comparable to DCM hearts. Similarly, Mettl3 KO mice present a dilated left ventricle and develop heart failure upon transverse aortic banding (Dorn et al. 2019).

Despite the opposing results from experiments with gain and loss of function of Mettl3 (Table 1), it is obvious that varying m^6A levels affect and modulate the

Table 1 Published effects on the heart of the manipulations on Mettl3 or FTO expression in mice

Cardiomyocyte-specific Mettl3 OE by AAV9 (Kmietczyk et al. 2019)	No effect on heart growth or function Reduced growth in response to pressure overload but no restoration of the heart function
Cardiomyocyte-specific Mettl3 KI mouse (OE) (Dorn et al. 2019)	Physiological hypertrophy at 8 months of age No change in growth or function after pressure overload
Cardiomyocyte-specific Mettl3 KO mouse (Dorn et al. 2019)	Decreased cardiac function in aged mice Decreased cardiac function after pressure overload
FTO KO (Carnevali et al. 2014)	Cardiac arrhythmias and hypertrophy
FTO OE (Mathiyalagan et al. 2018)	Improvement of cardiac function after myocardial infarction

growth and function of cardiomyocyte, presumably by post-transcriptional regulation of gene expression of a subset of genes that is involved in cardiomyocyte growth. Indeed, contractile transcripts are highly methylated in failing hearts and augmented m^6A levels lead to increases of arrhythmic events in cardiomyocytes (Mathiyalagan et al. 2018). However, Berulava et al. propose that alterations of m^6A do not correlate to transcript level changes since the m^6A transcriptome highly differs from the differentially expressed transcriptome in human failing hearts (Berulava et al. 2020).

This is further corroborated by results of Carnevallo et al. and Mathiyalagan et al. The loss of Fto in mice leads to cardiac arrhythmias and hypertrophy, although by the time it was published authors did not link these findings to its demethylase activity (Carnevali et al. 2014). Transcripts coding for contractile proteins are highly methylated in failing hearts and increased m^6A levels by Fto knockdown lead to increases of arrhythmic events in cardiomyocytes. On the contrary, the Fto overexpression in cardiomyocytes drastically improves the hypoxia-induced dysfunction and in mice cardiac function is improved along with induction of angiogenesis and reduction of fibrosis after myocardial infarction (Mathiyalagan et al. 2018).

6 Mechanisms

Highly methylated transcripts in hearts from mice have been found to be enriched for transcriptional regulation, cellular signaling, and cardiac muscular proteins, indicating a shift of the mRNA methylation toward transcripts of cardiomyocyte function and growth. In line with studies in other tissues and cell types, it has been suggested that m^6A regulates gene expression by influencing mRNA stability and translation efficiency of specific transcripts. In order to identify differential translated transcripts, ribosomal sequencing has been used in cardiomyocytes after Mettl3 overexpression, but the causal role of identified downstream targets dependent on Mettl3 in regulating cardiomyocyte cell size or cardiac function have not been

confirmed yet (Kmietczyk et al. 2019). Similar studies after knockout of Mettl3 or the other involved key proteins are still missing. However, like mentioned before in ischemic failing hearts it has been shown that expression of the demethylase Fto is reduced which is associated with an increase in m^6A mRNA levels. The overexpression of FTO enhances cardiac function and angiogenesis and reduces fibrosis. This can be partially explained by the finding that Fto selectively demethylates cardiac contractile transcripts such as Serca2a or RyR2 and thereby increases their expression and subsequently calcium handling and cardiac contractility. In ischemic failing hearts, this mechanism is impaired and can be rescued by the Fto overexpression, presenting a new therapeutic approach for heart failure (Mathiyalagan et al. 2018). Besides, decreased protein levels of calmodulin 1 (Calm1) that were not caused by transcriptional changes but rather by translational changes due to reduced methylation of the transcript were reported in human as well as mouse failing hearts (Berulava et al. 2020).

In ischemia/reoxygenated cardiomyocytes the increase of Mettl3 in parallel with augmented m^6A levels leads to reduced mRNA stability and expression of the transcription factor EB (TFEB) through the methylation in the 3′UTR. This is induced by binding of heterogeneous nuclear ribonucleoprotein D (HNRNPD) to the m^6A residues on the TFEB mRNA. TFEB plays a major role in lysosomal biogenesis and regulation of autophagy. Knockdown of Mettl3 in ischemia/reoxygenated cardiomyocytes leads to improved autophagy in combination with inhibition of apoptosis via increased expression of TFEB. Hence, the increased m^6A on TFEB mRNA caused by the augmented expression of Mettl3 in ischemia/reoxygenated cardiomyocytes leads to less expression of TFEB causing reduction of autophagy and augmented apoptosis. Interestingly, TFEB and Mettl3 regulate each other in a negative feedback loop. TFEB reduces *Mettl3* mRNA stability, and additionally it induces the transcription of m^6A demethylase *ALKBH5* by binding to its promoter (Song et al. 2019) (Fig. 2).

Whether writers or erasers compete with targets under different stress conditions and how this might be regulated is unknown and requires further studies. More studies that focus on mechanisms of m^6A-dependent control of cardiac gene expression are clearly needed in order to understand the molecular basis of heart failure development and progression. Use of cell-type-specific knockout mice targeting the different components of writers, erasers, and reader proteins should help to dissect different molecular mechanisms of mRNA modification in the diseased heart. Since the m^6A-reader proteins such as YTHDF1, 2 and 3, which bind to m^6A and regulate different aspects of the mRNA fate, are key regulators of methylated transcripts, future studies should be performed to reveal the involvement of those reader proteins in the m^6A-dependent cardiomyocyte growth, survival, and function.

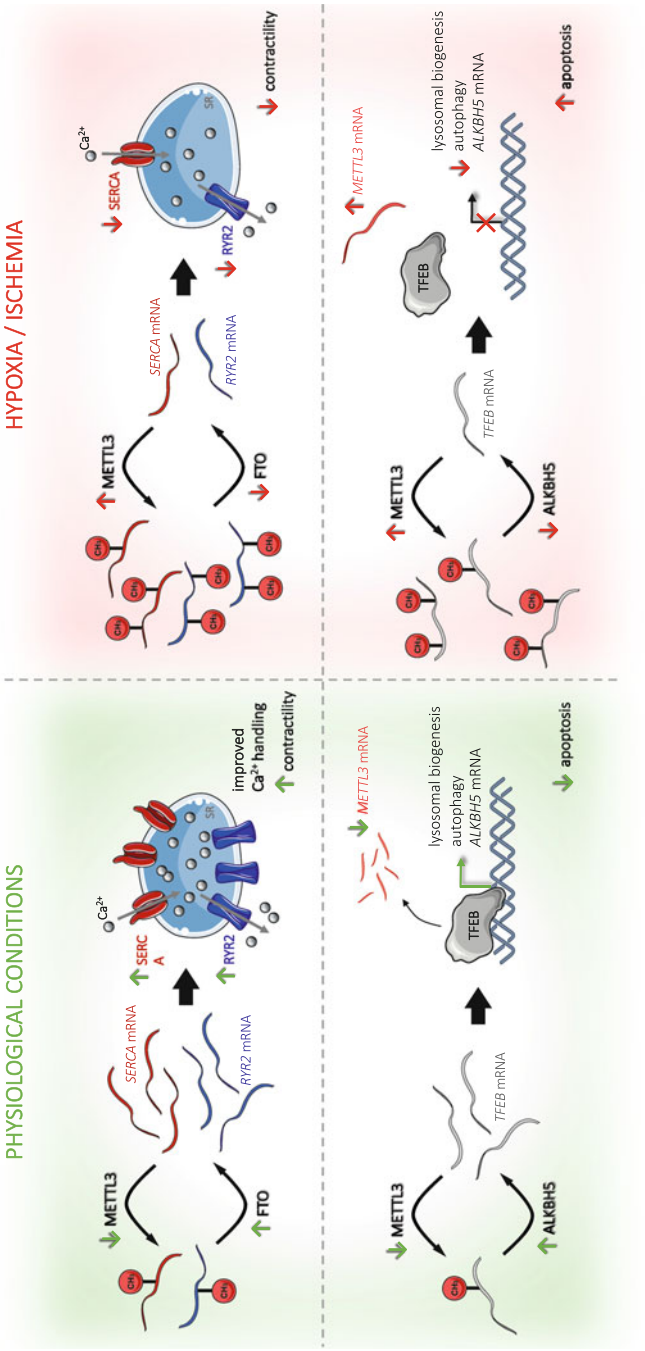


Fig. 2 Molecular mechanisms of m^{6}A-mediated regulation during cardiac ischemia. Reduced levels of FTO in ischemic failing hearts contribute to increased methylation of *SERCA* and *RYR2* transcripts. This leads to their decreased stability and reduced protein expression, resulting in aggravated Ca^{2+} cycling and impaired cardiac contractility (upper panels). SR: sarcoplasmic reticulum. Hypoxia-treated cardiomyocytes display increased levels of methylated *TFEB* mRNA, which reduces its stability and decreases protein expression. This leads to a reduction in lysosomal biogenesis and autophagy, thereby causing apoptosis. TFEB also activates the transcription of demethylase ALKBH5 and decreases the stability of *METTL3* mRNA, thus establishing two feedback loops (bottom panels)

7 m^6A in Non-cardiomyocyte Cells Important for the Cardiovascular System

Fibrosis is a common characteristic of various cardiovascular diseases including heart failure caused by myocardial infarctions. It is caused by irregular proliferation and migration of fibroblasts and the extensive deposition of extracellular matrix, especially in the injured area of the cardiac tissue. Cardiac fibrotic tissue contributes to the malfunction of failing hearts since its inability to contract or conduct electric impulses needed for the transfer of the contraction makes the heart stiffer and reduces its function. Mettl3 is highly upregulated in fibrotic tissue of hearts after myocardial infarction, and this is reduced by knocking down Mettl3 using a lentiviral system containing Mettl3 siRNA (Li et al. 2021).

Increased levels of m^6A mRNA are also found in pathological calcified arteries, a condition resulting in arterial elasticity and impairment of cardiovascular hemodynamics. However, this was rather attributed to induction of augmented Mettl14 and then Mettl3. Hypermethylation of vascular osteogenic transcripts led to impaired vascular repair (Chen et al. 2019).

Additionally, Mettl3 regulates the ability of adipose-derived stem cells to differentiate into vascular smooth muscle cells. Mettl3 is upregulated in differentiating cells, and knockdown of Mettl3 leads to less m^6A residues on smooth muscle marker transcripts resulting in their reduced expression (Lin et al. 2020).

m^6A mRNA levels are also increased in abdominal aortic aneurysm in comparison to healthy aortic tissue and the levels positively correlate to the risk of rupture of the aneurysm (He et al. 2019).

8 Conclusions and Outlook

This work highlights the post-transcriptional regulatory role of m^6A on gene expression in cardiac cells in response to different pathological stimuli. Consequently, mRNA modifications like m^6A represent an additional layer of controlling gene expression that might be targeted in order to prevent, treat, or even reverse heart failure and other heart diseases.

Up to this point, all published studies reported overall increased levels of m^6A in hypertrophying, ischemic, or otherwise stressed cardiomyocytes and hearts. Hence, the use of m^6A levels as a biomarker for various pathological cardiac conditions seems to be promising. However, this would only be advantageous if an invasive procedure for the assessment of the myocardial sample could be avoided. Hence, only in the case of m^6A-containing RNAs being released into the circulation in a disease-specific manner, m^6A levels would be conceivable biomarkers. So far, for cardiac diseases mainly circulating noncoding RNAs have been studied for the usage as suitable biomarker (Viereck and Thum 2017).

Furthermore, single transcripts could be targeted, either by inducing methylation or demethylation leading to changes of the transcript's stability or translation efficiency. Mechanistically, this could be approached by applying a programmable and nuclease-inactive Cas13 family system (Abudayyeh et al. 2017; Cox et al. 2017), fused to the functional domains of Mettl3 or Fto. This way, the site-directed RNA binding function of the Cas protein and its guide-RNA is combined with the functional effect of the fusion protein, leading to methylation or demethylation of the targeted mRNA. In the same way, m^6A-reader proteins could be directed to an mRNA of interest directly inducing decay or enhancing translation without the need of an m^6A site (Rauch et al. 2018). The use of this tool in cardiomyocytes could enable the manipulation of single RNAs in cardiac remodeling and represents a future tool as therapy but needs to be tested in future experiments.

Also, transcript methylation could be inhibited by base editing of the "methylatable" adenosine applying the adenosine deaminase acting on RNA (ADAR) class of enzymes. ADARs bind to double-stranded RNA and convert adenosine to inosine by deamination (Rees and Liu 2018). However, since inosine will be translated like a guanosine, in addition to the inhibition of methylation a change of the sequence takes place. ADAR can be directed toward adenosines that are naturally not being edited by an approach called site-directed RNA editing (SDRE) either by redirecting the endogenous ADAR to the target transcript using a complementary RNA oligonucleotide as type of gRNA or by using genetically engineered ADAR systems, which comprise the catalytic domain of ADAR fused to a protein domain that will lead the targeting process (Montiel-Gonzalez et al. 2019).

Manipulating and controlling protein expression at the mRNA level is advantageous and profitable as there is no interference with the genetic background and it bypasses the need to target two or more alleles (in case of bi-nucleated cardiomyocytes). Hence, the application of these tools in cardiomyocytes enables the manipulation of single RNAs in cardiac remodeling and represents a future tool as therapy.

Acknowledgments M.V, E.M., and V.K would like to thank members of the Völkers lab for stimulating discussions on the epitranscriptome of cardiomyocytes. The work of M.V. was kindly supported by the Deutsche Forschungsgemeinschaft, Baden-Württemberg Stiftung, and the Boehringer Ingelheim Foundation. V.K and M.V. acknowledge funding from the German Center of Cardiovascular Research (DZHK), Partner Site Heidelberg/Mannheim.

References

Abudayyeh OO, Gootenberg JS, Essletzbichler P et al (2017) RNA targeting with CRISPR-Cas13. Nature 550:280–284

Berulava T, Buchholz E, Elerdashvili V et al (2020) Changes in m6A RNA methylation contribute to heart failure progression by modulating translation. Eur J Heart Fail 22:54–66

Carnevali L, Graiani G, Rossi S et al (2014) Signs of cardiac autonomic imbalance and proarrhythmic remodeling in FTO deficient mice. PLoS One 9:e95499

Chen L, Sampson KJ, Kass RS (2016) Cardiac delayed rectifier potassium channels in health and disease. Card Electrophysiol Clin 8:307–322
Chen J, Ning Y, Zhang H et al (2019) METTL14-dependent m6A regulates vascular calcification induced by indoxyl sulfate. Life Sci 239:117034
Cohn JN, Ferrari, Sharpe N (2000) Cardiac remodeling—concepts and clinical implications: a consensus paper from an international forum on cardiac remodeling. J Am College Cardiol 35:569–582
Cox DBT, Gootenberg JS, Abudayyeh OO et al (2017) RNA editing with CRISPR-Cas13. Science 358:1019–1027
deAlmeida AC, Van Oort R, Wehrens XHT (2010) Transverse aortic constriction in mice. J Vis Exp 38:1729
Dewenter M, Von Der Lieth A, Katus HA et al (2017) Calcium signaling and transcriptional regulation in cardiomyocytes. Circ Res 121:1000–1020
Dirkx E, da Costa Martins PA, de Windt LJ (2013) Regulation of fetal gene expression in heart failure. Biochim Biophys Acta 1832:2414–2424
Dorn LE, Lasman L, Chen J et al (2019) The N(6)-methyladenosine mRNA methylase METTL3 controls cardiac homeostasis and hypertrophy. Circulation 139:533–545
Frey N, Olson EN (2003) Cardiac hypertrophy: the good, the bad, and the ugly. Annu Rev Physiol 65:45–79
Fu Y, Dominissini D, Rechavi G et al (2014) Gene expression regulation mediated through reversible m6A RNA methylation. Nat Rev Genet 15:293
Han P, Li W, Lin CH et al (2014) A long noncoding RNA protects the heart from pathological hypertrophy. Nature 514:102–106
He Y, Xing J, Wang S et al (2019) Increased m6A methylation level is associated with the progression of human abdominal aortic aneurysm. Ann Transl Med 7:797
Ke S, Alemu EA, Mertens C et al (2015) A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Gen Dev 29:2037–2053
Kmietczyk V, Riechert E, Kalinski L et al (2019) m6A-mRNA methylation regulates cardiac gene expression and cellular growth. Life Sci Alliance 2:e201800233
Li T, Zhuang Y, Yang W et al (2021) Silencing of METTL3 attenuates cardiac fibrosis induced by myocardial infarction via inhibiting the activation of cardiac fibroblasts. FASEB J 35:e21162
Lin J, Zhu Q, Huang J et al (2020) Hypoxia promotes vascular smooth muscle cell (VSMC) differentiation of adipose-derived stem cell (ADSC) by regulating Mettl3 and paracrine factors. Stem Cells Int 2020:2830565
Mathiyalagan P, Adamiak M, Mayourian J et al (2018) FTO-dependent m6A regulates cardiac function during remodeling and repair. Circulation 139:518–532
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149:1635–1646
Mill JG, Stefanon I, Dos Santos L et al (2011) Remodeling in the ischemic heart: the stepwise progression for heart failure. Braz J Med Biol Res 44:890–898
Montiel-Gonzalez MF, Quiroz JFD, Rosenthal JJC (2019) Current strategies for site-directed RNA editing using ADARs. Methods 156:16–24
Nag AC (1980) Study of non-muscle cells of the adult mammalian heart: a fine structural analysis and distribution. Cytobios 28:41–61
Nagai R, Low RB, Stirewalt WS et al (1988) Efficiency and capacity of protein synthesis are increased in pressure overload cardiac hypertrophy. Am J Phys 255:H325–H328
Rauch S, He C, Dickinson BC (2018) Targeted m(6)A reader proteins to study epitranscriptomic regulation of single RNAs. J Am Chem Soc 140:11974–11981
Rees HA, Liu DR (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. Nat Rev Genet 19:770–788
Roth GA, Abate D, Abate KH et al (2018) Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the global burden of disease study 2017. Lancet 392:1736–1788

Sequeira V, Nijenkamp LL, Regan JA et al (2014) The physiological role of cardiac cytoskeleton and its alterations in heart failure. Biochim Biophys Acta 1838:700–722

Song H, Feng X, Zhang H et al (2019) METTL3 and ALKBH5 oppositely regulate m(6)A modification of TFEB mRNA, which dictates the fate of hypoxia/reoxygenation-treated cardiomyocytes. Autophagy 1–19

Spruill LS, Baicu CF, Zile MR et al (2008) Selective translation of mRNAs in the left ventricular myocardium of the mouse in response to acute pressure overload. J Mol Cell Cardiol 44:69–75

Tan FL, Moravec CS, Li J et al (2002) The gene expression fingerprint of human heart failure. Proc Natl Acad Sci USA 99:11387–11392

Tang R, Long T, Lu KO, a. (2020) A roadmap for fixing the heart: RNA regulatory networks in cardiac disease. Mol Ther – Nucleic Acids 20:673–686

Van Heesch S, Witte F, Schneider-Lunitz V et al (2019) The translational landscape of the human heart. Cell 178:242–260.e29

Viereck J, Thum T (2017) Circulating noncoding RNAs as biomarkers of cardiovascular disease and injury. Circ Res 120:381–399

Völkers M, Tok H, Doroudgar S et al (2013) Pathological hypertrophy amelioration by PRAS40-mediated inhibition of mTORC1. Proc Natl Acad Sci USA 110:12661

Wang K, Liu F, Zhou LY et al (2014a) The long noncoding RNA CHRF regulates cardiac hypertrophy by targeting miR-489. Circ Res 114:1377–1388

Wang Y, Li Y, Toth JI et al (2014b) N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nat Cell Biol 16:191–198

WHO (1992) International classification of diseases and related health problems (10th revision). World Health Organization, Geneva

Zaccara S, Jaffrey SR (2020) A unified model for the function of YTHDF proteins in regulating m6A-modified mRNA. Cell 181:1582–1595.e18

From m^{6}A to Cap-Adjacent m^{6}Am and their Effects on mRNAs

Trinh T. Tat and Daniel L. Kiss

Contents

Abstract Although RNA modifications were discovered decades ago, the identification of enzymes that write, read, and erase RNA modifications enabled their functional study and spawned the field of epitranscriptomics. Coupling that

T. T. Tat
Center for RNA Therapeutics, Department of Cardiovascular Sciences, Houston Methodist Academic Institute, Houston, TX, USA

D. L. Kiss (✉)
Center for RNA Therapeutics, Department of Cardiovascular Sciences, Houston Methodist Academic Institute, Houston, TX, USA

Houston Methodist Cancer Center, Houston Methodist Research Institute, Houston, TX, USA

Weil Cornell Medical College, New York, NY, USA
e-mail: dlkiss@houstonmethodist.org

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_12

knowledge to new methods has enabled the precise pinpointing of epitranscriptomic modifications across the transcriptome and the elucidation of their functional consequences. PCIF1 (*P*hosphorylated *C*TD-*I*nteracting *F*actor *1*) was shown to add N6, 2′-*O*-dimethyladenosine (m^6Am) marks at the first nucleotide after the 5′ N7-methylguanosine (m^7G) cap. In this chapter, we discuss the epitranscriptomic regulation of mRNA in general and focus on m^7G cap-adjacent m^6Am in particular. m^6Am positions can now be distinguished from N6-methyladenosine (m^6A) using new techniques leveraging PCIF1-knockout cells. Although m^6Am modification sites can be detected precisely, conflicting data have been published regarding how cap-adjacent m^6Am marks affect their host mRNA. Discrepancies in the data mean that the effects of cap-adjacent m^6Am on mRNA stability, decapping, and translation continue to be debated. Finally, while PCIF1 is predominantly nuclear, a subset of results suggests a possible cytoplasmic role as well. Taken together, these contradictory results which employed different methodologies and cell lines mean that further experiments are required to determine the ultimate biological function (s) of m^7G cap-adjacent m^6Am.

Keywords Epitranscriptome · PCIF1 · N6 2′-*O*-dimethyladenosine · Cap-adjacent m^6Am · Cytoplasmic capping

1 The Epitranscriptome: Dynamic RNA Modifications that Regulate Gene Expression

1.1 An Introduction to Common RNA Modifications

The discovery of pseudouridine as the first structurally modified RNA nucleoside in the 1950s began over two decades of rapid advances where many chemically modified nucleotides were identified for the first time (Davis and Allen 1957). To date, about 160 distinct RNA modifications are catalogued in the MODOMICS database (http://genesilico.pl/modomics/) and map to many different types of cellular RNAs (Boccaletto et al. 2018; He 2010). Decades of data show that RNA modifications are common in *r*ibosomal *RNA* (rRNA), *t*ransfer *RNA* (tRNA), *s*mall *n*uclear *RNA* (snRNA), *s*mall *n*ucle*o*lar *RNA* (snoRNA), and *m*essenger *RNA* (mRNA) among other RNA types (Davis and Allen 1957; Goh et al. 2020; Grosjean et al. 1997). As tRNAs, snRNAs, snoRNAs, and rRNAs are both abundant and heavily (and specifically) modified in cells, RNA modifications were always of great interest in these fields (Bohnsack and Sloan 2018; Pan 2018). In fact, the proper modification of key nucleotides is critical to the functions of many of these *n*on-*c*oding *RNA*s (ncRNA). Compared to those ncRNAs, apart from the cap structures, mRNA base modifications were significantly understudied until ~10–15 years ago (He 2010; Saletore et al. 2012). To harness this growing interest, the name "RNA epigenetics" was suggested to describe the growing field focused on the study of

Table 1 Transcriptome-wide methods that target and map m^{6}A, m^{6}Am, and m^{7}G capping sites

RNA modification	Technique	References
All RNA modifications	Oxford Nanopore direct RNA sequencing	Viehweger et al. (2019)
m^{7}G	m^{7}G-MaP-seq	Enroth et al. (2019)
	AlkAniline-Seq	Marchand et al. (2018)
	Capped analysis of gene expression (CAGE)	Fejes-Toth et al. (2009)
m^{6}Am	m^{6}Am-Exo-Seq	Sendinc et al. (2019)
	m^{6}ACE-seq	Koh et al. (2019)
m^{6}A (most also detect m^{6}Am)	m^{6}A-seq	Dominissini et al. (2012)
	meRIP-seq	Meyer et al. (2012)
	m^{6}A-LAIC-seq	Molinie et al. (2016)
	miCLIP-seq	Hawley and Jaffrey (2019)
	PA-m^{6}A-seq	Chen et al. (2015)
	m^{6}A-CLIP-seq	Hsu and He (2019)
	m^{6}A-label-seq	Shu et al. (2020)
	m^{6}A-SEAL-seq	Wang et al. (2020)
	MAZTER-seq	Garcia-Campos et al. (2019)
	m^{6}A-REF-seq	Zhang et al. (2019b)
	DART-seq	Meyer (2019)
	meCLICK-seq	Mikutis et al. (2020)

RNA modifications (He 2010). Shortly thereafter, the field adopted epitranscriptomics as a more distinct RNA-focused identifier (Meyer et al. 2012; Saletore et al. 2012).

Several key advances have driven the growing interest in epitranscriptomics, mainly by removing barriers to their study. The first barrier fell when it was clear that non-cap-associated epitranscriptomic marks, including those on mRNAs, were both dynamic and had functional consequences (Roundtree et al. 2017). The continuing identification of the enzymes that added, interpreted, and removed epitranscriptomic marks also proved key (Bokar et al. 1997). Finally, the coupling of deep sequencing strategies with biochemical methods to purify modified RNAs yielded multiple methods that can recognize and pinpoint both the presence and prevalence of a diverse set of mRNA modifications (Anreiter et al. 2020). Table 1 presents the existing methodologies that target the three key RNA modifications in the focus of this chapter. Importantly, third-generation long-read sequencing technologies promise the next revolution in epitranscriptomics (Lorenz et al. 2020; Zhao et al. 2019). For example, RNA modifications can now be detected directly on their RNAs without using reverse transcription via direct RNA sequencing from Oxford Nanopore and similar methods are being designed to leverage PacBio long-read sequencing as well (Lorenz et al. 2020; Zhao et al. 2019).

1.2 Methylated RNA Bases

One of the most common family of RNA modifications is methylation, which is ubiquitous in life (Shi et al. 2019; Zaccara et al. 2019). In fact, according to the MODOMICS database, roughly 100 of the ~160 known modified RNA bases include at least one type of methylation event among the modifications (Boccaletto et al. 2018). RNA methylation predominantly occurs on nitrogen and carbon positions and/or amine groups outside the ring of purine and pyrimidine bases, plus the oxygen atom of the 2′-OH moiety of the ribose sugar (Boccaletto et al. 2018). Several types of methylated base modifications are common in eukaryotic mRNA. The m^7G (N7-methylguanosine) that constitutes the 5′ cap structure of mRNAs was among the first base modifications to be identified and characterized on mRNAs (Shatkin 1976). Besides the m^7G cap, m^6A (N6-methyladenosine) and m^6Am (N6,2′-*O*-dimethyladenosine) are two of the better characterized RNA methylation events and were also identified as abundant in mRNAs in the mid-1970s (Perry and Kelley 1974; Wei et al. 1975). The second of these, m^6Am, is common in the bodies of certain ncRNAs such as snRNAs and enriched directly adjacent to 5′ mRNA caps and imparts distinct functional properties to the mRNA (Keith et al. 1978; Wei et al. 1975).

Apart from the m^7G cap, m^6A, and m^6Am RNA modifications which will be covered in detail in Sects. 2–4, several other methylated RNA bases are common (Boccaletto et al. 2018). These include m^5C (5-methylcytosine), m^1A (N1-methyladenosine), $m^{6,6}A$ (N6, N6-dimethyladenosine), hm^5C (5-hydroxymethylcytosine), and the TMG (N2, N2, N7 trimethylguanosine) cap among many others (Boccaletto et al. 2018). These epitranscriptomic marks are known to play vital roles in altering RNA–protein interactions and RNA secondary structures and causing changes in RNA stability and/or translation efficiency (Shi et al. 2019; Zaccara et al. 2019). The TMG cap is found on snRNAs, snoRNAs, and certain other ncRNAs (Lamond 1990). m^1A modifications are found mainly in tRNAs, mRNAs, *l*ong *n*on-*c*oding *RNAs* (lncRNAs), and mitochondrial genes (Dominissini and Rechavi 2017). In 2017, m^1A was mapped near the *t*ranscription *s*tart *s*ites (TSS) and the first splice site in coding sequences and shown to increase translation efficiency through enabling the noncanonical binding of the exon–exon junction complex at the 5′ *unt*ranslated *r*egion (UTR) (Dominissini and Rechavi 2017). In addition, the methyl group of m^1A is known to block Watson-Crick base pairing and effectively terminates reverse transcription and disrupts translation (Dominissini and Rechavi 2017). Similar to m^1A, m^5C sites are mapped in human mRNA and lncRNA species; however, m^5C sites are mainly enriched in the 5′ UTRs before translation initiation sites, and in close proximity to the translation stop codon (Dong and Cui 2020; Roundtree et al. 2017; Squires et al. 2012). Changes in the level of NSUN2, a key m^5C methyltransferase, have been shown to strongly affect RNA metabolism and are linked to various human neurodegenerative diseases and cancers (Frye and Watt 2006; Hussain et al. 2013).

1.3 Focus and Scope of the Paper

Taken together, the abundance, sequence context, and chemical structures of RNA modifications create the epitranscriptomic landscape which can drive both molecular and cellular dynamics. We are now beginning to better understand key modifications in epitranscriptome and have begun unraveling their regulatory roles in biological processes of cells. Further, advances are continuously providing new precise, sensitive, and quantitative experimental and computational techniques to identify, pinpoint, and map individual epitranscriptomic modifications with single base resolution (Anreiter et al. 2020).

In this chapter, we focus on three RNA modifications: the m^{7}G cap, m^{6}A, and m^{6}Am and their effects on mRNA half-life and translation. We compare and contrast the "knowns" and "unknowns" regarding m^{6}A and m^{6}Am in particular. Table 1 lists the common techniques that are used to target the three epitranscriptomic marks described below (Anreiter et al. 2020). As a detailed description of these methodologies is beyond the scope of this chapter, please see these recent comprehensive reviews for more information (Anreiter et al. 2020). Finally, as this chapter focuses on m^{7}G cap-adjacent m^{6}Am marks on mRNAs and internal m^{6}Am marks are well documented for U2 snRNA and can be added to certain mRNAs under certain conditions, we will abbreviate m^{7}G *c*ap-*a*djacent *m^{6}Am* as CA-m^{6}Am hereafter (Goh et al. 2020; Mauer et al. 2019).

2 The m^{7}G Cap and Its Role in the Regulation of mRNAs

Likely because of its presence at the 5′ end of every RNA polymerase II-transcribed mRNA, the m^{7}G cap structure (Fig. 1) was among the first RNA modifications with a clearly defined function (Shatkin 1976). The *RN*A *g*uanylyl*t*ransferase and 5-′-*t*riphosphatase (RNGTT) uses a two-step process to add an inverted guanosine residue to the initiating nucleotide of the nascent mRNA via a 5′-5′ triphosphate linkage (Ensinger and Moss 1976; Shatkin 1976). This occurs co-transcriptionally in the nucleus as the nascent RNA is extruded from RNA Polymerase II as it transcribes mRNAs (Ramanathan et al. 2016). The final step of cap maturation occurs when *RN*A guanine-7 *m*ethyl*t*ransferase (RNMT), dimerized with *R*NMT-*a*ssociated *m*iniprotein (RAM), adds a methylation onto the N7 position of the inverted guanosine to complete the m^{7}G cap (Fig. 1, blue) (Ramanathan et al. 2016). This methyl group is a crucial feature, and protects the mRNA from degradation and enhances mRNA translation (Ramanathan et al. 2016). Notably, studies in the past decade have demonstrated that functional pools of RNGTT and RNMT are present in the cytoplasm and that a subset of uncapped human mRNAs can also be capped and methylated in the cytoplasm (Keith et al. 1978; Otsuka et al. 2009; Trotman et al. 2017).

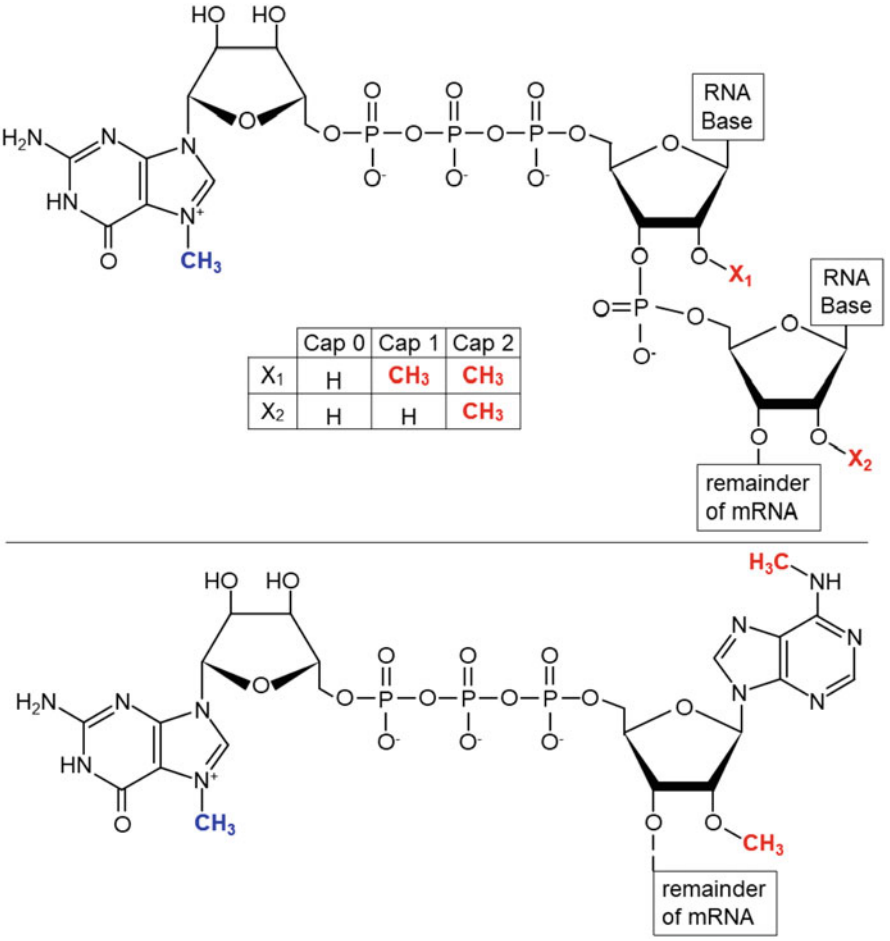


Fig. 1 A diagram showing the chemical structures of different cap structures observed in eukaryotes. (Top) Critical features of the cap structure include the methylation on the m^7G-cap (Blue) and possible methylations on the first (X_1) and second (X_2) transcribed nucleotides (Red). Cap 0 RNAs lack methyl groups at both X_1 and X_2, Cap 1 RNAs have methyl groups on X_1, but not X_2, while Cap 2 RNAs have methyl groups on both nucleotides (inset table). (Bottom) Diagram showing the chemical structure of cap-adjacent m^6Am. Key methylation events within the cap and the first transcribed nucleotide are highlighted in blue and red, respectively

Other early works also demonstrated that in addition to the m^7G cap, one or both of the first two transcribed nucleotides of an mRNA also modified in some organisms (Crain et al. 1978; Ensinger and Moss 1976). Together with the m^7G cap mRNAs were said to have Cap 0, Cap 1, or Cap 2 (Fig. 1) depending upon whether zero, one, or two transcribed RNA bases were methylated (Ramanathan et al. 2016; Wei et al. 1975). These methylations at the 2′ position on the ribose sugar of the first transcribed nucleotide are added in the nucleus by the actions of mRNA cap 2′-*O*-

methyltransferase, the first of which was identified in vaccinia virus (Barbosa and Moss 1978). In humans, the final methylation to complete Cap 2 structures is added in the cytoplasm by hMTr2 (Werner et al. 2011). The prevalence of these distinct mRNA cap structures depends on the organism, but in general, Cap 0 structures are present in lower eukaryotes, while Cap 1 and Cap 2 structures are more prevalent in more advanced eukaryotes (Ramanathan et al. 2016). Notably, different organisms such as trypanosomes often generate hypermethylated Cap 3 and Cap 4 structures where the third and fourth bases of their mRNAs are also methylated (Reolon et al. 2019). Cap 0 structures are essential to protect the mRNA from nucleases and are also required to enable efficient translation of mRNAs (Ghosh and Lima 2010; Ramanathan et al. 2016). Cap 1 and Cap 2 structures have been shown to be critical in designating an mRNA as "self" to escape the cellular innate immune response in humans (Leung and Amarasinghe 2016).

3 N6-Methyladenosine (m^{6}A)

As discussed above, N6-methyladenosine (m^{6}A) and N6, 2′-*O*-dimethyladenosine (m^{6}Am) are comparatively abundant RNA modifications (Fig. 2) in polyadenylated (poly(A)) mRNAs. Early works using P^{32}-labeled cellular RNA, nucleases, and thin-layer chromatography showed that m^{6}A was the most abundant internal mRNA modification and estimated that m^{6}A comprised ~0.125% of all bases in poly (A) mRNA (Perry and Kelley 1974). Those data were bolstered as m^{6}A was determined to comprise roughly one in every ~800 nucleotides in poly(A)-selected RNA species from both the cytoplasm and the nucleus (Lavi et al. 1977). They also showed that m^{6}A occurs roughly once in every 1800–3000 nucleotides, in non-polyadenylated, non-ribosomal RNAs (Lavi et al. 1977). The first consensus sequence motif candidates for m^{6}A addition were identified when ~70% of m^{6}A

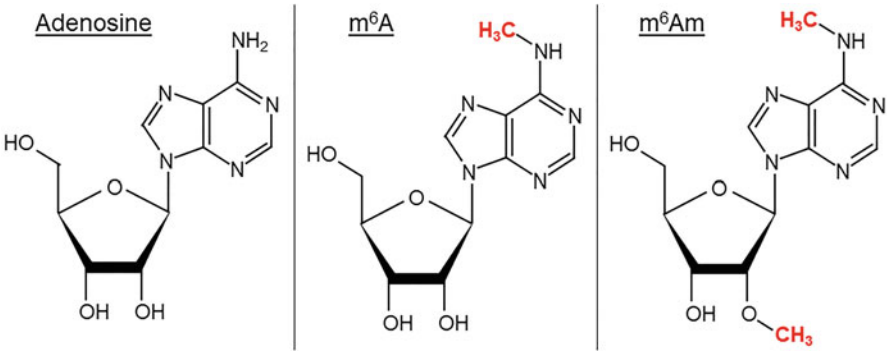


Fig. 2 Diagrams showing the chemical structures of adenosine, m^{6}A, and m^{6}Am. The line drawings showing the differences between unmodified adenosine (left), m^{6}A (middle), and m^{6}Am (right) are shown. The added methyl groups are highlighted in red on modified bases

modifications were shown to occur in the context of G(m^6A)C trinucleotides and the remaining 30% occurred in A(m^6A)C trinucleotides (Shi et al. 2019; Wei et al. 1976; Wei and Moss 1977). Finally, the increased prevalence of m^6A with a particular mRNA correlated with RNA instability (Sommer et al. 1978). Although the identities, relative frequency, sequence context, and general effect of m^6A mRNA modifications were known since the 1970s, they remained difficult to study as methods to definitively map their positions were limited to the extreme 5′ ends of mRNAs.

Advances in high-throughput sequencing technologies coupled with the advent of new biochemical reagents that target m^6A bases have allowed many groups to revisit and expand upon these early estimates. These methods (Table 1) now estimate that m^6A comprises about 0.2%–0.6% of all adenosines in mammalian mRNAs (Anreiter et al. 2020; Linder and Jaffrey 2019). Furthermore, they can provide a degree of certainty, with some methods offering single base resolution, as to where these mRNA modifications occur in the mRNA (Anreiter et al. 2020; Linder and Jaffrey 2019). Transcriptome-wide studies have convincingly shown that m^6A was enriched both near the stop codon and in 3′ UTRs of mammalian mRNAs (Dominissini et al. 2012; Meyer et al. 2012). Despite this progress, new methods which can more precisely verify the presence and positioning of m^6A modifications will continue to be in high demand.

The most consequential advances to define the function(s) of m^6A in vivo were made when the enzymes involved in adding and surveying m^6A were identified and characterized (Bokar et al. 1997). The cellular factors that place, interpret, and remove epitranscriptomic marks are generally referred to as writers, readers, and erasers, respectively. In this chapter, we discuss the effectors including writers, readers, and erasers of m^6A, and m^6Am.

3.1 m^6A Writers

Initially named MT-A, *met*hyl*t*ransferase-*l*ike protein *3* (METTL3) was the first m^6A writer to be identified (Bokar et al. 1997). Before the identification and cloning of METTL3, previous works had demonstrated that METTL3 was part of a multiprotein complex (Bokar et al. 1994). In fact, the efforts of multiple groups have shown that the m^6A methyltransferase complex consisting of METTL3, *met*hyl*t*ransferase-*l*ike protein *14* (METTL14), *W*ilms *t*umor 1-*a*ssociated *p*rotein (WTAP), *Vir*-like m^6A *M*ethyltransferase *A*ssociated (VIRMA, also called KIAA1429), and *R*NA *B*inding *M*otif Protein *15* (RBM15)/RBM15B is responsible for depositing m^6A in a co-transcriptional manner (Shi et al. 2019; Zaccara et al. 2019).

The majority of m^6A mRNA methylations are situated co-transcriptionally by methyltransferase writer complexes in a DRACH (D = A, G, or U, R = A or G, H = A, C, or U) sequence context (Edupuganti et al. 2017; Huang et al. 2018). Although METTL3 contains a nuclear localization signal (NLS), it is distributed distinctly among different cell lines (Xiang et al. 2020). METTL3 localizes

predominantly within the nucleus, with a visible enrichment in nuclear speckles where it interacts with WTAP to form a stable dimer with METTL14 in HeLa cells (Shi et al. 2019). A fraction of METTL3 is associated with the promoter regions of ~80 active genes marked by CEBPZ, independent of METTL14, suggesting a transcript-specific m^{6}A methylation activity (Barbieri et al. 2017). The recruitment of METTL3 to discrete chromatin loci in response to stress is dynamic, possibly via the action of epigenetic marks and/or transcription factors (Engel and Chen 2018). Furthermore, H3K36me3, a gene-body enriched histone modification, was shown to recruit METTL3 through interactions with METTL14 to deposit m^{6}A predominantly within mRNA open reading frames and 3′ UTRs (Huang et al. 2019).

Although the majority of METTL3 is found in the nucleus, it has been detected in the cytoplasm of several human cell lines and any feasible cytoplasmic function (s) remains unknown (Scholler et al. 2018). One possibility is that post-translational modifications change the interactions between METTL3 and its interactome leading to METTL3's cytoplasmic localization (Scholler et al. 2018). Cytoplasmic METTL3 may not be an m^{6}A writer, but rather functions as an m^{6}A reader (Lin et al. 2016). Using lung cancer cells, cytoplasmic METTL3 promoted the translation of a reporter mRNA when tethered to its 3′ UTR (Lin et al. 2016). Through post-translational modifications (such as SUMOylation) and interactions with other associated proteins, METTL3 could affect protein instability, localization, and the formation and catalytic activity of m^{6}A writer complexes (Shi et al. 2019; Zaccara et al. 2019).

Another m^{6}A writer, *met*hyl*t*ransferase-*l*ike protein *16* (METTL16), has a more restricted list of substrates including the hairpin (hp1) in the 3′ UTR of human *m*ethionine *a*denosyl*t*ransferase *2A* mRNA (*MAT2A*) that encodes the S-adenosylmethionine synthetase and the U6 snRNA (Pendleton et al. 2017; Ruszkowska et al. 2018). As with METTL3, at least a portion of METTL16 protein localizes to the cytoplasm (Nance et al. 2020). In addition, Ma et al. recently showed that ZCCHC4 is m^{6}A writer that methylates the A4220 on 28S rRNA, as well as interacts with a small group of mRNAs (Ma et al. 2019).

3.2 *m^{6}A Readers*

Several methods including the immunoprecipitation or pull down of methylated probes and quantitative protein mass spectrometry have been used to identify multiple m^{6}A readers (Linder and Jaffrey 2019). The first family of m^{6}A reader proteins contain *YT*521-B *h*omology (YTH) domains, including the *YTH d*omain *f*amily *1–3* (YTHDF1–3) and *YTH d*omain *c*ontaining *1–2* (YTHDC1–2) proteins in humans (Stoilov et al. 2002; Xu et al. 2014). Although belonging to the same broader protein family, several YTH domain-containing proteins have opposing effects when they recognize mRNAs with m^{6}A marks (Linder and Jaffrey 2019; Shi et al. 2019). For example, cytoplasmic YTHDF2 promotes mRNA deadenylation and degradation by recruiting deadenylase complexes (Roundtree et al. 2017). Two other m^{6}A readers, YTHDF1 and YTHDF3, promote the

translation of m^{6}A-containing mRNAs by recruiting translation initiation factors in HeLa cells (Shi et al. 2017; Xu et al. 2014). YTHDC2 also regulates both mRNA stability and translation, in addition to playing an important role in spermatogenesis (Huang et al. 2020). Finally, YTHDC1 localizes to the nucleus and helps regulate mRNA splicing, promotes mRNA export, and accelerates the decay of certain transcripts (Wang et al. 2015).

Another group of m^{6}A readers have common *R*NA *b*inding *d*omains (RBDs) such as a*r*ginine/*gly*cine-rich (RGG) domains, *R*NA *r*ecognition *m*otifs (RRM), and *K h*omology (KH) domains, to preferentially bind m^{6}A-containing RNAs (Liu et al. 2015). Having one RGG domain and three KH domains, *F*ragile X *m*ental *r*etardation *1* (FMR1) recruits YTHDF2 to affect the translation and stability of m^{6}A-containing mRNAs (Edupuganti et al. 2017). Several other m^{6}A readers such as *i*nsulin-like *g*rowth *f*actor *2* mRNA-binding *p*roteins *1–3* (IGF2BP1–3) or *pr*oline-*r*ich *c*oiled-coil *2A* (Prrc2a) have been reported to recognize and stabilize m^{6}A-bearing mRNAs (Huang et al. 2018). Multiple *h*eterogeneous *n*uclear *r*ibonucleo*p*roteins (HNRNP), including HNRNPC, HNRNPG, and HNRNPA2B1, are known to recognize and preferentially bind m^{6}A-containing ncRNAs in the nucleus (Shi et al. 2019; Zaccara et al. 2019). m^{6}A readers promote translation and/or alter mRNA stability depending on specific cellular contexts such as heat shock, viral infection, or other stresses (Shi et al. 2019; Zaccara et al. 2019).

Multiple studies have shown cross talk or competition between proteins that read m^{6}A marks (Panneerdoss et al. 2018). Reader proteins may also localize to specific subcellular compartments by interacting with other RNAs or RNA binding proteins. Several reader proteins YTHDF1–3, FMR1, and HNRNPA2B1 were found in the cores of mammalian stress granules while IGF2BP2–3 and HNRNPK were enriched in the protrusions of breast cancer cells (Mardakheh et al. 2015). Taken together, m^{6}A reader proteins comprise a network of physical and/or functional interactions that regulate the translation efficiency and stability of m^{6}A-bearing mRNAs in a context-dependent manner.

3.3 m^{6}A Erasers

Internal m^{6}A can be removed by one of two known demethylases FTO (*fat* mass and *o*besity-associated protein) and *AlkB h*omolog *5* (ALKBH5) (Jia et al. 2011; Zheng et al. 2013). The demethylase activity of both FTO and ALKBH5 serves to erase m^{6}A marks on RNAs (Jia et al. 2011; Zheng et al. 2013). Similar to the readers, most erasers also work in a context-dependent manner. FTO was the first enzyme shown to remove the methyl groups from m^{6}A in mRNA both in vitro and in vivo (Jia et al. 2011; Shi et al. 2019; Zaccara et al. 2019). In addition, using *c*ross-*l*inking *i*mmuno*p*recipitation followed by high-throughput *seq*uencing (CLIP-Seq), FTO has been demonstrated to demethylate CA-m^{6}Am (Mauer et al. 2017; Wei et al. 2018). FTO was established as an m^{6}A demethylase by a combination of cell culture-based assays that noted small changes in overall m^{6}A levels and experiments that showed

purified and/or recombinant FTO could demethylate m^{6}A RNA in vitro (Jia et al. 2011; Zou et al. 2016). FTO CLIP-Seq data from multiple cell lines also revealed that GAC- and/or GGAC-containing sequence motifs are significantly enriched in FTO-binding sites (Li et al. 2019).

Recently, the consensus that FTO is a dynamic m^{6}A demethylase has come under increased scrutiny (Meyer and Jaffrey 2017; Mauer and Jaffrey 2018). Me-RIP-Seq using material from FTO$^{-/-}$ mice showed that although a subset of m^{6}A-containing mRNAs showed changes, the global m^{6}A levels were essentially unchanged in these mice (Hess et al. 2013). Subsequent work supported this finding as m^{6}A consensus sequences were under-represented in mRNAs that were purified with CLIP experiments targeting FTO (Bartosovic et al. 2017). Together those data contradict the idea of FTO as an important m^{6}A demethylase (Bartosovic et al. 2017; Hess et al. 2013). FTO's role as an m^{6}A demethylase was further called into question when MATZER-seq studies showed little change in global m^{6}A in response to FTO depletion or overexpression (Garcia-Campos et al. 2019). Finally, in vitro assays showed that FTO strongly preferred m^{6}Am (and CA-m^{6}Am in particular) as a substrate rather than m^{6}A (Mauer et al. 2017). Collectively, these findings challenge the established model where FTO is an m^{6}A demethylase in vivo (Darnell et al. 2018; Mauer and Jaffrey 2018).

While it remains an open question, a substantial body of evidence does support a more limited or context-specific role for FTO as an m^{6}A demethylase. The conflict could possibly be explained, at least in part, by the compartmentalization of FTO activity. For example, the demethylation of internal m^{6}A mRNA and CA-m^{6}Am takes place in the cytoplasm, while the majority of m^{6}A removal happens in the nucleus (Wei et al. 2018). This interpretation is reasonable as FTO is predominantly a nuclear protein, although it does localize both to the nucleus and to the cytoplasm in certain cell lines (Aas et al. 2017; Jia et al. 2011). A crystal structure of human FTO with a 6 mA-modified single-stranded DNA bound in its active site provided additional mechanistic insights regarding FTO activity (Zhang et al. 2019a). Further modeling of the FTO crystal structure coupled to directed point mutations showed the mechanism by which FTO could demethylate both m^{6}A and m^{6}Am (Zhang et al. 2019a). They also demonstrated that both the sequence and secondary structure contexts of the m^{6}A modification are key determinants of FTO activity (Zhang et al. 2019a).

Another possible resolution to this controversy is that FTO works in concert with other proteins to mediate its m^{6}A demethylase activity (Song et al. 2020). Using cross-linking IP coupled to mass spectrometry FTO was shown to interact with over a dozen proteins, including six known RNA binding proteins such as *S*plicing *F*actor *P*roline and *Glutamine* Rich (SFPQ) (Song et al. 2020). RNA is hypomethylated in the vicinity of SFPQ binding sites and FTO to RNA interactions were greatly enriched near SFPQ binding sites (Song et al. 2020). The idea that FTO could be recruited near internal m^{6}A sites by an RNA binding protein could explain how FTO could still recognize and demethylate m^{6}A despite the enzyme's ~10-fold preference for CA-m^{6}Am (Mauer et al. 2017; Song et al. 2020).

3.4 *The Effects of m^{6}A on mRNA*

Numerous studies showed that mammalian m^{6}A modifications are highly regulated and have profound effects on the cellular heat-shock response, stem cell proliferation and differentiation, the DNA damage response, and tumorigenesis (Shi et al. 2019; Zaccara et al. 2019). The first evidence of m^{6}A causing mRNA instability was obtained using radioisotope metabolic labeling (Sommer et al. 1978). By comparing the half-lives of two populations of mRNAs (with and without m^{6}A) m^{6}A inclusion was demonstrated to prominently decrease mRNA half-lives in HeLa cells (Sommer et al. 1978). In addition, depletion of METTL3, m^{6}A writer, resulted in the increase of mRNA stability of m^{6}A-modified mRNAs in the cytoplasm (Ke et al. 2017). Multiple studies have shown that m^{6}A does not alter the steady-state level of cytoplasmic mRNAs; however, it serves as an imprint to mark the short half-life transcripts when they reach the cytoplasm (Barbieri et al. 2017; Darnell et al. 2018).

m^{6}A facilitates translation via different mechanisms. m^{6}A was reported to modulate mRNA translation efficiency through interactions between an m^{6}A reader, YTHDF1, and *e*ukaryotic translation *i*nitiation *f*actor *3* (eIF3) which then recruits the small ribosomal subunit to mRNAs (Wang et al. 2015). In addition, m^{6}A within the 5′ UTRs of stress and heat-shock protein-coding mRNAs can directly bind to eIF3, bypassing the normal requirement of *e*ukaryotic translation *i*nitiation *f*actor *4E* (eIF4E), and potentially enhance their translation during stress (Meyer et al. 2015). The third mechanism involves the interaction between METTL3, eIF3, and mRNA cap-associated proteins present in the cytosol. These interactions may allow ribosomes paused at stop codons to reload onto the 5′ UTR of transcripts while mRNAs are being translated (Lin et al. 2016).

When m^{6}A demethylases such as FTO and Alkbh5 were identified, the precise modification sites of m^{6}A as well as their biological functions were broadly revealed (Annapoorna et al. 2019; Mauer et al. 2019). A bird's-eye view of the m^{6}A epitranscriptomic landscape has become comprehensible and conclusively shows that m^{6}A is mainly distributed in the coding and 3′ untranslated regions with a significant enrichment just upstream of the stop codon (Koh et al. 2019; Wang et al. 2020). Therefore, the continued development of new, more sensitive technologies that can more precisely label, detect, and/or positionally pinpoint m^{6}A/m modifications are continuously in high demand (Anreiter et al. 2020; Mikutis et al. 2020).

4 Cap-Adjacent m^{6}Am

Unlike m^{6}A which is generally situated within the body of mRNAs, CA-m^{6}Am (Fig. 1, bottom) occurs at the first transcribed nucleotide of mRNAs (Hawley and Jaffrey 2019; Wang et al. 2019). The first attempts to identify the writer of CA-m^{6}Am took place shortly after the modification was discovered. The enzymatic activity responsible for adding a methyl group to the N6 position of cap-adjacent

2′-*O*-*m*ethyl*a*denosine (Am) to create the CA-m^{6}Am modification was isolated from fractionated HeLa cell extracts in the late 1970s (Keith et al. 1978). Their work further showed that the enzymatic activity was specific for m^{7}G cap-adjacent adenosines and did not methylate adenosines within the body of the mRNA. Despite their thorough work, the constraints imposed by the methods available at the time prevented them from cloning and identifying the protein(s) responsible. The identity of the CA-m^{6}Am methyltransferase would only be elucidated about four decades later.

4.1 PCIF1, the Writer of Cap-Adjacent m^{6}Am

In contrast to m^{6}A, which is added by a complex of proteins, CA-m^{6}Am is added to RNA by a single protein, *p*hosphorylated *C*TD-*i*nteracting *f*actor *1* (PCIF1, also called CAPAM for *cap*-specific *a*denosine *m*ethyltransferase) (Akichika et al. 2019; Boulias et al. 2019; Sendinc et al. 2019; Sun et al. 2019). For continuity, we will refer to this protein as PCIF1 hereafter (see Box 1 for an important note concerning another protein named PCIF1). Several independent groups published studies identifying PCIF1 as the enzyme responsible for CA-m^{6}Am addition in quick succession (Akichika et al. 2019; Boulias et al. 2019; Sendinc et al. 2019; Sun et al. 2019). Each group took a slightly different track to identify the writer of m^{6}Am. The fractions containing CA-m^{6}Am-adding enzymatic activity were isolated from HEK293 cell extracts following the same workflow devised four decades earlier (Keith et al. 1978; Sun et al. 2019). Next, mass spectrometry was used to identify candidate proteins that co-fractionated with the CA-m^{6}Am-adding activity (Sun et al. 2019). Among the proteins in their list, they focused on PCIF1 since its evolutionary conservation suggested that it possessed methyltransferase activity (Iyer et al. 2016; Sun et al. 2019). They validated their result when they observed a decrease in CA-m^{6}Am when LC-MS/MS was performed on mRNA harvested from cells where PCIF1 was knocked down with *s*mall *i*nterfering *RNAs* (siRNAs). They cross-validated this observation by demonstrating that recombinant PCIF1 could methylate a target RNA in vitro while active site point mutants could not. Finally, m^{6}A-seq studies in PCIF1 knockdown and control cells observed a loss of signal only in the 5′ UTR of mRNAs (Sun et al. 2019).

Box 1 Two Human Proteins Are Currently Named PCIF1

A distinct arc of papers follows a different PCIF1 protein that is not involved in the methylation of CA-m^{6}Am. Those papers also refer to *P*DX1 *C*-terminal *i*nhibiting *f*actor *1*, the human homolog of SPOP (speckle-type POZ protein) as PCIF1 (Bjorkbacka et al. 2017; Claiborn et al. 2010; Hensel et al. 2015; Klein et al. 2016; Liu et al. 2004). Unfortunately, the two different proteins

(continued)

Box 1 (continued)
appear to have been named PCIF1 in quick succession (PCIF1, phosphorylated CTD-interacting factor 1) in 2003 and SPOP/PCIF1 in 2004 (Fan et al. 2003; Liu et al. 2004). Our research shows that SPOP/PCIF1 (HGNC: 11254, Entrez Gene: 8405, Ensembl: ENSG00000121067) and PCIF1/CAPAM (HGNC: 16200, Entrez Gene: 63935, Ensembl: ENSG00000100982) are in fact distinct genes with distinct protein products observed as 374 (~42 kDa) and 704 (~80 kDa) amino acids, respectively. Indeed, western blots from these works show a ~45 kDa band for epitope-tagged SPOP/PCIF1 that matches expectations for SPOP rather than PCIF1 (Klein et al. 2016; Liu et al. 2004). Further, the papers mentioned above show that the untagged, recombinant PCIF1 that can generate CA-m^{6}Am is ~80 kDa (Sendinc et al. 2019). Although Ensembl does list PDX1 C-terminal inhibiting factor 1 as an alias for PCIF1/CAPAM, epitranscriptomics researchers should ensure that their reagents target the proper protein.

CRISPR-mediated deletions of PCIF1 in cultured cells coupled to rescue experiments with exogenous functional or mutated PCIF independently confirmed PCIF1 as the methylase required to add CA-m^{6}Am marks (Akichika et al. 2019; Boulias et al. 2019; Sendinc et al. 2019). Although the underlying approaches were consistent, each of these studies asked slightly different questions. First, RNA mass spectrometry was used to precisely compute m^{6}Am methylation sites in the 5-′-terminal cap structures of the capped mRNAs in normal and PCIF1-deleted cells (Akichika et al. 2019). Importantly, they also solved a high-resolution structure that delineated the mechanism by which PCIF1 uses *S*-adenosylmethionine to catalyze the N6-methylation of cap-adjacent-Am to form CA-m^{6}Am (Akichika et al. 2019). miCLIP experiments in WT and PCIF1 knockout cells complemented those experiments and observed that PCIF1 overexpression increased the prevalence of CA-m^{6}Am in cultured cells (Boulias et al. 2019). Finally, a new method called m^{6}Am-Exo-Seq, which relies on exonucleolytic digestion of uncapped RNAs, mapped the transcriptome-wide distribution of m^{6}Am vs. m^{6}A (Sendinc et al. 2019). Their data confirmed earlier reports by showing that the signals from m^{6}A and m^{6}Am sites did not overlap, suggesting that m^{6}Am has a function distinct from m^{6}A (Sendinc et al. 2019). Collectively all four groups showed that PCIF1 is required for mRNA m^{6}Am methylation in vivo and that recombinant PCIF1 can methylate capped mRNA in vitro. Together, these data show that PCIF1 is both necessary and sufficient to add CA-m^{6}Am to mRNAs. The final proof was provided by Pandey et al. when, perhaps surprisingly, they succeeded in generating *PCIF1*$^{-/-}$ mice (Pandey et al. 2020). Their work confirmed PCIF1 as the sole enzyme responsible for CA-m^{6}Am addition in vivo as *PCIF1*$^{-/-}$ mice completely lacked CA-m^{6}Am (Pandey et al. 2020).

4.2 FTO, an m^{6}Am Eraser

While there is some controversy as to whether FTO demethylates m^{6}Am, m^{6}A, or both in vivo, there is broad agreement that FTO demethylates m^{6}Am and CA-m^{6}Am in different types of RNA (Shi et al. 2019; Zaccara et al. 2019). By combining different methods FTO was convincingly shown to remove methyl groups from m^{6}Am in different contexts. As described above, the structural basis for FTO's recognition of CA-m^{6}Am has been established (Zhang et al. 2019a). In vitro assays showed that FTO has a much higher affinity for m^{6}Am, particularly CA-m^{6}Am, as opposed to m^{6}A (Mauer et al. 2017). In fact, when recombinant FTO was added to an equimolar mixture of m^{6}A- and m^{6}Am-containing RNA oligonucleotides, only m^{6}Am was demethylated (Mauer et al. 2017). Others have posited that the subcellular localization of FTO could play a role in regulating its activity (Wei et al. 2018). That reasoning is supported by work which showed that FTO could demethylate both internal m^{6}Am and CA-m^{6}Am from snRNAs and CA-m^{6}Am from mRNAs (Wei et al. 2018). Supporting this finding, FTO was independently demonstrated to reversibly demethylate CA-m^{6}Am snRNAs (Mauer et al. 2019). Deletion of FTO in adult neurons resulted in m^{6}Am-focused epitranscriptomic changes (Engel et al. 2018). Their final observation was that deletion of FTO identified 1801 putative m^{6}Am peaks which were enriched in developmental and DNA–RNA related genes by gene ontology (Engel et al. 2018).

4.3 Functions of CA-m^{6}Am

All investigators in the field agree that the identity and methylation status of the cap-adjacent nucleotide influence the mRNA's characteristics and several experimental systems have been established to help elucidate the function(s) of CA-m^{6}Am (Mauer et al. 2017; Pandey et al. 2020; Sikorski et al. 2020). This consensus was built upon data from targeted and transcriptome-wide mapping techniques. First, overexpression of FTO alters the ratio of m^{6}Am to Am in cells (Mauer et al. 2017). Next, once PCIF1 was identified as the writer of CA-m^{6}Am, wild-type and *Pcif1*-knockout cells made it possible to separate internal m^{6}A and CA-m^{6}Am marks on their respective mRNAs (Boulias et al. 2019; Mauer et al. 2017). Overexpression of PCIF1 in HEK293T cells led to a ~3-fold increase in the m^{6}Am to Am ratio showing that overexpression studies could also help determine the in vivo functions of CA-m^{6}Am (Boulias et al. 2019). Finally, altering the levels of CA-m^{6}Am has effects on mRNA metabolism in vivo. For example, PCIF1$^{-/-}$ mice are viable but show a pronounced growth defect (Pandey et al. 2020). Further, stress and glucocorticoid exposure can change m^{6}Am and m^{6}A marks and their regulatory network in a gene-specific manner (Engel et al. 2018). FTO's demethylase activity has also been linked to the repression of the stem-like phenotype in colorectal cell cancers (Relier et al. 2020).

However, despite the available tools, methods, and data focusing on CA-m^{6}Am, the current consensus regarding the function(s) of CA-m^{6}Am in vivo is that there is no consensus (Doxtader and Nam 2019). As described below, the data from different but complementary methods detail a general disagreement as to the function(s) of CA-m^{6}Am and its effects on mRNA stability and translation in vivo (Mauer et al. 2017; Sendinc et al. 2019; Sikorski et al. 2020). In fact, every function is attributed to CA-m^{6}Am; the modification's effects on mRNA decapping, mRNA stability, and mRNA translation all require further examination and clarification.

4.3.1 The Effects of CA-m^{6}Am on Decapping

CA-m^{6}Am has been shown to resist the activity of a key decapping enzyme Dcp2 and was initially thought to promote RNA stability (Mauer et al. 2017). Importantly, those data are bolstered as the analysis of transcriptomic data from mouse tissues and showed evidence that CA-m^{6}Am stabilized transcripts by inhibiting the action of DCP2, an mRNA decapping enzyme (Pandey et al. 2020). Despite these results, CA-m^{6}Am had little effect on the decapping activity of Dcp2 in vitro (Sikorski et al. 2020). That work showed that after 30 min of exposure to purified Dcp2, 25-mer RNAs beginning with three similar trinucleotide cap structures m^{7}G-A-G, m^{7}G-Am-G, and m^{7}G-m^{6}Am-G all showed similar levels (~65–75%) of decapping. Surprisingly, their data showed that, regardless of methylation status, RNAs beginning with an A (~70% decapped after 30 min) were much more susceptible to decapping by Dcp2 than RNAs beginning with G, C, or U (~25%, ~30%, and ~45% decapped, respectively). A key caveat is that these assays were performed entirely using an in vitro system with a short (25-mer) RNA (Sikorski et al. 2020). Therefore, their system does not account for cellular factors (such as cap binding proteins) and/or RNA secondary structures which could inhibit Dcp2 activity by obscuring the m^{7}G cap in vivo.

4.3.2 The Effect of CA-m^{6}Am on mRNA Levels

As mentioned above, CA-m^{6}Am was shown to correlate with an increase in the stability of CA-m^{6}Am-bearing mRNAs (Mauer et al. 2017). mRNAs beginning with CA-m^{6}Am were also somewhat resistant to microRNA-induced degradation (Mauer et al. 2017). Those data agreed with earlier work showing a similar increase in mRNAs with m^{6}A marks near their 5′ ends (Schwartz et al. 2014). An important note is that these earlier works were published prior to the identification of PCIF1, and therefore, their methods could not differentiate CA-m^{6}Am, m^{6}Am, or m^{6}A (Schwartz et al. 2014). Next, in vivo labeling experiments showed that preventing the addition of CA-m^{6}Am by knocking out PCIF1 significantly reduced stability of a subset of m^{6}Am-annotated mRNAs in HEK293 and HeLa cells (Boulias et al. 2019). In particular, two classes of CA-m^{6}Am-containing transcripts existed. A small group of transcripts with both very high copy number and very long (24+ h) half-lives were

not affected strongly by PCIF1 knockout. The second class consisted of less abundant transcripts that were particularly destabilized by the loss of CA-m^{6}Am (Boulias et al. 2019). This transcript-specific difference in mRNA stability may suggest that other factors work in concert with CA-m^{6}Am to influence mRNA stability.

CA-m^{6}Am differentially regulates transcript levels in *Pcif1*$^{-/-}$ mouse tissues, with starkly different numbers of changed mRNAs in the testes (~12,000), brain (~1500), and spleen (~750) (Pandey et al. 2020). *Pcif1*$^{-/-}$ mouse tissues also revealed the dysregulation of many pseudogenes and predicted gene transcripts. In addition, transcripts with a TSS adenosine were predominantly downregulated in transcriptome-wide measurements of RNA from *Pcif1*$^{-/-}$ mouse tissues. An important caveat regarding these data is that while most downregulated mRNAs began with adenosines, which was decidedly the case in the testes; however, on balance across all tissues, the majority of upregulated mRNAs began with adenosines as well (Pandey et al. 2020). The authors suggest that the regulation imparted by CA-m^{6}Am depends upon other, likely tissue-specific, factors which confer a multi-tiered and tunable regulation to their host mRNAs.

In contrast to the data showing CA-m^{6}Am as a stabilizer of mRNAs, others have shown that CA-m^{6}Am has either the opposite effect or no effect on mRNA stability. Steady-state measurements of RNA levels showed that only ~60 mRNAs changed substantially upon knockout of PCIF1, suggesting that the presence of CA-m^{6}Am had little bearing on mRNA stability (Akichika et al. 2019). m6Am-Exo-Seq was developed to accurately map CA-m^{6}Am, and was able to identify a subset of CA-m^{6}Am-bearing transcripts (Sendinc et al. 2019). The combination of m^{6}Am-Exo-Seq studies and sample-matched PRO-Seq experiments showed that m^{6}Am does not alter mRNA stability. Rather, the changes in steady-state levels of CA-m^{6}Am-bearing mRNAs were fully accounted for by changes to their basal transcription rates. While the effects of CA-m^{6}Am on mRNAs remain debated, to date, this study offers the most complete answer as it was the only one to control for mRNA levels by assaying the transcription rates of the changed genes (Sendinc et al. 2019).

4.3.3 The Translation of CA-m^{6}Am-Bearing mRNAs

Recent works used a combination of reporter assays, ribosome profiling, and mass spectrometry to assess the effects of CA-m^{6}Am on translation (Mauer et al. 2017; Sendinc et al. 2019). As with cap binding and mRNA stability above, their data have failed to produce a consensus as to the effect(s) of CA-m^{6}Am on translation. First, ribosome profiling data taken from HEK293T cells showed that mRNAs with CA-m^{6}Am were translated more efficiently than other mRNAs (Mauer et al. 2017). Once PCIF1's activity was identified, additional ribosome profiling data from WT and PCIF1 knockout HEK293T cells showed that the translation efficiency of CA-m^{6}Am-bearing mRNAs decreased in cells where PCIF1 was deleted (Akichika et al. 2019). Further, their data showed that the translation of upstream

open reading frames and the distribution of ribosomes were not affected by deleting PCIF1 (Akichika et al. 2019).

The influence of CA-m^{6}Am on translation was further tested by transfecting meticulously purified in vitro-transcribed luciferase mRNAs into three different cell lines (Sikorski et al. 2020). They reported that mRNAs with CA-m^{6}Am mRNAs were translated more efficiently in different cell lines than mRNAs with other beginning nucleotides. The experiment centered on transfecting identical mRNAs that differed only in the identity and methylation status of the first transcribed nucleotide. All their readings were normalized against luciferase mRNA possessing an adenosine in a Cap 0 context, a curious choice, since such a cap structure represents a small minority of natively transcribed mRNAs in mammalian cells. Particularly strong increases (~7 fold) in the translation of CA-m^{6}Am-containing mRNA (measured by relative luciferase signals) were observed in JAWS II (immortalized immature mouse dendritic) cells with a smaller increase (~1.5-fold) in HeLa cells and no change in 3T3-L1 cells. As summarized above, their data show large differences between cell types. For example, CA-m^{6}Am-bearing mRNAs were translated at a ~4-fold higher rate when comparing to the same mRNA with a Cap 1 guanosine in 3T3-L1 and HeLa cells, but they report a ~60-fold range for the same comparison in JAWS II cells (Sikorski et al. 2020). This difference is startling as the transfected mRNAs differ only by their first nucleotide and could evince an unknown translational control mechanism in JAWS II cells.

The analysis of ribosome profiling data from *Pcif1*$^{-/-}$ mouse brain tissue showed either up- or downregulation of translation depending upon the mRNA (Pandey et al. 2020). A comparatively small number of mRNAs exhibited increased or decreased translational efficiency with similar numbers of mRNAs showing increased or decreased translation. However, they found no correlation between changes in translation rates and the first transcribed nucleotide of the affected mRNA, suggesting that the observed change in translation was independent of CA-m^{6}Am (Pandey et al. 2020). Another ribosome profiling study also showed that the translation rates and protein levels of high confidence CA-m^{6}Am mRNAs were essentially unchanged in PCIF1 knockout HEK293T cells (Boulias et al. 2019).

Contradicting those results, several methods showed that CA-m^{6}Am marks negatively influenced the translation of their mRNAs (Sendinc et al. 2019). In a similar experiment to the one described above, purified in vitro-transcribed EGFP mRNAs beginning with either m^{7}G-cap-m^{6}Am or m^{7}G-cap-Am were transfected into WT and PCIF1-deleted MEL624 cells. The coupling of fluorescence microscopy with flow cytometry showed that CA-m^{6}Am-bearing mRNAs produced quantitatively lower GFP signals. Next, by adding an in vitro-transcribed dual luciferase reporter RNA to a common rabbit reticulocyte lysate translation system, CA-m^{6}Am was shown to decrease the translation of the reporter in a cap-dependent manner. Finally, mass spectrometry experiments comparing WT and PCIF1 knockout MEL624 cells showed that the levels of over 500 proteins increased, compared to 17 decreases, when PCIF1 was deleted. Taken together, their data show that CA-m^{6}Am negatively impacts cap-dependent translation of methylated mRNAs in MEL624 cell line (Sendinc et al. 2019).

In summary, as with the effect of CA-m^{6}Am on decapping and mRNA stability, the data regarding this epitranscriptomic mark's role in translation are contradictory and require further investigation and clarification.

5 Unanswered Questions Regarding Cap-Adjacent m^{6}Am

As described in detail above, many questions regarding the biological function(s) of CA-m^{6}Am lack clear answers. Currently, it is thought that yet to be identified cell-type specific factors are the likeliest drivers of these divergent results (Doxtader and Nam 2019). As with the controversy regarding FTO as an eraser of m^{6}A marks in vivo, the hope is that newer, more sensitive methods will help resolve the apparent conflicts with the reported data (Mauer and Jaffrey 2018). The identification of PCIF1 as the writer of CA-m^{6}Am and the availability of PCIF1$^{-/-}$ cells and mice have opened the door to asking many new questions (Table 2) regarding the role of CA-m^{6}Am in vivo. We discuss two of these unanswered questions in greater detail.

5.1 Is CA-m^{6}Am Addition by PCIF1 Truly a Co-transcriptional Event?

The presence of PCIF1's WW domain and the papers showing interactions with the phosphorylate C-terminal of RNA polymerase II, it has been assumed that CA-m^{6}Am addition is co-transcriptional (Akichika et al. 2019; Fan et al. 2003). Supporting this idea, exogenously expressed, epitope-tagged PCIF does localize predominantly to the nucleus, although cytoplasmic staining is visible for some cells, particularly for inactive point mutations of PCIF1 (Sendinc et al. 2019). Indirect immunofluorescence shows that PCIF1 is predominantly nuclear in most mouse tissues, although as with other works, some degree of cytoplasmic staining is evident in some of the images presented (Pandey et al. 2020; Sendinc et al. 2019). A careful reading of the older literature revealed that the CA-m^{6}Am adding activity had been isolated from the cytoplasm of HeLa cells (Keith et al. 1978). By coupling differential centrifugation to multiple rounds of column chromatography CA-m^{6}Am addition was performed by a cytoplasmic enzyme which was not associated with ribosomes, the mitochondria, or nuclei (Keith et al. 1978). Confirming that result, the first demonstration of PCIF1 as the CA-m^{6}Am methyltransferase isolated the activity from HEK293 cell cytoplasmic extracts (Sun et al. 2019). Re-examination of the other recent studies revealed that all experiments measuring CA-m^{6}Am deposition and PCIF1 activity were performed with whole cell lysates or extracts or with tagged constructs rather than the endogenous proteins (Akichika et al. 2019; Boulias et al. 2019; Mauer et al. 2017; Pandey et al. 2020; Sendinc et al. 2019). The demonstration that PCIF1 co-immunoprecipitates the phosphorylated C-terminal domain of RNA

Table 2 Salient questions regarding cap-adjacent m^6Am

Unanswered question	Reasoning/Implication
What is/are the role(s) of CA-m^6Am in vivo?	This fundamental question is still up for debate as several studies have yielded conflicting data.
Precisely how much of m^6A signal is actually CA-m^6Am?	The current assumption is that ~100% of the m^6A signal mapping to TSS and across the 5′ UTR is actually CA-m^6Am. Is this true?
What is the role of CA-m^6Am in stress?	Loss of PCIF1 has been shown to sensitize cells to oxidative stress. What mechanism surveys CA-m^6Am in stress? Does it apply to other stressors?
Which other decapping enzymes also have difficulty with removing CA-m^6Am?	Many decapping enzymes are known in eukaryotes, most of which are poorly characterized. Could one or more of these enzymes serve as CA-m^6Am readers?
Do any decapping enzymes exhibit higher affinity for RNAs with CA-m^6Am?	
Are all other cap binding proteins also CA-m^6Am readers?	The affinity of both eIF4E and Dcp2 for capped mRNAs is affected by the presence of CA-m^6Am. Do additional proteins (cap-binding or other) serve as CA-m^6Am readers?
What other cellular factors function as CA-m^6Am readers?	
Is FTO the only CA-m^6Am demethylase?	m^6A appears to have multiple functional demethylases. Is the same true for CA-m^6Am?
Does a particular FTO-interacting protein target it to CA-m^6Am?	Interactions with another protein could offer a broader regulatory potential by fine-tuning FTO's CA-m^6Am demethylase activity.
Is there an interplay between CA-m^6Am and other RNA modifications or the proteins that recognize them?	Interactions between proteins that recognize CA-m6Am and other epitranscriptomic marks would expand their regulatory potential.
Can cap-adjacent Am be methylated to form CA-m^6Am in the cytoplasm?	As most mature mRNAs localize to the cytoplasm, cytoplasmic addition of CA-m^6Am would offer more tunable regulation of the targeted mRNAs.

polymerase II offers the most direct proof that PCIF1 works co-transcriptionally (Akichika et al. 2019). However, those data were obtained using whole cell extracts, opening the possibility that the interaction with the C-terminal domain of RNA polymerase II could be an artifact caused by the destruction of the nuclear membrane during cell lysis (Akichika et al. 2019). By showing that PCIF1 is predominantly localized in the cytoplasm of HUVECs (Fig. 3) our data are consistent with a cytoplasmic role for PCIF1.

5.2 Could PCIF1 Function in Concert with Cytoplasmic Capping?

A cytoplasmic complex that adds a cap onto 5′-monophosphate RNAs and is capable of restoring m^7G caps to mRNAs in the cytoplasm was identified in 2009 (Otsuka

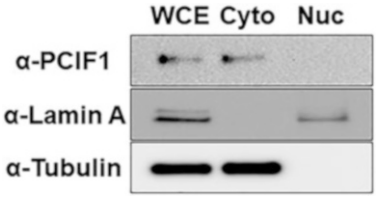


Fig. 3 Western blots demonstrate that PCIF1 is cytoplasmic in HUVEC cells. HUVEC cells were cultured in vascular cell basal medium (ATCC PCS-100-030) supplemented with Endothelial Cell Growth Kit-VEGF (ATCC PCS-100-041) at 37 °C and 5% CO_2. ~80% confluent cultures were rinsed with PBS and harvested using a cell lifter. Cell pellets were resuspended in 0.9 ml of lysis buffer (PBS pH 7.4, 0.1% NP40 (Thermofisher), 0.1 M PMSF (Sigma), protease inhibitor cocktail (Sigma), and phosphatase inhibitor (Sigma)) for 10 min. 300μl cell lysate was collected as whole cell extract (WCE) and sonicated for an hour at 4C using a Bioruptor Plus (Diagenode). The remaining cell lysate (600μl) was then centrifuged for 1 min at 21,000 xG and the supernatant was transferred to a new tube as cytoplasmic extracts (Cyto). The pelleted nuclei were rinsed once with lysis buffer, resuspended in fresh lysis buffer, and sonicated for an hour. Equal amounts of protein were separated using Mini-PROTEAN TGX Stain-free AnyKD gels (Biorad) and blotted onto TransBlot Turbo PVDF Membrane (Biorad). Blots were blocked using 5% skim milk and probed with α-PCIF1 (Abcam, ab205016), α-Lamin A (Invitrogen, MA1-06101), and α-Tubulin (Proteintech 66031-I-Ig). Blots probed with Lamin A and Tubulin serve as nuclear and cytoplasmic markers, respectively. Data presented are a single representative experiment from independent biological triplicate experiments

et al. 2009). The cytoplasmic capping complex includes RNGTT, *NCK* Adaptor Protein *1* (NCK1), an unidentified 5′-monophosphate kinase, and a heterodimer of RNMT with its activating subunit RAM (Trotman and Schoenberg 2018). NCK1 is a scaffold protein to coordinate the activities of RNGTT; a monophosphate kinase and the RNMT:RAM heterodimer interact to form the active complex in the cytoplasm (Trotman and Schoenberg 2018). Importantly, reported cell fractionation results provide strong supporting evidence for functional cytoplasmic capping as cytoplasmic extracts were shown to possess a methyltransferase activity capable of converting a G-capped RNA into a proper m^{7}G cap (Keith et al. 1978; Trotman et al. 2017). Inhibition of cytoplasmic cap methylation was used to identify 5′ *t*erminal *o*ligo*p*yrimidine (TOP)-containing mRNAs as cytoplasmic capping targets and uncovered cytoplasmic capping sites downstream of canonical 5′ ends (Del Valle Morales et al. 2020). Although the overall biological significance of cytoplasmic capping remains poorly understood, several reports show that cytoplasmic capping targets are enriched in mRNAs involved in mitotic cell cycle control, cellular stress responses, and development (Berger et al. 2019).

We have long thought that epitranscriptomic modifications may be among the keys to better understanding cytoplasmically capped mRNAs. For this reason, we are examining whether m^{6}A and/or m^{6}Am play an important role in cytoplasmically capped mRNAs. Possibly supporting this idea, numerous internally mapped m^{6}Am sites (16.7% of total) have been identified (Boulias et al. 2019). While internally mapping m^{6}Am sites were interpreted as arising from alternative TSSs, such CA-m^{6}Am sites could also arise from the cytoplasmic capping of truncated

mRNAs (Berger et al. 2019; Del Valle Morales et al. 2020). By showing that PCIF1 localizes to the cytoplasm (Fig. 3), our cell fractionation data agree with two papers demonstrating CA-m^{6}Am-adding activity in the cytoplasm (Keith et al. 1978; Sun et al. 2019). Together, these data imply that PCIF1 functions in the cytoplasm, either in addition to, or instead of, the nucleus. If confirmed, the cytoplasmic addition of CA-m^{6}Am could serve as a consequential and dynamic epitranscriptomic mark that helps regulate the translation and stability of mRNAs.

6 Closing Remarks

The field of epitranscriptomics has advanced greatly since the discovery of the first modified RNA nucleotide in 1957 (Davis and Allen 1957). While roughly 160 different RNA base modifications are currently known, most of them are poorly characterized (Boccaletto et al. 2018). Furthermore, the functions and the enzymes that write, read, and erase the majority of RNA modifications remain unknown (Boccaletto et al. 2018). This void of knowledge and the contradictory nature of some of the results are both certainly contributors to some of the recent skepticism regarding a functional and dynamic epitranscriptome (Darnell et al. 2018). As the field of epitranscriptomics continues to grow rapidly, we should expect (indeed, we should welcome) seemingly contradictory findings such as the apparently opposing effect(s) of CA-m^{6}Am on mRNA decapping, stability, and translation, the compartmentalization of PCIF1 activity, or the target(s) of the FTO demethylase. While such conflicting results can be confusing, they provide singular opportunities to better understand the fundamental biological mechanism(s) underlying the contradiction. In general, such conflicts can be resolved as new tools, techniques, and insights enable a more complete investigation of the systems involved. The multitude of unanswered questions ensures that advances in epitranscriptomics will continue to yield impactful findings for years to come.

Acknowledgments The authors would like to thank their colleagues in the Kiss lab and the Department of Cardiovascular Sciences at the Houston Methodist Research Institute (HMRI) for creating an innovative and challenging yet welcoming scientific environment. TTT was supported by funds from the Cancer Prevention and Research Institute of Texas (CPRIT, grant RP150611 to J.P. Cooke, MD, PhD.). DLK was supported by startup funds provided by the HMRI (to DLK) and a grant from the National Institutes of Health (1R35GM137819 to DLK). The content presented here is solely the responsibility of the authors and does not represent the official views of the HMRI, CPRIT, or the National Institutes of Health.

References

Aas A, Isakson P, Bindesboll C et al (2017) Nucleocytoplasmic shuttling of FTO does not affect starvation-induced autophagy. PLoS One 12:e0168182

Akichika S, Hirano S, Shichino Y et al (2019) Cap-specific terminal N (6)-methylation of RNA by an RNA polymerase II-associated methyltransferase. Science 363:eaav0080

Annapoorna PK, Iyer H, Parnaik T et al (2019) FTO: an emerging molecular player in neuropsychiatric diseases. Neuroscience 418:15–24

Anreiter I, Mir Q, Simpson JT et al (2020) New twists in detecting mRNA modification dynamics. Trends Biotechnol S0167-7799(20):30166–30169

Barbieri I, Tzelepis K, Pandolfini L et al (2017) Promoter-bound METTL3 maintains myeloid leukaemia by m(6)A-dependent translation control. Nature 552:126–131

Barbosa E, Moss B (1978) mRNA(nucleoside-2′-)-methyltransferase from vaccinia virus. Purification and physical properties. J Biol Chem 253:7692–7697

Bartosovic M, Molares HC, Gregorova P et al (2017) N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3′-end processing. Nucleic Acids Res 45:11356–11370

Berger MR, Alvarado R, Kiss DL (2019) mRNA 5′ ends targeted by cytoplasmic recapping cluster at CAGE tags and select transcripts are alternatively spliced. FEBS Lett 593:670–679

Bjorkbacka H, Yao Mattisson I, Wigren M et al (2017) Plasma stem cell factor levels are associated with risk of cardiovascular disease and death. J Intern Med 282:508–521

Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307

Bohnsack MT, Sloan KE (2018) Modifications in small nuclear RNAs and their roles in spliceosome assembly and function. Biol Chem 399:1265–1276

Bokar JA, Rath-Shambaugh ME, Ludwiczak R et al (1994) Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. J Biol Chem 269:17697–17704

Bokar JA, Shambaugh ME, Polayes D et al (1997) Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA 3:1233–1247

Boulias K, Toczydlowska-Socha D, Hawley BR et al (2019) Identification of the m(6)Am methyltransferase PCIF1 reveals the location and functions of m(6)Am in the transcriptome. Mol Cell 75:631–643.e638

Chen K, Lu Z, Wang X et al (2015) High-resolution N(6) -methyladenosine (m(6) A) map using photo-crosslinking-assisted m(6) A sequencing. Angew Chem Int Ed Engl 54:1587–1590

Claiborn KC, Sachdeva MM, Cannon CE et al (2010) Pcif1 modulates Pdx1 protein stability and pancreatic beta cell function and survival in mice. J Clin Invest 120:3713–3721

Crain PF, Choi YC, Busch H et al (1978) Characterization of N6, O2-dimethyladenosine from nuclear RNA of Novikoff hepatoma. Nucleic Acids Res 5:771–776

Darnell RB, Ke S, Darnell JE Jr (2018) Pre-mRNA processing includes N(6) methylation of adenosine residues that are retained in mRNA exons and the fallacy of "RNA epigenetics". RNA 24:262–267

Davis FF, Allen FW (1957) Ribonucleic acids from yeast which contain a fifth nucleotide. J Biol Chem 227:907–915

Del Valle Morales D, Trotman JB, Bundschuh R et al (2020) Inhibition of cytoplasmic cap methylation identifies 5′ TOP mRNAs as recapping targets and reveals recapping sites downstream of native 5′ ends. Nucleic Acids Res 48:3806–3815

Dominissini D, Rechavi G (2017) Loud and clear epitranscriptomic m(1)A signals: now in single-base resolution. Mol Cell 68:825–826

Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485:201–206

Dong Z, Cui H (2020) The emerging roles of RNA modifications in glioblastoma. Cancers (Basel) 12:736

Doxtader KA, Nam Y (2019) Manipulation by methylation: garnishing mRNAs with m(6)Am. Mol Cell 75:417–418

Edupuganti RR, Geiger S, Lindeboom RGH et al (2017) N(6)-methyladenosine (m(6)A) recruits and repels proteins to regulate mRNA homeostasis. Nat Struct Mol Biol 24:870–878

Engel M, Chen A (2018) The emerging role of mRNA methylation in normal and pathological behavior. Genes Brain Behav 17:e12428
Engel M, Eggert C, Kaplick PM et al (2018) The role of m(6)A/m-RNA methylation in stress response regulation. Neuron 99:389–403.e389
Enroth C, Poulsen LD, Iversen S et al (2019) Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing. Nucleic Acids Res 47:e126
Ensinger MJ, Moss B (1976) Modification of the 5′ terminus of mRNA by an RNA (guanine-7-)-methyltransferase from HeLa cells. J Biol Chem 251:5283–5291
Fan H, Sakuraba K, Komuro A et al (2003) PCIF1, a novel human WW domain-containing protein, interacts with the phosphorylated RNA polymerase II. Biochem Biophys Res Commun 301:378–385
Fejes-Toth K, Sotirova V, Sachidanandam R et al (2009) Post-transcriptional processing generates a diversity of 5′-modified long and short RNAs. Nature 457:1028–1032
Frye M, Watt FM (2006) The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. Curr Biol 16:971–981
Garcia-Campos MA, Edelheit S, Toth U et al (2019) Deciphering the "m(6)A Code" via antibody-independent quantitative profiling. Cell 178:731–747.e716
Ghosh A, Lima CD (2010) Enzymology of RNA cap synthesis. Wiley Interdiscip Rev RNA 1:152–172
Goh YT, Koh CWQ, Sim DY et al (2020) METTL4 catalyzes m6Am methylation in U2 snRNA to regulate pre-mRNA splicing. Nucleic Acids Res 48:9250–9261
Grosjean H, Szweykowska-Kulinska Z, Motorin Y et al (1997) Intron-dependent enzymatic formation of modified nucleosides in eukaryotic tRNAs: a review. Biochimie 79:293–302
Hawley BR, Jaffrey SR (2019) Transcriptome-wide mapping of m(6) A and m(6) Am at single-nucleotide resolution using miCLIP. Curr Protoc Mol Biol 126:e88
He C (2010) Grand challenge commentary: RNA epigenetics? Nat Chem Biol 6:863–865
Hensel J, Duex JE, Owens C et al (2015) Patient mutation directed shRNA screen uncovers novel bladder tumor growth suppressors. Mol Cancer Res 13:1306–1315
Hess ME, Hess S, Meyer KD et al (2013) The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. Nat Neurosci 16:1042–1048
Hsu PJ, He C (2019) High-resolution mapping of N (6)-methyladenosine using m(6)A crosslinking immunoprecipitation sequencing (m(6)A-CLIP-Seq). Methods Mol Biol 1870:69–79
Huang H, Weng H, Sun W et al (2018) Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol 20:285–295
Huang H, Weng H, Zhou K et al (2019) Histone H3 trimethylation at lysine 36 guides m(6)A RNA modification co-transcriptionally. Nature 567:414–419
Huang T, Liu Z, Zheng Y et al (2020) YTHDF2 promotes spermagonial adhesion through modulating MMPs decay via m(6)A/mRNA pathway. Cell Death Dis 11:37
Hussain S, Aleksic J, Blanco S et al (2013) Characterizing 5-methylcytosine in the mammalian epitranscriptome. Genome Biol 14:215
Iyer LM, Zhang D, Aravind L (2016) Adenine methylation in eukaryotes: apprehending the complex evolutionary history and functional potential of an epigenetic modification. BioEssays 38:27–40
Jia G, Fu Y, Zhao X et al (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887
Ke S, Pandya-Jones A, Saito Y et al (2017) m(6)A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. Genes Dev 31:990–1006
Keith JM, Ensinger MJ, Moss B (1978) HeLa cell RNA (2′-O-methyladenosine-N6-)-methyltransferase specific for the capped 5′-end of messenger RNA. J Biol Chem 253:5033–5039
Klein S, Meng R, Montenarh M et al (2016) The phosphorylation of PDX-1 by protein kinase CK2 is crucial for its stability. Pharmaceuticals (Basel) 10:2

Koh CWQ, Goh YT, Goh WSS (2019) Atlas of quantitative single-base-resolution N(6)-methyladenine methylomes. Nat Commun 10:5636

Lamond AL (1990) The trimethyl-guanosine cap is a nuclear targeting signal for snRNPs. Trends Biochem Sci 15:451–452

Lavi U, Fernandez-Munoz R, Darnell JE Jr (1977) Content of N-6 methyl adenylic acid in heterogeneous nuclear and messenger RNA of HeLa cells. Nucleic Acids Res 4:63–69

Leung DW, Amarasinghe GK (2016) When your cap matters: structural insights into self vs non-self recognition of 5′ RNA by immunomodulatory host proteins. Curr Opin Struct Biol 36:133–141

Li Y, Wu K, Quan W et al (2019) The dynamics of FTO binding and demethylation from the m(6)A motifs. RNA Biol 16:1179–1189

Lin S, Choe J, Du P et al (2016) The m(6)A methyltransferase METTL3 promotes translation in human cancer cells. Mol Cell 62:335–345

Linder B, Jaffrey SR (2019) Discovering and mapping the modified nucleotides that comprise the epitranscriptome of mRNA. Cold Spring Harb Perspect Biol 11

Liu A, Desai BM, Stoffers D (2004) Identification of PCIF1, a POZ domain protein that inhibits PDX-1 (MODY4) transcriptional activity. Mol Cell Biol 24:4372–4383

Liu N, Dai Q, Zheng G et al (2015) N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature 518:560–564

Lorenz DA, Sathe S, Einstein JM et al (2020) Direct RNA sequencing enables m(6)A detection in endogenous transcript isoforms at base-specific resolution. RNA 26:19–28

Ma H, Wang X, Cai J et al (2019) N(6-)Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. Nat Chem Biol 15:88–94

Marchand V, Ayadi L, Ernst FGM et al (2018) AlkAniline-Seq: profiling of m(7) G and m(3) C RNA modifications at single nucleotide resolution. Angew Chem Int Ed Engl 57:16785–16790

Mardakheh FK, Paul A, Kumper S et al (2015) Global analysis of mRNA, translation, and protein localization: local translation is a key regulator of cell protrusions. Dev Cell 35:344–357

Mauer J, Jaffrey SR (2018) FTO, m(6) Am, and the hypothesis of reversible epitranscriptomic mRNA modifications. FEBS Lett 592:2012–2022

Mauer J, Luo X, Blanjoie A et al (2017) Reversible methylation of m(6)Am in the 5′ cap controls mRNA stability. Nature 541:371–375

Mauer J, Sindelar M, Despic V et al (2019) FTO controls reversible m(6)Am RNA methylation during snRNA biogenesis. Nat Chem Biol 15:340–347

Meyer KD (2019) DART-seq: an antibody-free method for global m(6)A detection. Nat Methods 16:1275–1280

Meyer KD, Jaffrey SR (2017) Rethinking m(6)A readers, writers, and erasers. Annu Rev Cell Dev Biol 33:319–342

Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149:1635–1646

Meyer KD, Patil DP, Zhou J et al (2015) 5′ UTR m(6)A promotes cap-independent translation. Cell 163:999–1010

Mikutis S, Gu M, Sendinc E et al (2020) meCLICK-Seq, a substrate-hijacking and RNA degradation strategy for the study of RNA methylation. ACS Central Science

Molinie B, Wang J, Lim KS et al (2016) m(6)A-LAIC-seq reveals the census and complexity of the m(6)A epitranscriptome. Nat Methods 13:692–698

Nance DJ, Satterwhite E, Bhaskar B et al (2020) Characterization of METTL16 as a cytoplasmic RNA binding protein. PLoS One 15:e0227647

Otsuka Y, Kedersha NL, Schoenberg DR (2009) Identification of a cytoplasmic complex that adds a cap onto 5′-monophosphate RNA. Mol Cell Biol 29:2155–2167

Pan T (2018) Modifications and functional genomics of human transfer RNA. Cell Res 28:395–404

Pandey RR, Delfino E, Homolka D et al (2020) The mammalian cap-specific m(6)Am RNA methyltransferase PCIF1 regulates transcript levels in mouse tissues. Cell Rep 32:108038

Panneerdoss S, Eedunuri VK, Yadav P et al (2018) Cross-talk among writers, readers, and erasers of m(6)A regulates cancer growth and progression. Sci Adv 4:eaar8263
Pendleton KE, Chen B, Liu K et al (2017) The U6 snRNA m(6)a methyltransferase METTL16 regulates SAM synthetase intron retention. Cell 169:824–835.e814
Perry RP, Kelley DE (1974) Existence of methylated messenger RNA in mouse L cells. Cell 1:37–42
Ramanathan A, Robb GB, Chan SH (2016) mRNA capping: biological functions and applications. Nucleic Acids Res 44:7511–7526
Relier S, Ripoll J, Guillorit H et al (2020) FTO-mediated cytoplasmic m^6A_m demethylation adjusts stem-like properties in colorectal cancer cell. bioRxiv, 2020.2001.2009.899724
Reolon LW, Vichier-Guerre S, de Matos BM et al (2019) Crystal structure of the Trypanosoma cruzi EIF4E5 translation factor homologue in complex with mRNA cap-4. Nucleic Acids Res 47:5973–5987
Roundtree IA, Evans ME, Pan T et al (2017) Dynamic RNA modifications in gene expression regulation. Cell 169:1187–1200
Ruszkowska A, Ruszkowski M, Dauter Z et al (2018) Structural insights into the RNA methyltransferase domain of METTL16. Sci Rep 8:5311
Saletore Y, Meyer K, Korlach J et al (2012) The birth of the epitranscriptome: deciphering the function of RNA modifications. Genome Biol 13:175
Scholler E, Weichmann F, Treiber T et al (2018) Interactions, localization, and phosphorylation of the m(6)A generating METTL3-METTL14-WTAP complex. RNA 24:499–512
Schwartz S, Mumbach MR, Jovanovic M et al (2014) Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5′ sites. Cell Rep 8:284–296
Sendinc E, Valle-Garcia D, Dhall A et al (2019) PCIF1 catalyzes m6Am mRNA methylation to regulate gene expression. Mol Cell 75:620–630.e629
Shatkin AJ (1976) Capping of eucaryotic mRNAs. Cell 9:645–653
Shi H, Wang X, Lu Z et al (2017) YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. Cell Res 27:315–328
Shi H, Wei J, He C (2019) Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers. Mol Cell 74:640–650
Shu X, Cao J, Cheng M et al (2020) A metabolic labeling method detects m(6)A transcriptome-wide at single base resolution. Nat Chem Biol 16:887–895
Sikorski PJ, Warminski M, Kubacka D et al (2020) The identity and methylation status of the first transcribed nucleotide in eukaryotic mRNA 5′ cap modulates protein expression in living cells. Nucleic Acids Res 48:1607–1626
Sommer S, Lavi U, Darnell JE Jr (1978) The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. J Mol Biol 124:487–499
Song H, Wang Y, Wang R et al (2020) SFPQ is an FTO-binding protein that facilitates the demethylation substrate preference. Cell Chem Biol 27:283–291.e286
Squires JE, Patel HR, Nousch M et al (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40:5023–5033
Stoilov P, Rafalska I, Stamm S (2002) YTH: a new domain in nuclear proteins. Trends Biochem Sci 27:495–497
Sun H, Zhang M, Li K et al (2019) Cap-specific, terminal N(6)-methylation by a mammalian m(6) Am methyltransferase. Cell Res 29:80–82
Trotman JB, Schoenberg DR (2018) A recap of RNA recapping Wiley Interdiscip Rev RNA e1504
Trotman JB, Giltmier AJ, Mukherjee C et al (2017) RNA guanine-7 methyltransferase catalyzes the methylation of cytoplasmically recapped RNAs. Nucleic Acids Res 45:10726–10739
Viehweger A, Krautwurst S, Lamkiewicz et al (2019) Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. Genome Res 29:1545–1554
Wang X, Zhao BS, Roundtree A et al (2015) N(6)-methyladenosine modulates messenger RNA translation efficiency. Cell 161:1388–1399

Wang J, Alvin Chew BL, Lai Y et al (2019) Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. Nucleic Acids Res 47:e130

Wang Y, Xiao Y, Dong S et al (2020) Antibody-free enzyme-assisted chemical approach for detection of N(6)-methyladenosine. Nat Chem Biol 16:896–903

Wei CM, Moss B (1977) Nucleotide sequences at the N6-methyladenosine sites of HeLa cell messenger ribonucleic acid. Biochemistry 16:1672–1676

Wei C, Gershowitz A, Moss B (1975) N6, O2′-dimethyladenosine a novel methylated ribonucleoside next to the 5′ terminal of animal cell and virus mRNAs. Nature 257:251–253

Wei CM, Gershowitz A, Moss B (1976) 5′-terminal and internal methylated nucleotide sequences in HeLa cell mRNA. Biochemistry 15:397–401

Wei J, Liu F, Lu Z et al (2018) Differential m(6)A, m(6)Am, and m(1)A demethylation mediated by FTO in the cell nucleus and cytoplasm. Mol Cell 71:973–985 e975

Werner M, Purta E, Kaminska KH et al (2011) 2′-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. Nucleic Acids Res 39:4756–4768

Xiang S, Liang X, Yin S et al (2020) N6-methyladenosine methyltransferase METTL3 promotes colorectal cancer cell proliferation through enhancing MYC expression. Am J Transl Res 12:1789–1806

Xu C, Wang X, Liu K et al (2014) Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. Nat Chem Biol 10:927–929

Zaccara S, Ries RJ, Jaffrey SR (2019) Reading, writing and erasing mRNA methylation. Nat Rev Mol Cell Biol 20:608–624

Zhang X, Wei LH, Wang Y et al (2019a) Structural insights into FTO's catalytic mechanism for the demethylation of multiple RNA substrates. Proc Natl Acad Sci USA 116:2919–2924

Zhang Z, Chen LQ, Zhao YL et al (2019b) Single-base mapping of m(6)A by an antibody-independent method. Sci Adv 5:eaax0250

Zhao L, Zhang H, Kohnen MV et al (2019) Analysis of transcriptome and epitranscriptome in plants using PacBio Iso-Seq and nanopore-based direct RNA sequencing. Front Genet 10:253

Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29

Zou S, Toh JD, Wong KH et al (2016) N(6)-Methyladenosine: a conformational marker that regulates the substrate specificity of human demethylases FTO and ALKBH5. Sci Rep 6:25677

RNA m^{6}A Modification: The Mediator Between Cellular Stresses and Biological Effects

Lingfang Wang, Yasen Maimaitiyiming, Kunhui Su, and Chih-Hung Hsu

Contents

Lingfang Wang, Yasen Maimaitiyiming and Kunhui Su contributed equally with all other contributors.

L. Wang · K. Su · C.-H. Hsu (✉)
Women's Hospital, and Institute of Genetics, and Department of Environmental Medicine, Zhejiang University School of Medicine, Hangzhou, China
e-mail: ch_hsu@zju.edu.cn

Y. Maimaitiyiming
Department of Hematology of First Affiliated Hospital, and Department of Public Health, Zhejiang University School of Medicine, Hangzhou, China

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_13

Abstract Cells are constantly exposed to ubiquitous threats from the external and intracellular sources, including ultraviolet (UV), temperature switch, pathogen infection, starvation, etc. These adverse conditions would disturb cellular homeostasis by dysregulation of RNA metabolisms, such as transcription, splicing, translation, and so forth. Similar to DNA and proteins, RNA is subject to various (over 160) covalent modifications, among which m^6A is the most abundant internal modification on messenger RNA (mRNA) and plays crucial roles in regulation of RNA-related bioprocesses. Recently, increasing evidence indicated that RNA modifications could be the "sensor" to recognize and respond to external and intracellular stresses. For example, we found that UV exposure rapidly and transiently induced the m^6A on RNA at DNA damage sites to recruit Pol κ for efficient DNA repair. Several studies also showed RNA modifications responding to other stresses such as starvation, heat shock, and pathogen infection. For instance, heat shock could directly or indirectly affect distribution and abundance of m^6A modification, regulated by m^6A modifiers, which in turn influenced the expression of specific genes (*HSPs*, *MYC*, *circE7*) and the downstream bioprocesses to respond to the temperature stress. In this section, we summarize the involvement of m^6A RNA modification in regulation of distinct stress responses and discuss the recent advances in the underlying molecular mechanisms involved in these regulations to get a comprehensive picture of the functions of RNA m^6A modifications in response to cellular stresses.

Keywords RNA modification · Stress response · N^6-methyladenosine (m^6A) · UV · DNA damage · Virus · Heat shock · Starvation · Hypoxia

1 Introduction

Similar to DNA and proteins, RNA carries a diverse array of post-transcriptional chemical modifications (over 160), including N^6-methyladenosine (m^6A), N^1-methyladenosine (m^1A), 5-methylcytosine (m^5C), 5-hydroxymethylcytosine (hm^5C), 2′-O-methylation (Nm), isomerization of uridine to pseudouridine (Ψ), as well as a newly identified acetylation of cytidine (ac^4C), which play crucial roles in RNA metabolism (Esteve-Puig et al. 2020). Among these, m^6A is the most abundant and most studied internal modification on messenger RNA (mRNA) in many eukaryotic species, including yeast (Schwartz et al. 2013), zebrafish (Zhao et al. 2017), plants (Luo et al. 2014), and mammals (Dominissini et al. 2012; Meyer et al. 2012). The methylation at the N^6 position of adenosine is a dynamic and reversible process, and its function is mainly regulated by three kinds of "m^6A modifiers," including m^6A methyltransferases (writers), m^6A binding proteins (readers), and m^6A demethylases (erasers) (Liang et al. 2020).

The components of the major "m^6A writer" complex include methyltransferase like factor 3 (METTL3), methyltransferase like factor 14 (METTL14), and Wilms tumor associated protein 1 (WTAP). In this "m^6A writer" complex, METTL3 is the

only subunit with m^{6}A catalytic activity (Ping et al. 2014; Wang et al. 2016). WTAP promotes m^{6}A modification by recruiting METTL3 and METTL14 into nuclear speckles, and METTL14 associates with METTL3 and enhances its activity (Chen et al. 2019b). Apart from METTL14 and WTAP, several METTL3 associated factors which regulated METTL3 functions were reported recently. For example, VIRMA-METTL3 preferentially mediates mRNA methylation near the 3′-UTR and stop codon regions, and RNA binding motif protein 15 (RBM15/15B) binds uridine rich region to recruit METTL3 complex to methylate adjacent DRACH motif (where D denotes A/G/U, R denotes A/G, and H denotes A/C/U) (Patil et al. 2016; Yue et al. 2018). Besides METTL3-containing m^{6}A writer complexes, methyltransferase like factor 16 (METTL16) was found to participate in catalyzation of m^{6}A methylation on U6 snRNA (Pendleton et al. 2017). Notably, the presence of m^{6}A was still observed on poly(A) enriched RNA in METTL3 KO cells (Xiang et al. 2017), suggesting the existence of unidentified m^{6}A writers.

After m^{6}A deposited on RNA by writers, m^{6}A readers directly bind to m^{6}A-containing RNA to regulate methylated RNA metabolism and determine RNA fate. YTH domain family proteins are the well-known m^{6}A readers; each of them possesses distinct functions in regulating m^{6}A-containing RNA, such as the ability of YTHDC1 on regulating alternative splicing (Roundtree and He 2016), YTHDC2 on stabilizing m^{6}A-mofidied mRNA (Kretschmer et al. 2018), YTHDF1 and YTHDF3 on promoting translation efficiency (Wang et al. 2015; Shi et al. 2017), and YTHDF2 on accelerating mRNA degradation (Wang et al. 2014). A recent study found another reader family, IGF2BP family proteins, mainly working on stabilizing the m^{6}A-containing mRNA recognized by them and displaying certain oncogenic functions (Huang et al. 2018). Currently identified m^{6}A erasers include alkylation repair homolog protein 5 (ALKBH5) as well as fat mass and obesity-associated protein (FTO), both of which could remove m^{6}A modification from RNA; therefore, m^{6}A RNA modification is reversible and adjustable (Jia et al. 2011; Zheng et al. 2013).

Increasing evidence showed that the steady state of m^{6}A, regulated by m^{6}A writers, readers, and erasers, plays a critical role in multiple bioprocesses including self-renewal (Cui et al. 2017), circadian clock regulation (Fustin et al. 2013), maternal-zygotic transition (Zhao et al. 2017), tumorigenesis (Deng et al. 2018), and so on. Recently, more and more studies indicated the biological function of RNA modifications on stress response (Shi et al. 2019; Song et al. 2019; Wang et al. 2015; Xiang et al. 2017). In fact, RNA modification is not only a "sensor" of cellular stresses but is also a "regulator" for cellular homeostasis upon stresses (Feng et al. 2020; Xiang et al. 2017; Zhang et al. 2016a); for example, m^{6}A plays an important role in DNA damage response (DDR) system for repairing UV-induced damage (Sun et al. 2018; Svobodová Kovaříková et al. 2020; Xiang et al. 2017; Zhang et al. 2020a). Besides UV, m^{6}A responded to a wide range of stresses like heat shock (Huang et al. 2018; Zhou et al. 2015), pathogen infection (Kennedy et al. 2016), starvation (Yang et al. 2019b), etc. Apart from m^{6}A, several other RNA modifications have been demonstrated to respond to stresses as well. For instance, UVC or UVA exposure reduced the levels of m^{1}A and 2,2,7-trimethylguanosine (m^{3}G/

TMG) modifications, and hypoxia decreased m^5C and m^1A levels in total RNA (Svobodová Kovaříková et al. 2020).

Interestingly, the regulation between RNA modification and stress response is usually mutual. Stresses affect the abundance, localization, and existing time of RNA modifications; on the other hand, modified RNAs regulate gene expression and cellular bioprocesses to modulate cellular stress responses and stress adaptation. Additionally, while over-affordable stresses would be harmful, modest stresses under certain conditions would be beneficial to organisms. For example, moderate heat (hyperthermia) could be used for treatment of several types of cancers and other diseases (Snider et al. 2016); starvation took advantage of the reduction of basic cellular metabolism to promote cell survival in nutrient deficient conditions (Ravanan et al. 2017; Shi et al. 2012). As RNA modification is a newly emerging component of stress-response system, further investigation will be needed, which will provide better understanding for developing the strategies to reduce stress-induced adverse effects and enhance stress-mediated benefits.

In this chapter, we summarized the current understanding of the roles of m^6A RNA modification in stress responses. Additionally, we would discuss the recent advances in the underlying molecular mechanisms involved in the mutual regulations between modified RNA and stresses to have a comprehensive picture of the functions of m^6A RNA in response to cellular stresses.

2 Ultraviolet (UV)

UV light as a form of electromagnetic radiation is a kind of common environmental stresses coming from sunlight and certain artificial illuminant sources like welding torches, etc. Although moderate exposure to UV has some benefits for people such as stimulation of vitamin D production (Wacker and Holick 2013), prolonged and excessive UV exposure would lead to damage of cells and even result in syndromes and diseases, including the development of premature aging (Magimaidas et al. 2016), potentially blinding diseases (Dave et al. 2019), and skin cancers (Cadet and Douki 2018; D'Orazio et al. 2013), which are due to the accumulation of UV-induced DNA damage and mutations in UV-exposed cells (Moriwaki and Takahashi 2008). It was found that UV irradiation generated several types of modifications/damages on RNA including photochemical modifications, crosslinking, and oxidative damage (Svobodová Kovaříková et al. 2020). In our previous study, we demonstrated UV-induced m^6A as a new DNA damage response (DDR) factor which responded to UV and facilitated DNA repair (Xiang et al. 2017). This work and subsequent studies suggested a potential research direction for investigation of the correlation between DNA damage response and RNA modifications (Robinson et al. 2019; Svobodová Kovaříková et al. 2020; Zhang 2017). Here, we review the current understanding about the mutual regulation between UV exposure and RNA modifications.

2.1 UV Affects Both Abundance and Distribution of RNA Modifications

UV stress dynamically affected the overall levels of RNA m^{6}A modification in a time- and dosage-dependent manner. Besides m^{6}A, it was also demonstrated that UV exposure could alter the levels of other modifications on RNA (Sun et al. 2018; Svobodová Kovaříková et al. 2020; Xiang et al. 2017; Zhang et al. 2020a). In *Escherichia coli* tRNA model, UV exposure led to a remarkable and successive decrease in the level of most sulfur-containing modifications, including s^{2}C, s^{4}U, mnm^{5}s^{5}U, cmnm^{5}s^{2}U, and ms^{2}i^{6}A, along with an induction of mn^{5}U, nm^{5}U, and ho^{5}U (Sun et al. 2018). Notably, m^{1}G and m^{6}A were degraded under UV exposure in vitro, while they were unchanged even after 2-hours exposure under in vivo conditions, suggesting a specific category of factors interacting with these two modifications and regulating their stability in vivo (Sun et al. 2018). Although the modifications are known to modulate structural stability of tRNA and its decoding efficiency of mRNA to affect gene expression, why and how UV alters tRNA modifications is still elusive, and further investigation will be needed to elucidate the underlying mechanism. Similarly, UV affects RNA modifications in eukaryotic cells (Svobodová Kovaříková et al. 2020; Xiang et al. 2017; Zhang et al. 2020a). While N^1-methyladenosine (m^{1}A) and 2,2,7-trimethylguanosine (m^{3}G/TMG) were reduced under the exposure of UVC or UVA, RNA m^{6}A was upregulated (Svobodová Kovaříková et al. 2020; Xiang et al. 2017; Zhang et al. 2020a). However, it is not clear whether these distinct modifications could mutually regulate each other, and what kind of RNA species (mRNA, tRNA, snRNA, etc.) carried UV-response modifications in eukaryotes. Therefore, to understand the RNA modification-mediated UV-exposed response circuitry, further examination will be required.

Besides directly altering the overall levels, UV stress also influenced the distribution of RNA modifications. Remarkably, a notable UV-induced m^{6}A relocalization was reported by three independent groups, which found an obvious accumulation of m^{6}A-containing RNA at UV-irradiated damage sites (Svobodová Kovaříková et al. 2020; Xiang et al. 2017; Zhang et al. 2020a). Our previous work demonstrated that m^{6}A modification rapidly and transiently existed at UV-irradiated DNA damage sites, and the UV-stimulated m^{6}A-modified polyA(+) RNAs played a key role for local DNA repair. Furthermore, sequence analysis of m^{6}A-modified RNAs revealed the known METTL3 target site "GGACU" and an additional degenerate "AACUG" motif on UV-stimulated m^{6}A-containing RNA. Together, these results showed that UV affected both the subcellular localization of m^{6}A-containing RNA and the m^{6}A-modified sites on a single m^{6}A-containing RNA transcript (Xiang et al. 2017). Svobodová Kovaříková et al. (2020) observed that a relatively high level of m^{6}A-containing RNA was present in the cytoplasm in HaCaT, MEFs, and HeLa cells in normal condition, while the m^{6}A-containing RNA diffused from cytoplasm to nucleus in UV-exposed cells. Although the biological significance is unclear, UV altered the distribution of m^{1}A as well. The tiny foci of

m^1A RNA detected inside the cell nucleus disappeared after UV irradiation (Svobodová Kovaříková et al. 2020). Zhang et al. (2020a) reported that m^6A-modified RNA accumulates at the foci of DNA double-strand breaks (DSBs) and forms RNA-DNA hybrids to promote homologous recombination-mediated repair.

The relocalization of RNA modifications upon UV irradiation might be due to the redistribution of RNA modifiers. Immediately after UV exposure, PARP1-mediated poly(ADP-ribos)ylation occurred on DNA damage sites for recruiting METTL3 and METTL14 but not WTAP to catalyze m^6A on DNA damage-associated RNA, and then around 4 min after UV irradiation, FTO could be detected at DNA damage sites. Since METTL3 localization preceded that of FTO at damage sites, which is likely to allow a brief window of m^6A accumulation. This would explain why the presence of m^6A at damage sites is rapid but transient (Xiang et al. 2017). On the other hand, besides working as a "damage sensor" for immediate and transient presence at UV damage sites, m^6A RNA also localized at damage sites at later stages. METTL16, a methyltransferase for U6 snRNA, highly expressed in around 10% of microirradiated cells, showed an accumulation at UV-irradiated sites at later stages (18–20 min), suggesting METTL16-mediated m^6A-modified snRNA participating in the DNA repair at this period (Svobodová Kovaříková et al. 2020).

2.2 *UV-Regulated RNA Modifications Play Significant Roles for Cell Homeostasis*

Organisms evolved various pathways to sense and overcome DNA damage, for example, single-strand damages are generally restored by nucleotide excision repair (NER), base excision repair (BER), mismatch repair, and so on, while double-strand breaks are fixed by non-homologous end joining (NHEJ), microhomology-mediated end joining (MMEJ), and homologous recombination (HR). In addition to the repair systems, the translesion synthesis (TLS), which is a DNA damage tolerance process, was also evolved to allow the DNA replication machinery to replicate past DNA lesions such as AP sites (apurinic/apyrimidinic sites) or cyclobutane pyrimidine adducts (CPD). Actually, m^6A would take part in multiple DNA damage repair pathways to restore cellular homeostasis.

In our previous study, we found that m^6A serves as an important mediator in UV-induced DNA damage repair (Xiang et al. 2017). Specifically, in mammalian cells, the early DNA damage responder PARP but not γH2A.X induced by UV exposure led to the accumulation of METTL3/METTL14 complex to DNA damage sites, thereby contributing to the deposition of m^6A-modified RNA at damage sites to promote the removal of UV-induced CPD. Depletion of METTL3 resulted in reduction of m^6A and delayed removal of UV-induced CPD. Notably, only wild type but not catalytic mutant METTL3 rescued the CPD-removal defect in METTL3 depleted cells, suggesting the importance of m^6A for response to UV stress. In

addition, DNA polymerases κ (Pol κ), which operates in both nucleotide excision repair (NER) and translesion synthesis (TLS) pathways, acted as a downstream factor of METTL3-m^{6}A-mediated DNA repair pathway, as loss of m^{6}A led to reduction of Pol κ recruitment which impaired DNA repair (Xiang et al. 2017). These results raise a new METTL3-m^{6}A-Pol κ UV response axis (Xiang et al. 2017) (Fig. 1, left panel).

Recent studies revealed that m^{6}A is also involved in the DNA double-strand break (DSB) repair. DSB works as an initiator for R-loop (a kind of three-stranded nucleic acid structures with a DNA-RNA heteroduplex and a displaced DNA single strand, also termed as DNA-RNA hybrids), which plays an important role for responding to DNA DSB repair process. Zhang et al. (2020a) found that the RNA moiety of R-loop can be methylated. They found an ataxia telangiectasia mutated (ATM)-triggered phosphorylation of METTL3 at S43 by DSBs, which facilitated METTL3 associating with RNA Pol II and being recruited to DSB sites for m^{6}A modification on local RNA. The m^{6}A-containing RNA facilitated the recruitment of m^{6}A reader protein YTHDC1, playing a critical role in protection of m^{6}A-containing DNA-RNA hybrids, which then guaranteed efficient recruitment of RAD51 and BRCA1 for subsequent homologous recombination (HR)-mediated repair (Zhang et al. 2020a) (Fig. 1, right panel). However, another study demonstrated that instead of promoting the formation of R-loop, m^{6}A modification occurred at the RNA moiety of R-loop leads to its degradation (Abakir et al. 2020). In this study, authors found the RNA moiety of R-loop can be m^{6}A modified; however, their data showed that the m^{6}A-RNA reader YTHDF2 but not YTHDC1 was recruited to R-loop and subsequently destabilized DNA-RNA hybrids, which may represent a novel mechanism of preventing accumulation of co-transcriptional R-loops during mitosis. Knocking down YTHDC1 didn't influence the formation and stability of R-loop. A possible explanation is that the involvement of m^{6}A in DNA damage repair is a dynamic process. First, DNA damage caused DNA double-strand break which led to the formation of R-loop. Protection of m^{6}A-containing RNA-DNA hybrids at the early stage by the recognition of YTHDC1 promoted recruitment of DNA repair factors for subsequent homologous recombination-mediated repair (Zhang et al. 2020a) (Fig. 1, right panel). Second, after the DNA repair, the m^{6}A-RNA reader YTHDF2 is recruited to R-loop and subsequently destabilizes DNA-RNA hybrids, which prevented accumulation of co-transcriptional R-loops during mitosis. Apart from m^{6}A writers and m^{6}A readers, m^{6}A erasers might participate in DNA damage response. For example, osteoblasts were more susceptible to genotoxic agents (UV and H_2O_2) when FTO was depleted, suggesting that FTO-mediated reduction of m^{6}A might play a role in UV and H_2O_2 response pathways in osteoblasts (Zhang et al. 2019).

Besides m^{6}A, a recent study demonstrated that m^{5}C can be locally induced by DSBs at sites of DNA damage and take part in homologous recombination (HR) (Chen et al. 2020). Authors found TRDMT1, a tRNA (cytosine(38)-C(5))-methyltransferase, is recruited to the DNA-RNA hybrids at sites of DNA damage, promoting the m^{5}C modification on RNA moiety in DNA-RNA hybrids, through which facilitate the recruitment of RAD51 and RAD52 to sites of DNA damage.

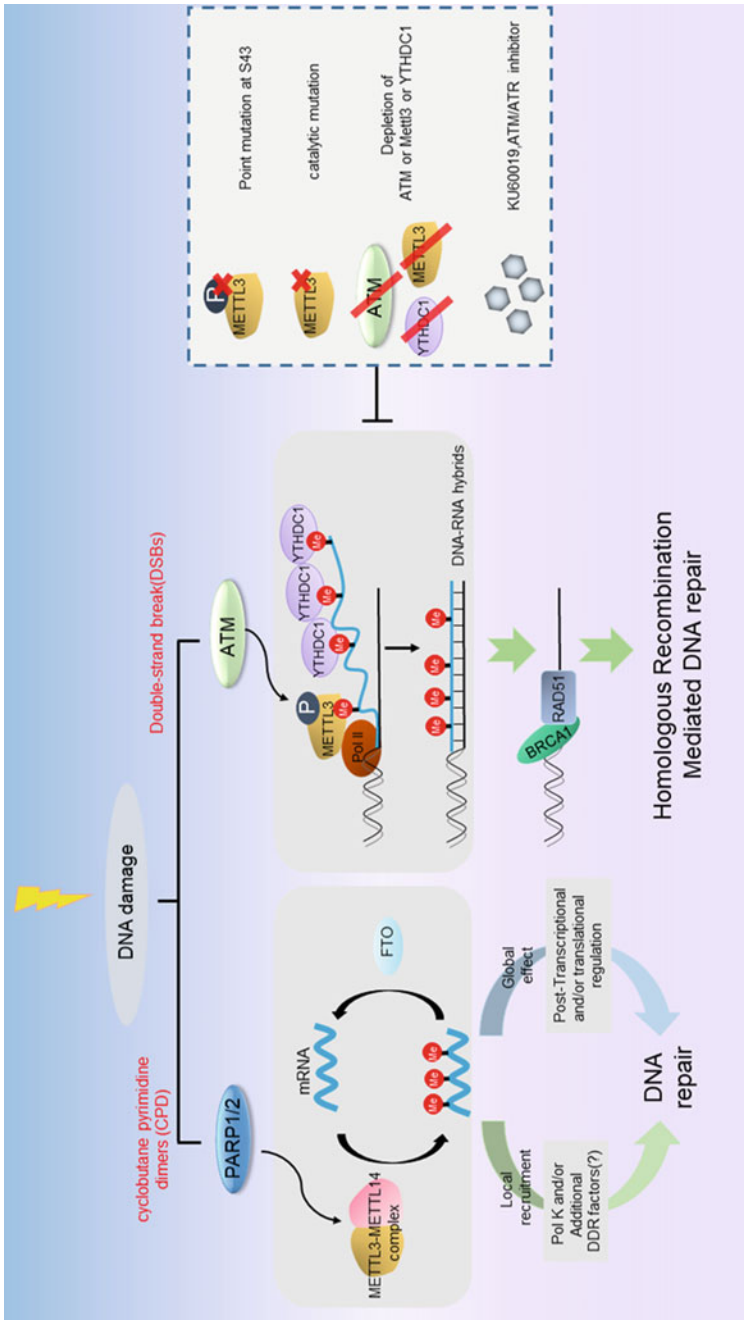


Fig. 1 Working model of m^{6}A RNA-mediated regulation in DNA repair. m^{6}A modification participated in the CPD removal and double-strand breaks (DSBs) repair upon DNA damage. CPD accumulation induced by UV irradiation resulted in the recruitment of METTL3/14 complex to DNA damage sites, which led to an accumulation of m^{6}A-containing RNA at the damage regions, then recruiting DNA polymerase k (Pol k) for DNA repair. On the other hand, ataxia telangiectasia mutated (ATM) triggered by DSBs phosphorylated METTL3 at S43, which facilitated METTL3 associating with RNA Pol II to be recruited to DSBs. RNAs at DSBs were catalyzed by activated METTL3 and then recognized and protected by YTHDC1. This METTL3-m^{6}A-YTHDC1 axis increased the accumulation of DNA-RNA hybrids at DSBs which promoted the recruitment of RAD51 and BRCA1 for homologous recombination-mediated repair

Thus, the RNA moiety of DNA-RNA hybrids at DNA damage sites could be regulated by different RNA modifications, which played important roles in efficient DNA repair. Additionally, it was found that the s^{4}U modification in tRNA protects bacteria from near-UV light. UV led to a photochemical reaction resulting in intramolecular cross-linking between s^{4}U and specific cytosine, and then promoted deposition of guanosine 5′-diphosphate-3′-diphosphate (ppGpp), which is a stress-response factor in triggering growth delay (Kramer et al. 1988).

Together, UV exposure affected cell homeostasis resulting in dysregulation of RNA modifiers (enzymes and readers) as well as downstream RNA modifications, and modified RNA showed ability to respond to adverse stress for repairing DNA and maintaining homeostasis. Therefore, investigating the mutual regulation between RNA modifications and UV-stimulated damage would promote better understanding of UV stress response circuitry.

3 Pathogen

Pathogen infection (e.g., bacteria, virus) has long been a serious public health and clinical challenge, and it triggers cellular stress response as well (Ruggieri et al. 2012). Infection of pathogens is a game of survival between the host and pathogens to let pathogens achieve a successful infection and adapt to a hostile environment, or let the host eliminate infected pathogens. Recent findings suggested that this kind of game is hinged on RNA modifications. Generally, the pathogen-induced alteration of RNA modifications occurred on both pathogens and host RNA, which could change the outcome of this game. RNA modifications have been confirmed to be deposited on the RNAs of a number of pathogens, such as human immunodeficiency virus 1 (HIV-1), influenza A virus (IAV), human respiratory syncytial virus (RSV), enterovirus 71 (EV71), human metapneumovirus (HMPV), Zika virus (ZIKV), hepatitis C virus (HCV), etc., and have different roles on regulating transcripts of these viruses (Table 1). In this part, we summarize the involvement of RNA modifications during pathogen infection and host anti-infection response, so as to provide novel insights into coping with pathogen-related adverse health effects.

3.1 Viral Infection-Stimulated RNA Modifications Regulate Viral Replication

Recent studies found that increased m^{6}A modification on viral RNAs would regulate viral replication. The HIV-1 viral mRNAs are subjected to m^{6}A modification at 3′-UTR that strongly enhanced viral replication through the m^{6}A reader protein YTHDF2. When YTHDF2 was overexpressed, viral mRNA was stabilized, and viral replication was also increased in CD4$^+$ human T cell line CEM-SS, while

Table 1 Involvement of m^6A modification during viral infection and host antiviral defense

Virus		Pattern of involvement	References
RNA viruses	HIV-1	Enrichment of m^6A modification at 3′-UTR strongly increases viral replication through the m^6A reader protein YTHDF2	(Kennedy et al. 2016)
		m^6A modification in the Rev response-element region of HIV-1 increases its binding to Rev and facilitates nuclear export of viral RNA, which at last enhances HIV-1 replication	(Lichinchi et al. 2016a)
		YTHDF3 is incorporated into viral particles which permits itself limit infection in the next cycle of infection	(Jurczyszak et al. 2020)
		YTHDF1–3 decrease viral genomic RNA (gRNA) levels and inhibit both early and late reverse transcription of HIV-1	(Lu et al. 2018)
		HIV-1 infection increases m^6A levels of cellular RNA mainly mediated by the binding of HIV-1 gp120 and CD4 receptor	(Tirumuru and Wu 2019)
		Acetylation of the N4 position of cytidine (ac4C) on HIV-1 RNAs by NAT10 enhances viral RNA stability, and the expression of HIV-1 can be inhibited by Remodelin, an inhibitor of NAT10	(Tsai et al. 2018)
		Addition of m^5C modification on HIV-1 RNA not only prompts viral translation by increased ribosome binding but increases alternative splicing of HIV-1 transcripts	(Courtney et al. 2019)
	IAV	Addition of m^6A modification on IAV RNA not only increases viral gene expression and replication but also its pathogenicity mainly mediated by YTHDF2	(Courtney et al. 2017)
	RSV	Addition of m^6A modification on RSV transcripts enhances viral replication	(Xue et al. 2019)
	EV71	EV71 infection alters the expression patterns of m^6A methyltransferases, demethylases, and binding proteins to facilitate the interaction between m6a modifiers and EV71 RNA, which at last enhances EV71 replication	(Yao et al. 2020); (Hao et al. 2019)
	HMPV	Addition of m^6A on HMPV RNA helps itself mimic cellular RNA to escape recognition by RNA sensor RIG-I and avoid detection by innate immunity	(Lu et al. 2020)
	ZIKV	M6A modification inhibits ZIKV replication	(Lu et al. 2020)
		ZIKV infection affects human and viral RNA by altering the topology and m^6A modification	(Lichinchi et al. 2016a)
	HCV	m6A modification negatively regulates HCV viral particle production in a YTHDF1/2/3-dependent manner	(Gokhale et al. 2020)
DNA viruses	SV40	m^6A modifications occurred on SV40 enhance viral replication mainly mediated by YTHDF2, but its mechanisms remain unknown	(Tsai et al. 2018)
	KSHV	Binding of YTHDF2 to m^6A-modified KSHV transcripts leads to degradation of KSHV RNA which at last inhibits lytic replication	(Tan et al. 2018)

(continued)

Table 1 (continued)

Virus		Pattern of involvement	References
		Depletion of YTHDF2 eliminates lytic entry and virion production of KSHV in iSLK.219 cells but increases ORF50 abundance in the B cell line TREX-BCBL-1	(Hesser et al. 2018)
		Lytic switch protein RTA can induce m6A modification at its own pre-mRNA which enhances its splicing via YTHDC1	(Ye et al. 2017)
	HBV	Binding of YTHDF2 to m^{6}A-modified HBV transcripts inhibits RIG-1 recognition and achieves immune evasion	(Kim et al. 2020a)
		YTHDF2 recognizes m^{6}A-modified HBV transcripts and recruits ISG20 to degrade these HBV transcripts	(Imam et al. 2020)
		m6A at 3′-UTRs of HBV RNA decreases the stability of HBV by YTHDF2. However, m6A at the 5′ epsilon loop promotes reverse transcription of pgRNA	(Imam et al. 2018)

knockdown of YTHDF2 showed the opposite effect (Kennedy et al. 2016). Interestingly, although YTHDF3 reduced the infectivity of newly produced viruses, this antiviral m^{6}A reader protein is incorporated into viral particles and cleaved by the HIV protease (Jurczyszak et al. 2020). Consistently, other studies also reported that HIV-1 infection caused a massive increase in m^{6}A modification on viral RNAs, promoting viral protein translation, virion particle production, and virus replication (Lichinchi et al. 2016b; Tirumuru et al. 2016). In HIV-1 infected T cells, a total of 14 methylation peaks were identified in coding and non-coding regions of viral mRNA (Lichinchi et al. 2016b). The authors pinpointed that the m^{6}A site A7883 in the Rev response element (RRE) bulge region is pivotal for viral RNA nuclear export and HIV-1 replication. Replication of influenza A virus (IAV), human respiratory syncytial virus (RSV), Kaposi's sarcoma-associated herpesvirus (KSHV), enterovirus 71 (EV71), and human metapneumovirus (HMPV) was also positively regulated by m^{6}A modification on viral transcripts (Baquero-Perez et al. 2019; Courtney et al. 2017; Hao et al. 2019; Lu et al. 2020; Xue et al. 2019; Yao et al. 2020). Simian virus 40 (SV40) transcripts were subjected to m^{6}A modification by the host writer proteins, enhancing translation of viral late transcripts in a YTHDF2-dependent manner and promoting viral replication (Tsai et al. 2018). Although multiple mechanisms are involved, increased m^{6}A modification of viral RNAs by host m^{6}A modifiers promoted replication of abovementioned viruses. Thus, suppression of viral RNA m^{6}A levels through interfering host enzymes could be a potential strategy for treatment of infections with these viruses.

On the other hand, in certain viruses, increased m^{6}A modification of viral RNAs could be a negative regulator for viral processes. Hepatitis B virus (HBV) is one of the major causes of hepatocellular carcinoma. Recently, it is reported that m^{6}A-modified HBV transcripts are selectively recognized and subjected to degradation by interferon-stimulated gene 20 (ISG20) in a YTHDF2-dependent manner (Imam et al. 2020). Zika virus (ZIKV) is a mosquito-borne RNA virus associated with severe neurological disorders. Infection with ZIKV is known to induce cellular

stress through phosphorylation of eIF2α and shutting-off host protein synthesis (Oyarzún-Arrau et al. 2020). However, the cellular response mechanism to ZIKV infection is poorly elucidated. Lichinchi et al. (2016a) found that ZIKV infection affected human and viral RNAs by altering the topology of m^6A modification. A total of 12 m^6A peaks were identified across the full length ZIKV RNA, and m^6A modification of ZIKV RNA was controlled by host methyltransferases METTL3/14 as well as demethylases ALKBH5 and FTO. YTHDF1/2/3 recognized m^6A modifications and destabilized ZIKV mRNA, inhibiting viral replication (Lichinchi et al. 2016a). Infection of *Flaviviridae* viral RNAs including human hepatitis C virus (HCV) is also negatively regulated by m^6A modification in a YTHDF1/2/3-dependent manner (Gokhale et al. 2020).

3.2 *Viral Infection Affects RNA Modifications on Host Transcripts*

In addition to alteration of modifications on viral RNA, viral infection affected m^6A modification on host transcripts, which again could positively or negatively regulate viral processes dependent on the context. It was found that the helicase DEAD-box (DDX) family member DDX46 recruits ALKBH5 following vesicular stomatitis virus (VSV) infection, leading to demethylation of m^6A-modified antiviral transcripts, enforcing their retention in the nucleus and preventing their translation, inhibiting interferon production, and thereby facilitating viral replication (Zheng et al. 2017). Similarly, HIV-1 infection led to unique m^6A modification on 56 human gene transcripts to promote viral gene expression and virus replication (Lichinchi et al. 2016b). In contrast, another study reported that in response to viral infection (e.g., VSV; Sendai virus, SeV; encephalomyocarditis virus, EMCV; herpes simplex virus type 1, HSV-1), host cells impaired the enzymatic activity of the m^6A demethylase ALKBH5, increasing m^6A modification on α-ketoglutarate dehydrogenase (OGDH) and reducing its mRNA stability and protein expression (Liu et al. 2019b). Reduction in OGDH level resulted in decreased production of the metabolite itaconate that is important for virion particle production and virus replication. Intriguingly, infection by *Flaviviridae* family viruses (e.g., dengue virus, DENV; Zika virus, ZIKV; HCV) activated endoplasmic reticulum (ER) stress response and increased m^6A modification of *RIOK3* mRNA and protein expression, which in turn increased production of DENV and ZIKV particles but restrained HCV replication (Gokhale et al. 2020). Rubio et al. (2018) reported that replication of the cytomegalovirus (HCMV) is regulated by host m^6A methyltransferase METTL14 and demethylase ALKBH5. Mechanistically, HCMV infection altered the expression of m^6A writers METTL3/14, erasers ALKBH5 and FTO, as well as reader proteins, stimulating type I interferon (IFN) production as an innate immune response. METTL3/14 depletion downregulated m^6A modification and stabilized interferon (INFB1) mRNA, leading to increased INFB1 production and restrained viral

replication, while silencing ALKBH5 showed opposite effects. Infection with Kaposi's sarcoma-associated herpesvirus (KSHV) altered both host and viral transcripts' m^{6}A level; intriguingly, depletion of m^{6}A machinery displayed pro- or antiviral effects on viral gene expression depending on the infected cell types (Baquero-Perez et al. 2019; Hesser et al. 2018).

Notably, mutual regulation between RNA m^{6}A modification and viral infection was also recorded. For example, enterovirus 71 (EV71) would affect host m^{6}A modifiers, and the replication of infected EV71 was also modulated by host m^{6}A modifiers (Hao et al. 2019). Briefly, EV71 viral RNA contained m^{6}A modification which modulated viral replication. Knockdown of host m^{6}A writer (METTL3), m^{6}A eraser (FTO), and m^{6}A readers (YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2) influenced EV71 replication. Moreover, the outcomes were similar between depletion of METTL3 and mutation of the m^{6}A methylated sites in the infectious clone; both would decrease the production of EV71 progeny, which suggested that the host m^{6}A modifier-mediated m^{6}A modification on EV71 RNA regulated the fate of infected EV71. On the other hand, the expression of m^{6}A modifiers was affected after EV71 infection. The key components (METTL3 and METTL14) of m^{6}A writer complexes and certain m^{6}A readers (YTHDF1, YTHDF2, YTHDF3, and YTHDC1) were upregulated, but the expression of m^{6}A eraser (FTO) was decreased by EV71 infection. Markedly, both upregulation of m^{6}A writer and downregulation of m^{6}A eraser benefited EV71 replication and production. In addition, this study showed that METTL3 could regulate EV71 replication via a m^{6}A-independent manner. They observed that host METTL3 associated with viral RdRp 3D protein and modulated its modification to facilitate viral replication.

3.3 2019-nCoV and RNA Modifications

Several studies demonstrated that the coronavirus family members including 2019 novel coronavirus (2019-nCoV or SARS-CoV-2), SARS, and MERS viruses were also subjected to RNA modifications. A recent study identified an interaction of COVID-19 NSP5 (C145A) with tRNA methyltransferase 1 (TRMT1), which is responsible for synthesis of the dimethylguanosine (m2,2G) on both nuclear and mitochondrial tRNAs (Gordon et al. 2020), which altered the function of tRNA. Another study found at least 41 RNA modification sites on COVID-19 viral transcripts, with the most frequent motif being "AAGAA," while the type of modifications was yet to be elucidated (Kim et al. 2020b). They also found the RNA modifications on COVID-19 RNA were negatively correlated to the length of poly (A) tails. While further studies will be needed, given the important role of poly (A) tails in RNA turnover, it is likely that these modifications are involved in the stability control of SARS-CoV-2 RNA.

3.4 *Bacteria, Host, and RNA Modifications*

Besides viral infections, RNA modifications also participated in regulation of bacterial infections. A recent study found that heat-killed *Salmonella typhimurium* (HKST) infection induced m^6A modification on RNAs related to the cytokine production and the inflammatory signaling pathways mainly by increasing expression of WTAP. These m^6A-modified transcripts were subjected to regulation by YTHDF2 (Wu et al. 2020). Specifically, YTHDF2 bound to and destabilized the *KDM6B* transcripts via an m^6A-dependent manner, and loss of YTHDF2 specifically enhanced the expression of a subset of proinflammatory cytokines, such as IL-6 and IL-12B, during bacterial infection by facilitating H3K27me3 demethylation at their promoters through KDM6B. These findings put forward a new concept that there is a crosstalk between histone and RNA modifications in bacteria-induced inflammatory response. In addition, although the detailed mechanisms are elusive, mice gut microbiome regulated the m^6A level and remodeled m^6A landscape in several organs, including the brain, intestine, cecum, liver, etc., which might play certain roles to regulate physiological conditions of the host (Jabs et al. 2020; Wang et al. 2019b).

Overall, RNA modifications are key regulators in the survival game between pathogens and host during infection within a complicated regulatory network. Further understanding of the infection-induced alteration of epitranscriptomic modifications will be crucial in revealing the potential RNA-modification-based treatment against pathogenic infection.

4 Heat Stress

Heat shock/stress response (HSR) is a self-protective reaction of cells and organisms following exposure to elevated temperatures beyond normal. HSR is mainly regulated by the transcription factor HSF1 (heat shock factor 1). Under normal conditions, HSF1 localizes at the cytoplasm as an inactive-monomer form; upon heat stress, HSF1 rapidly relocates into the nucleus and trimerizes to associate with HSEs (heat-shock response elements) in the promoter regions of target chaperone genes, leading to prompt and massive expression of these chaperones termed as HSPs (heat shock proteins) to promote refolding of misfolded proteins and protect cells (Anckar and Sistonen 2011). Overwhelming heat stress or improper activation of HSR would cause protein unfolding, entanglement, and unspecific aggregation, leading to the disruption of cytoskeleton, nuclear components, and other cellular organelles, which in turn resulted in irreversible damage or even cell death (Richter et al. 2010). Increasing number of studies suggest an interaction between RNA modification and HSR (Dominissini et al. 2012; Schwartz et al. 2014; Meyer et al. 2015; Zhou et al. 2015; Yu et al. 2018; Feng et al. 2020). For instance, mRNAs of *HSF1* and *HSPs* are rich in m^6A sites, and their m^6A levels change upon heat shock

(Dominissini et al. 2012; Feng et al. 2020; Meyer et al. 2015; Schwartz et al. 2014; Yu et al. 2018; Zhou et al. 2015). Both temperature sensitivity of cells and the expression of *HSF1* as well as *HSPs* are regulated by m^{6}A modification, indicating m^{6}A playing an important role in the regulation of HSR. Besides alteration of RNA modifications, several studies independently demonstrated that heat stress affected the expression and intracellular localization of RNA modifiers (writers, erasers, readers).

In this part, we would focus on recent advances with respect to RNA modifications and heat stress response and systematically summarize the underlying molecular mechanisms involved in the mutual regulation of RNA modification and HSR, so as to provide a comprehensive picture of the interaction between RNA m^{6}A modification and HSR.

4.1 Heat Stress Regulates RNA Modifications

Growing evidence suggested that heat stress not only altered the overall level but also the subcellular localization of RNA modifications. For example, it was found that the tissues of both abdominal fat and liver taken from piglets raised in the heat-room had a higher level of m^{6}A-modified RNA compared to the control (Heng et al. 2019), which was consistent with another study conducted on sheep (Lu et al. 2019). Abundance of other RNA modifications, such as m^{1}A or m^{7}G, was also increased upon heat shock in cells (Alriquet et al. 2020; Malbec et al. 2019). Since RNA modifications are regulated by RNA modifiers (writers, erasers, and readers), heat shock might modulate RNA modifiers to alter RNA modifications. Indeed, the expression levels of the m^{6}A writers (METTL3, METTL14, WTAP), m^{6}A erasers (FTO), and m^{6}A readers (YTHDF2, YTHDF3) were regulated following heat shock in a tissue-specific manner. For instance, YTHDF2 was elevated in both abdominal adipose and liver, while METTL3 and METTL14 were only elevated in abdominal adipose following exposure to heat (Heng et al. 2019); the results from sheep model were very similar, and it was further found that YTHDF3 protein decrease upon heat shock as well (Lu et al. 2019). On the other hand, although Zhou et al. (2015) similarly found a significant increase in the mRNA and protein levels of YTHDF2 following heat shock in MEF and HeLa cells, none of the other m^{6}A modifiers were changed in their experimental systems. These findings suggest that the effects of heat shock/stress on m^{6}A modifiers are context specific. However, the underlying mechanisms involved in the tissue- and cell-specific regulation of m^{6}A modifiers' expression upon heat shock/stress remain largely unknown.

Recently, several studies showed that certain heat shock/stress response (HSR) factors might have ability to alter m^{6}A through regulation of m^{6}A modifiers. HSF1 is the key factor to modulate HSR, and it was shown that m^{6}A might be involved in HSF1-mediated HSR. Depletion of HSF1 partially impaired heat-shock-stimulated expression of YTHDF2 (Zhou et al. 2015), indicating that YTHDF2 might act as a downstream effector of HSF1 during HSR. Since YTHDF2 directly bound to and

regulated the stability of m^{6}A-continaing RNA, HSF1-mediated alteration of YTHDF2 expression would alter the abundance and distribution of m^{6}A-continaing RNA. HSPs exert broad functions including post-translational regulation of target proteins by de novo folding and refolding of stress-denatured proteins, oligomeric assembly, intracellular protein transport, and assistance in protein stabilization or proteolytic degradation (Vabulas et al. 2010). Considering the discovery of WTAP stabilization by forming a complex with HSP90 (Kuai et al. 2018), it is possible that HSP90 alters the expression of m^{6}A modifier, WTAP, and subsequent m^{6}A levels upon HSR. Thus, heat stress-stimulated HSF1- and HSPs-mediating regulation of RNA modifiers partially explained the alteration of m^{6}A modifications upon heat shock/stress; nevertheless, the detailed mechanisms regarding m^{6}A modifiers' regulation by heat stress need to be further studied.

Heat shock/stress also influenced the distribution of RNA modification on a single RNA transcript. In normal condition, most m^{6}A modifications are located at the coding sequence (CDS), stop codon, and 3′-UTRs. Only a few m^{6}A sites are present in 5′-UTRs (Meyer et al. 2012). The different location of m^{6}A bases correlates with varying RNA fate. For instance, methylation at exons is involved in splicing, while the methylation near stop codons associates with translational control (Darnell et al. 2018; Meyer et al. 2012). Recent studies have confirmed that the distribution pattern of m^{6}A on transcripts was affected upon heat shock/stress (Meyer et al. 2015; Zhou et al. 2015). Although the m^{6}A modification appeared in 5′-UTR is much less than that in other regions within transcripts, it has been reported the dynamic changes of m^{6}A at 5′-UTR is a key regulator of HSR (Meyer et al. 2015; Zhou et al. 2015). Meyer et al. (2015) found a marked enrichment of m^{6}A in the 5′-UTR after heat shock in MEF and HepG2 cell lines, and the enrichment of m^{6}A at 5′-UTR initiated the cap-independent translation (Fig. 2). Consistently, Knuckles et al. (2017) observed a similar phenomenon that m^{6}A at 5′-UTR promotes cap-independent translation. An earlier study performed by Zhou et al. (2015) also showed an elevation of m^{6}A level at 5′-UTR by heat shock in HeLa and MCF cells, and they further demonstrated an overexpression and the translocation of YTHDF2 from cytosol to nucleus, which protected m^{6}A modification at 5′-UTR from FTO-mediated demethylation, leading to the characteristic re-distribution of m^{6}A within individual transcripts (Fig. 2). Therefore, the heat shock-YTHDF2-cap-independent translation axis indicated a potential mechanism by which heat shock/stress responsive genes can be preferentially expressed. However, how YTHDF2 specifically interacts with m^{6}A at 5′-UTR without influencing m^{6}A modification at other locations in the same transcript remains to be elucidated. In contrast to m^{6}A modification, m^{7}G was enriched in 5′-UTR and AG-rich contexts of mRNA under normal conditions, but it was remarkably accumulated in CDS and 3′-UTR region after heat shock to enhance translation efficiency (Malbec et al. 2019).

In addition to regulating the distribution of the modified sites on a single transcript, heat stress also regulated the subcellular localization of modified RNAs. It was demonstrated that heat shock induces METTL3 translocation from cytosol to the nucleus to facilitate the relocalization of DGCR8 and METTL3 to heat-shock genes for methylation of these genes (Knuckles et al. 2017) (Fig. 2). Two

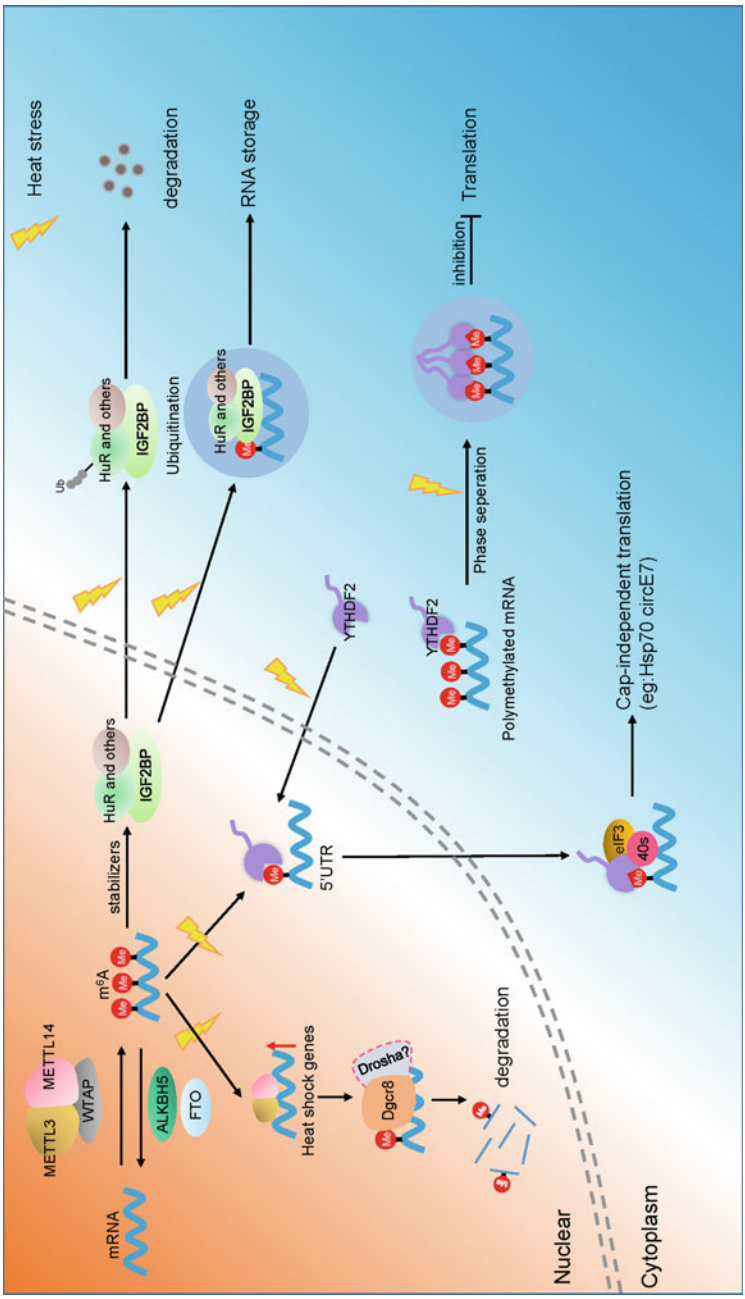


Fig. 2 Regulation of m6A modification by heat stress. mRNAs were de novo methylated by the "m6A writer" complex (METTL3, METTL14, and WTAP), and demethylated by the "m6A eraser" FTO or ALKBH5. The m6A-modified RNAs would be recognized by various "m6A readers" (such as IGF2BPs, YTHDFs) which regulated the fate of reader-bound m6A-containing RNAs. Heat stress affected both the expression level of and the distribution of m6A modifiers (writer, eraser, and reader), which in turn modulated the levels of m6A modification, the destitution of m6A-containgin RNAs, and the subsequent m6A-participated bioprocesses. Currently demonstrated heat stress-m6A regulatory pathways are listed here. Heat stress elevated m6A level at 5′-UTR, which prompted the cap-independent translation in a m6A-dependent way. Heat stress caused radical relocalization of METTL3 and DGCR8 to m6A-containing RNA for subsequent degradation. IGF2BPs-associated RNAs shuttled to stress granules after heat stress, which is a stress-stimulated RNA storage manner. Heat

independent research groups observed that heat stress induces the translocation of YTHDF2 from cytoplasm to nucleus, protecting the 5′-UTR m^6A from FTO-mediated demethylation; this YTHDF2-dependent fashion increased the levels of m^6A in nucleus (Zhou et al. 2015). Intriguingly, the study conducted by Ries et al. (2019) neither observed alteration of YTHDF2 expression nor detected any enrichment of YTHDF2 in the nucleus after heat shock; instead, they noted an enrichment of YTHDF2 in stress granules (SGs) following heat stress and found a ~ 40 kDa prion-like domain within YTHDF proteins, which was essential for the initiation of phase separation.

Stress granules (SGs) are a kind of RNA-protein condensates that assemble in the cytoplasm of cells in response to various environmental stresses including heat shock, virus infection, UV exposure, starvation, and so on. The formation of SGs is important for RNA fate determination. On one hand, SGs function to prioritize translation of recruited mRNAs; on the other hand, SGs also promote the abrogation of other mRNAs and RNA binding proteins from the harmful cellular environment and cease their translation. When cells were exposed to heat stress, poly-m^6A-methylated mRNAs would act as a multivalent scaffold to enable the binding of YTHDF proteins, leading to the formation of phase-separated YTHDF-m^6A-mRNA complexes which were subsequently partitioned into SGs, a kind of phase-separated structures in cells, and subsequently influenced the translational efficiency of m^6A-modified mRNAs (Fig. 2). Other studies (Gao et al. 2019; Liu et al. 2020; Wang et al. 2020b) regarding phase separation also drew same conclusions with Ries et al. (2019), and these works were summarized by Liu et al. (2020). In addition to YTHDF2, another group of proteins, IGF2BPs, which were confirmed as new m^6A reader family proteins by Huang et al. (2018), translocated into SGs in response to heat shock as well, in line with the previous studies (Stöhr et al. 2006; Wächter et al. 2013) (Fig. 2), which in turn regulated the stability and the translational efficiency of its target genes in a m^6A-dependent manner. Besides m^6A RNA modifiers, a recent study reported that the m^1A writer, TRMT6/61A methyltransferase, localized to stress granules (SGs) under heat shock, explaining the alteration of m^1A RNA level upon temperature stress and suggesting the involvement of m^1A in stress-induced granulation (Alriquet et al. 2020).

4.2 RNA Modifications Modulate HSR Factors

As discussed above, RNA modification could be regulated by heat stress in many ways, which, on the other hand, suggests that RNA modification might have a role in heat shock/stress response (HSR). Several independent studies found that *HSPs* and

Fig. 2 (continued) shock led to the formation of phase-separated YTHDF-m6A-mRNA complex that then partitioned into stress granules, and inhibits the translational efficiency of m6A-containing mRNAs

HSF1 transcripts are rich in m^{6}A sites (Dominissini et al. 2012; Feng et al. 2020; Meyer et al. 2015; Schwartz et al. 2014; Yu et al. 2018; Zhou et al. 2015). Miao et al. (2019) observed that the elevated translation of both HSP70 and HSP40 induced by heat shock was mediated by m^{6}A modification. Using m^{6}A-seq, Yu et al. (2018) reported the profiling pattern of m^{6}A modification in HepG2 cells. They found that the m^{6}A sites of *HSPA1B*, *HSPB1*, *HSPA9*, *HSP90AA1*, *HSPD1*, and *HSF1* transcripts were mainly distributed on exons and around stop codons, despite certain differences. Briefly, there was an enrichment of m^{6}A in both the 5′-UTR and 3′-UTR on the transcripts of *HSPA1B*, *HSPB1*, *HSPA9*, and *HSPD1*, while the enrichment was found only in the 5′-UTR in *HSP90AA1* transcript and only in the 3′-UTR in *HSF1* transcript (Yu et al. 2018). These results were similar with the latter study conducted on chicken (Feng et al. 2020), suggesting the conservation of m^{6}A-mediated regulation among different species.

Additionally, the different distributions and abundances of m^{6}A sites on *HSPs*' transcripts would lead to different effects of m^{6}A-mediated regulation of HSPs' expression. For instance, knockdown of YTHDF2 would significantly increase the mRNA expression of *HSP90AA1*, *HSPD1*, and *HSPB1*, but only have little influence on the expression of *HSPA1B* and *HSPA9* (Meyer et al. 2015; Miao et al. 2019; Zhou et al. 2015). Similarly, Meyer et al. (2012) found an enrichment of m^{6}A at the 5′-UTR of *HSP70*, while no m^{6}A site was found at the 5′-UTR of another heat shock gene *HSP25*. Furthermore, their results showed a dramatic heat-shock-induced translation of HSP70 in MEF cell after knockdown of FTO, while no change was observed on HSP25 level in the same condition, supporting the concept that the distribution of m^{6}A on HSPs' transcripts indeed affects m^{6}A-mediated regulation of HSPs' expression. Zhou et al. (2015) also showed that the induced expression of *HSPA1A*, a member of *HSP70* gene family, mainly resulted from the increased m^{6}A modification occurred at 5′-UTR of its transcript upon heat shock. On the contrary, another *HSP70* gene *HSPA8*, lack of m^{6}A site at 5′-UTR, showed only minor increase in both the mRNA and protein levels upon heat shock. In addition, they further confirmed that the enhanced translation of *HSP70* was resulted from the overexpression and translocation of YTHDF2, which was discussed above. Together, m^{6}A modifications occurred on and played crucial roles to regulate *HSPs*' transcripts.

Intriguingly, although Knuckles et al. (2017) similarly found an elevation of m^{6}A modification on *HSP70* in post heat shock mouse embryonic stem (mES) cells, a rapid degradation of *HSP70* mRNA rather than an increase in translation efficiency was noted following the elevation of m^{6}A modification. They observed that heat shock induced a radical relocalization of DGCR8 and METTL3 from miRNA, snoRNA, and protein-coding genes to the major heat-shock genes in mES cells for the subsequent degradation of these transcripts. Combination of these studies, upon heat shock, the enhancement of m^{6}A modification at 5′-UTR increased the translation of *HSP70* in a cap-independent manner, which guaranteed the production of sufficient amounts of HSP70 protein in response to heat shock (Zhou et al. 2015). In addition, when heat stress is ceased (post heat shock period), the m^{6}A-modified

HSP70 transcripts underwent rapid degradation to avoid the aberrant accumulation of *HSP70* mRNA (Knuckles et al. 2017).

Apart from m^6A, other RNA modifications also participated in heat shock response. For example, modified RNA could modulate the transportation of *HSP*s mRNA upon heat shock. Mobility of transcripts in graft plants plays a critical role to regulate development and growth. When m^5C methyltransferases *dnmt2* (*trdmt1*) and *nsun2b* (*trm4b*) were double-knockout, transport of the endogenously produced heat shock cognate protein (HSC70.1) was impaired, indicating the importance of m^5C modification in mRNA transport (Yang et al. 2019a). Taken together, these findings display that the dynamic regulation of RNA modification on HSR is a precise, elaborate, and complicated process, which merits further exploration.

4.3 RNA Modification Participates in HSR-Mediated Bioprocesses

Increasing evidence showed the involvement of RNA modification in HSR-mediated bioprocesses. For instance, the m^5C methyltransferase Dnmt2-mediated tRNA stability and DNA repeat integrity in *Drosophila* were mostly heat shock dependent (Genenncher et al. 2018); the internal m^7G enrichment increased downstream *PCNA* mRNA translation efficiency under heat stress (Malbec et al. 2019); m^6A modification at 5′-UTR could be induced by heat shock, leading to cap-independent translation of HSPs' mRNA. Besides HSPs, It was found that the cap-independent translation of circRNAs is enhanced following heat shock in a m^6A-depenednt manner (Yang et al. 2017).

Heat stress directly and indirectly regulates the distribution and expression of RNA modifiers, which subsequently regulate RNA modifications and closely relate to many bioprocesses. For example, WTAP played important roles in regulating cell proliferation and differentiation, and was considered as an oncogene in leukemia; YTHDC1/2 took part in mouse germline development and they were vital for germ cell maturation; the energy homeostasis regulated by FTO were associated with type 2 diabetes mellitus (T2DM) and obesity; these modifier-associated functions were well-summarized in a review by Chen et al. (2019b). Here we focus on the biological effects resulting from the axis of HSR-RNA modifiers. Cellular m^6A-modified mRNA (e.g., *MYC*) could be recognized and stabilized by IGF2BPs through HUR in a m^6A-dependent fashion. However, elevating temperature decreased the stability of these mRNAs and led to its degradation by translocated IGF2BPs into SGs (Huang et al. 2018). Since IGF2BPs regulate various genes, such as *FSCN1*, *TK1*, and *MARCKSL1*, which are involved in multiple biological processes including DNA replication, cell cycle, and cell proliferation, heat stress-mediated relocalization of IGF2BPs would affect the bioprocesses exerted by IGF2BPs-bound RNA. Additionally, m^6A modification could regulate the translation efficiency directly through the recruitment of eIF4G2 by YTHDF3 (Shi et al. 2017).

Since YTHDF3 was affected by HSR, it is very likely that HSR could modulate translation via a YTHDF3-dependent manner (Zhou et al. 2015). Hence, HSR-mediated alteration of m^{6}A modifiers (e.g., IGF2BPs, YTHDF3) might in turn regulate downstream bioprocesses to affect cellular homeostasis.

Previous studies found that the overall level of m^{6}A modification was elevated following heat shock. Recently, it was demonstrated that the m^{6}A levels in stress-granule mRNA were higher than that in total cellular mRNA after heat shock, indicating that heat stress could trigger triaging of m^{6}A-modified mRNA to stress granules (SGs) (Ries et al. 2019). The authors took poly-methylated mRNAs *FIGNL1* and *FEM1B* and non-methylated mRNAs *GRK6* and *POLR2A* as study objects and found that mRNAs with m^{6}A modification are substantially enriched in SGs, while non-methylated mRNAs had no enrichment in SGs upon heat shock. Another study discovered that the relocalization of m^{6}A-modified mRNAs into SGs is region-selective (Anders et al. 2018). They found the m^{6}A-modified mRNA, *ARL4C*, was recruited to SGs following heat shock. Interestingly, deleting the methylation sites of 5′-vicinity in coding sequences of *ARL4C* blocked its colocalization with SGs, suggesting the importance of the region-specific m^{6}A modification for the localization of m^{6}A-containing RNA. It is well known that SGs are repositories for the storage of mRNAs under stress and influence the fate of mRNAs. During a stress-recovery period, some intact mRNAs stored in SGs reenter translation via a process facilitated by chaperones such as HSP101, while other mRNAs released from SGs may be subjected to degradation (Chantarachot and Bailey-Serres 2018). Therefore, future studies are strongly recommended to figure out which mRNAs are recruited into SGs in the m^{6}A-dependent manner and how about their fate after stress.

On the other hand, investigating the HSR-stimulated m^{6}A-mediated bioprocesses might benefit clinical practice. Currently, hyperthermia is proven as an effective adjuvant to chemotherapy or radiotherapy during cancer treatment (Soares et al. 2012), while there is limited evidence regarding the mechanism. Whether modified RNA plays a role in hyperthermia merits further investigation.

Collectively, we summarized the mutual regulation between RNA modification and HSR in this part. Broadly speaking, heat stress directly or indirectly regulates the distribution and abundance of RNA modification mainly by affecting the distribution and expression of the RNA modifiers. The alteration of RNA modification induced by heat shock not only influenced the expression of HSPs but also affected various downstream bioprocesses.

5 Starvation

Starvation is one of the most physiologically relevant stresses associated with severe nutrient (e.g., carbohydrates, amino acids, and trace elements) deprivation, which could be induced by fasting, fasting mimicking diet (FMD), or specific material or element depletion. The conditions of starvation are often sensed by the protein

named target of rapamycin (TOR), which mainly functions in regulating the synthesis and metabolism of proteins inside the cell. In the absence of nutrients, the activity of TOR is inhibited, leading to the induction of autophagy in the cell. Subsequently, the activated autophagy machinery by starvation accelerates the degradation and recycling of specific cellular components and provide essential material for starvation adaptation and cell survival, while severe starvation could lead to cell death (Oliva Trejo et al. 2020).

Recently, it is reported that increased m^6A modification by knockdown of m^6A eraser, FTO, negatively regulates ATG5 and ATG7 expression through a YTHDF2-dependent manner to control autophagy (Wang et al. 2020a), indicating a potential role of RNA m^6A modification in starvation response. Consistently, Yang et al. (2019b) reported that the m^6A eraser FTO is upregulated by serum starvation (0.2% FBS), decreasing m^6A level in melanoma cells. Intriguingly, FTO expression might be induced through both autophagy and NF-κB pathways, since knockdown of ATG5, ATG7, or the NF-κB subunit p65 blocked starvation-induced FTO expression (Yang et al. 2019b). These findings suggest a feedback control loop between autophagy induction and FTO expression, and FTO probably acts as a downstream effector of starvation-induced autophagy.

5.1 RNA m^6A Modification Modulates Cellular Response to Starvation

Activating transcription factor 4 (ATF4), a well-defined regulator modulating the transcription of key genes pivotal for adaptative functions, is an important mechanism of cellular response to stresses including starvation. Conventionally, it is believed that upregulation of ATF4 during starvation occurs via increased translation efficiency of the ATF4 mRNA upon eIF2α phosphorylation (Suragani et al. 2012), albeit the precise mechanism remains elusive. Activated ATF4 turns on expression of several downstream genes such as C/EBP Homologous Protein-10 (CHOP), to help cells cope with various stresses. Zhou et al. (2018) reported that apart from the eIF2α signaling pathway, m^6A modification also regulates ATF4 expression following starvation. They found that the m^6A demethylase ALKBH5 is appreciably increased in MEF cells upon amino acid starvation, resulting in reduced m^6A methylation of 5′-UTR (upstream open reading frame 2) on *ATF4* mRNA, controlling start codon selection and promoting translation of the downstream *ATF4* main CDS. Proving this notion, knockdown of either the m^6A writers METTL3/14 or the erasers ALKBH5/FTO did not affect phosphorylated eIF2α or *ATF4* mRNA level, but markedly influenced ATF4 translation. Collectively, these findings shed light on the importance of mRNA m^6A modification during translational regulation upon starvation and suggest that RNA m^6A modification is a vital part of starvation-response circuitry.

5.2 *RNA m^{6}A Modification Contributes to Intracellular SAM Level Maintenance During Methionine Starvation*

S-adenosylmethionine (SAM) is the well-known methyl donor for almost all cellular methylation reactions including DNA methylation, histone methylation, and RNA m^{6}A modification (Green et al. 2019). Faithful maintenance of SAM level is necessary to sustain normal bioprocesses and cellular homeostasis. SAM is generally produced from methionine (Met) by SAM synthetase Met adenosyltransferase (MAT) in the presence of ATP (Wang and Breaker 2008). MAT includes homologous enzymes MAT1A and MAT2A. MAT1A is only expressed in adult liver cells, and MAT2A is ubiquitously expressed in all cells apart from liver cells (Wang and Breaker 2008). It was reported that MAT2A expression was facilitated through mRNA stabilization upon SAM depletion (induced by Met starvation or MAT inhibition), but the mechanism is not yet fully elucidated (Martínez-Chantar et al. 2003).

Recently, two research groups independently demonstrated that m^{6}A writer METTL16 is involved in MAT2A protein upregulation following SAM depletion by Met starvation or MAT inhibitor cycloleucine (cLEU) treatment (Pendleton et al. 2017; Shima et al. 2017). The *MAT2A* gene encodes a cytoplasmic isoform (*MAT2A*) and a nuclear-retained intron isoform (*MAT2A-RI*). Pendleton et al. (2017) found that upon Met starvation, *MAT2A-RI* mRNA is rapidly decreased following appreciable increase in *MAT2A* mRNA due to altered splicing but not mRNA transcription rates. Specifically, Met starvation induced SAM depletion, leading to prolonged METTL16 occupancy of the 3′-UTR and m^{6}A modification at the A4 position of hairpin 1 (UACAGAGAA) in *MAT2A* and inhibiting intron retention during pre-mRNA splicing, so as to increase cytoplasmic *MAT2A* mRNA and protein level to upregulate SAM production. Consistently, Shima et al. (2017) observed that m^{6}A modification in the 3′-UTR of *MAT2A* by *METTL16* and its recognition by the reader YTHDC1 is essential for cLEU treatment- (SAM depletion) induced *MAT2A* expression. Together, these studies unveiled the elaborate mechanism of intracellular SAM monitoring and maintenance by m^{6}A modification.

5.3 *Starvation, RNA m^{6}A Modification, and Cancer*

The role of starvation in cancer progression and treatment remains largely elusive (Raffaghello et al. 2008; Buono and Longo 2018; Niu et al. 2019; Yang et al. 2019b). Recent studies suggested m^{6}A modification might play a role in starvation-mediated malignance of cancer. In melanoma cells, starvation-induced FTO upregulation led to increased PD-1 expression through decreasing m^{6}A modification, thereby promoting melanoma tumorigenesis and anti-PD-1 therapy resistance (Yang et al. 2019b). Similarly, upregulation of FTO expression in breast cancer led to promotion of tumor progression through inhibiting the tumor suppressor BNIP3, a responder to

amino acid starvation (Niu et al. 2019). On the other hand, starvation could be beneficial for cancer treatment in different context. Cancerous cells present upmost desire for nutrients due to rapid and abnormal self-renewal as well as metastasis; thus depletion of nutrients or simply fasting has been deemed an effective approach to reduce the risk factors or reverse the symptoms of cancers. For example, compared to untreated cells, short-term starved cancer cells are much more sensitive to chemotherapy, which is called differential stress resistance (DSR) (Buono and Longo 2018; Raffaghello et al. 2008).

As a whole, these studies showed that starvation alters mRNA m^6A modifiers and modification, which affected the fate of m^6A-containing RNA and downstream cellular processes. However, due to the diversity of m^6A RNA fate regulated by distinct m^6A modifiers, and the double-edged effects of starvation in cancer, further studies will be needed for understanding the association among m^6A modification, starvation, and treatment response in different scenarios.

6 Hypoxia

Hypoxia is a phenomenon of lower oxygen level in the whole or specific region of the body than normal. As oxygen is important in various life activities, insufficient oxygen supply will cause abnormal cellular metabolism, vascular and pulmonary disease, cancers, and so on. Hypoxia-inducible factors (HIFs) are transcription factors which promote expression of specific genes to regulate cellular response to hypoxia, and they belong to the basic helix-loop-helix (HLH)-PER-ARNT-SIM (bHLH-PAS) protein family (Bersten et al. 2013). HIFs are heterodimeric complexes composed of two subunits, the α subunit (HIF-α) and β subunit (aryl hydrocarbon receptor nuclear translocator, ARNT) (Jaakkola et al. 2001). Under normoxic conditions, HIF-α is hydroxylated by prolyl-4-hydroxylases (PHDs) which requires molecular oxygen as a substrate (Schofield and Ratcliffe 2004), and then the hydroxyl signal is recognized by ubiquitin E3 ligase von Hippel-Lindau (VHL) to mediate ubiquitin-proteasome degradation of HIF-α (Flügel et al. 2007), while in hypoxia conditions, hydroxylation of HIF-α is impeded due to the lack of oxygen, leading to dimerization of stable HIF-α with ARNT to activate transcription of target genes. In mammals, HIF-α has three isoforms, including HIF-1α, HIF-2α, and HIF-3α. Some hypoxia-responsive genes, such as erythropoietin (EPO) (Kuhrt and Wojchowski 2015), vascular endothelial growth factor (VEGF) (Nicolas et al. 2019), and platelet-derived growth factor (PDGF) (Xiao et al. 2017), have hypoxia-responsive elements (HREs) in their promoter regions. HIF-1α can recognize HREs and regulate expression of these genes upon hypoxia.

It was reported that hypoxia correlates with epigenetic modification. For example, in gastric cancer, hypoxia silenced tumor suppressor RUNX3 through histone di-methylation and H3 deacetylation (Batie et al. 2019), while another study reported that hypoxia reduces ten-eleven translocation (TET) enzymes expressed in cancer which catalyze DNA m^5C demethylation to activate tumor suppressor genes

(Thienpont et al. 2016). Compared to the well-known epigenetic modification, epitranscriptomic modification is a relatively new aspect of hypoxia field. However, recently, increasing number of studies investigated the role of RNA modifications in response to hypoxia, and here we would focus on discussion of m^{6}A.

6.1 *Hypoxia Regulates RNA m^{6}A Modification and m^{6}A Modifiers*

Recent works indicated hypoxia regulates m^{6}A levels in a context-dependent manner. For example, Fry et al. (2018) established two kinds of human mammary epithelial cells (HMECs) as breast cancer progression model and found that mRNA m^{6}A levels of both genetically defined immortalized HMECs and oncogenically transformed HMECs increase by hypoxia. These results are consistent with their earlier study (Fry et al. 2017), which revealed that hypoxia increases mRNA m^{6}A level in HEK293 cells. Pan et al. (2020) also showed upregulation of m^{6}A levels of wheat root in response to hypoxia, indicating hypoxia also plays a regulatory role in m^{6}A modification of plants . However, other studies found hypoxia decreasing m^{6}A levels in breast cancer cells and lung adenocarcinoma cells, respectively (Zhang et al. 2016a; Chao et al. 2020). Chao et al. (2020) also found that the levels of m^{6}A in lung adenocarcinoma cells and nude mice subcutaneous tumor tissues were both decreased under intermittent hypoxia. Collectively, since hypoxia-mediated m^{6}A modification showed cell context-dependent manner, further investigation will be needed to understand the detailed regulatory mechanism in each scenario.

Similar to the response to other stresses, dynamic changes of RNA modification play important roles in hypoxia stress as well, and these changes might be mediated by altering expression of RNA modifiers via hypoxia. Indeed, several studies observed that hypoxia would lead to upregulation of m^{6}A erasers and downregulation of m^{6}A writers. It was found that ALKBH5 is the target of HIF, and the expression of ALKBH5 is induced by sustained hypoxia in a HIF-1α- and HIF-2α-dependent fashion (Zhang et al. 2016b). Consistently, certain more recent studies also observed upregulation of ALKBH5 by hypoxia (Zhang et al. 2020b; Panneerdoss et al. 2018). In addition, a transcription factor, ZNF217, which blocks METTL3 methyltransferase activity, was also induced in a HIF-dependent manner by hypoxia in breast cancer cells (Fry et al. 2018; Zhang et al. 2016b). Conversely, under different context, hypoxia showed opposite effects to upregulate m^{6}A writers. For example, Lin et al. (2020a) discovered that hypoxia facilitates adipose-derived stem cell (ADSC) differentiation into smooth muscle cells (VSMCs) through upregulation of METTL3. Another component of m^{6}A writer complex, METTL14, was also upregulated under hypoxic conditions in breast cancer cells (Panneerdoss et al. 2018). Besides m^{6}A writer and eraser, hypoxia could alter m^{6}A levels via affecting m^{6}A readers, however, still in a context dependent manner. Since

animals in high altitude regions are prone to hypoxia, Shi and colleagues compared genes involved in carcinogenesis between six Tibetan mammals (dog, horse, pig, cattle, sheep, and goat) and corresponding species in low altitude and found YTHDF1 mRNA expression decreased in liver and kidney tissue of Tibetan cattle, while YTHDF2 and YTHDF3 had no changes (Shi et al. 2019). They depleted YTHDF1 in normal human bronchial epithelium cells (BEAS-2B) and found the cells are resistant to hypoxia-induced apoptosis, indicating low expression of YTHDF1 associates with hypoxic adaption. However, in other studies, YTHDF2 expression has been reported to be abrogated by hypoxia in hepatocellular carcinoma (HCC) cells and promoted HCC exacerbation (Hou et al. 2019; Zhong et al. 2019). Hypoxia-mediated downregulation of YTHDF3 was observed in breast cancer cells (Panneerdoss et al. 2018). These findings suggest that hypoxia regulates "writers," "readers," and "erasers" of RNA m^6A modification in a context-dependent manner, which might explain the controversial effects of hypoxia on regulation of m^6A levels.

6.2 *RNA m6A Modification Regulates Hypoxia-Inducible Factors*

RNA m^6A modification also showed the ability to participate in and modulate hypoxia response via regulating hypoxia-inducible factors (HIFs). For example, HIF-α subunits especially HIF-1α are the key transcriptional factor regulating gene expression during the bioprocess of human and other mammals' response to hypoxia, and several studies demonstrated that m^6A could regulate the function and expression of HIF-1α. First, the m^6A modification would stabilize *lnc-Dpf3* mRNA, and lnc-Dpf3 could directly bind to HRE motif of HIF-1α downstream genes to inhibit the transcription of HIF-1α-activated glycolytic-related gene, *LINHA*, and suppress HIF-1α-mediated dendritic cell migration (Liu et al. 2019a). The E3 ligase Von Hippel-Lindau (VHL) is categorized as a tumor suppressor, which promotes HIF-1α ubiquitination and proteasomal degradation under normoxia. Zhu et al. (2019) found that VHL promotes follicular helper T cell (Tfh) development via VHL-HIF-1α-ICOS (inducible costimulator) axis with m^6A involvement. Briefly, depletion of VHL would increase the expression of METTL3/14, which methylated the transcripts of *ICOS* to decrease its expression, and then led to negative regulation of Tfh development. These findings revealed the involvement of m^6A modification in HIF-1α mediated glycolysis pathways.

The m^6A readers also play important roles for m^6A-mediated regulation of HIF-α mRNA function and stability. For example, Tanabe et al. (2016) demonstrated that m^6A-modified *HIF-1α* is a target of YTHDC2. They found YTHDC2 influenced the translation efficiency without affecting the transcription of HIF-α mRNA under hypoxia, and demonstrated that 5′-UTR unwinding by YTHDC2 is important for HIF-1α translation (Tanabe et al. 2016). Methylenetetrahydrofolate dehydrogenase

2 (MTHFD2) is a mitochondrial enzyme involved in one-carbon metabolism. Recently, it was reported that MTHFD2 contributes to the progression of renal cell carcinoma (RCC) via a novel epitranscriptomic mechanism that involves HIF-2α (Green et al. 2019). Authors found that MTHFD2 was upregulated in RCC, which facilitates replenishing of intracellular SAM and promotes global m^{6}A RNA levels, including the m^{6}A methylation of *HIF-2α* mRNA, thereby results in enhanced translation of *HIF-2α*. Subsequently, the enhanced expression of HIF-2α promotes the aerobic glycolysis, which would enhance hypoxic condition of tumor microenvironment and promote RCC progression. Taken together, RNA m^{6}A modification directly or indirectly modulates the function and expression of HIFs.

6.3 RNA m6A Modification, HIFs, and Cancer

RNA m^{6}A modification not only regulates HIFs' function and expression but also is involved in both short-term and long-term effects of hypoxia. For example, hypoxia upregulates METTL3 to catalyze m^{6}A modification of paracrine factors, including *VEGF*, *HGF*, *TGF-β*, *GM-CSF*, *bFGF*, and *SDF-1*, thereby facilitating the differentiation of adipose-derived stem cells (ADSCs) into smooth muscle cells (VSMCs) (Lin et al. 2020a). YTHDF2 transcription is regulated by HIF-2α in human hepatocellular carcinoma (HCC), and this reader protein mediated the degradation of m^{6}A-modified *IL11* and *SERPINE2* mRNAs, which were responsible for the inflammation-mediated malignancy and disruption of vascular normalization (Hou et al. 2019). These findings suggest that RNA m^{6}A machinery would cooperate with HIFs to mediate various HIF-associated bioprocesses.

Hypoxia is a typical feature of tumor microenvironment, and HIFs increase vascularization around tumor to ensure adequate blood and oxygen supply as well as promote invasion and metastasis of neoplasms. In breast cancer cells, it was found that hypoxic stress induces m^{6}A modification, leading to stabilization of several mRNAs, including *GLUT1*, *JUN*, *MYC*, *VHL*, and *DUSP*1(Fry et al. 2017; Fry et al. 2018), and upregulation of their protein expression. The m^{6}A levels of these specific mRNAs were controlled by HIFs and promoted breast cancer to a more aggressive phenotype. Another study documenting the role of hypoxia in breast cancer reported that ALKBH5, a target of HIF, inhibited m^{6}A modification of *NANOG* mRNA and increased its stability and protein expression, thereby leading to the generation of breast cancer stem cell (BCSC) phenotype (Zhang et al. 2016a). In addition, METTL3 is significantly downregulated in HCC to promote sorafenib resistance. Mechanistically, low level of METTL3 under hypoxic conditions leads to decreased m^{6}A modification of *FOXO3* mRNA at 3′-UTR and decreases its YTHDF1-dependent stabilization to promote sorafenib resistance in cultured HCC cells (Lin et al. 2020b). In other cancers, such as renal cell cancer and lung adenocarcinoma, ALKBH5 respectively regulated mRNA levels of *AURKB* and *FOXM1* through decreasing m^{6}A modification to promote cancer proliferation and invasion under hypoxia (Chao et al. 2020). The m^{6}A reader YTHDC2 enhanced HIF-1α translation

efficiency inducing colon cancer metastasis in hypoxic conditions (Tanabe et al. 2016). Thus, RNA m^6A modification is closely associated with cancer initiation, progression, metastasis, and drug-resistance under hypoxic conditions. As m^6A is a newly emerging regulator of hypoxia response, further examinations are needed to get profound understanding for exploiting novel treatment approaches of hypoxia-related cancers.

7 Other Stresses

Apart from the abovementioned stresses, cells are subjected to other insults such as heavy metals, oxidative stress, mental stress, and various toxic agents including environmental pollutants. In some cases, it is hard to differentiate the cellular response of one stress from the other since they are interconnected (one stress may elicit one or more stress response). For instance, the heat shock response (HSR), an evolutionarily conserved cell protective mechanism, is often activated upon various protein damage causing stresses (not limited to heat shock), to maintain protein homeostasis in basically all eukaryotic species. Since the mechanisms of some stress-induced outcomes are currently unknown, understanding the post-transcriptional gene regulatory aspects of these stresses might fill in this "gap."

Environmental hazardous substances including various gaseous compounds and particulate matters cause acute or chronic threat to the health of humans and other animals. As one of the regulators of gene expression at post-transcriptional level, RNA m^6A modification is also implicated during the adverse effects of these compounds. For instance, exposure to endocrine disrupting chemicals (EDCs) is closely related to induction of lipid-metabolism diseases including obesity, nonalcoholic fatty liver disease (NAFLD), and others. Chen et al. (2019a) compared the effects of three EDCs (triclosan, bisphenol A, and fluorene-9-bisphenol) on lipid metabolism in zebrafish. They found that EDCs exposure led to a significant decrease in global m^6A levels and abnormal expression of m^6A modifiers in larvae, and this might be associated with dysregulation of genes' expression controlling lipid metabolism. Cobalt (Co) is believed to show adverse impact on the nervous system, but its detailed mechanisms are yet to be determined. A recent report used CL57BL/6 mice and human neuroblastoma H4 cells as study objects and found that $CoCl_2$ exposure reduced global RNA m^6A modification in the cerebral cortex by downregulating expression of m^6A writer proteins METTL3, METTL14, as well as WTAP and increasing expression m^6A erasers FTO and ALKBH5, and thereby caused disorder in the expression of neurodegenerative diseases-associated genes (Chen et al. 2019a). These findings not only revealed the mechanism behind neurotoxicity of Co but also provided a research direction for decoding the pathways of heavy metal-induced adverse health effects.

Arsenite is not only an effective drug for acute promyelocytic leukemia (APL) treatment but also a common carcinogen widely distributed in the environment (Maimaitiyiming et al. 2020). The double-edged effects of arsenite is termed as

hormesis, and it is reported that m^{6}A modification participates in this process (Chen et al. 2019a). Low concentration of arsenite increased METTL3/METTL14/WTAP expression and inactivated FTO in human keratinocytes, promoting m^{6}A modification and exerting cytoprotective effects against oxidative stress, while high doses of arsenite reduced m^{6}A modification and enhanced oxidative stress, showing inhibitive effects on cell viability and proliferation (Chen et al. 2019a). Another study also showed that arsenite-induced oxidative stress resulted in augmented m^{6}A methylation through elevated expression of WTAP and METTL14 in human keratinocytes (Zhao et al. 2019). Together, these data demonstrate that "epitranscriptomics" hold the key for resolving currently unknown mechanism of environmental toxicants.

Oxidative stress is a critical contributor during pathogenesis of many diseases, and m^{6}A plays an important role in response to this threat as well. Generally, the imbalance between free radicals (e.g., reactive oxygen species) and antioxidants in the cells or organisms caused oxidative stress. It has been reported that colistin-induced oxidative stress reduced METTL3 expression leading to a decrease in m^{6}A level of pre-miR-873-5p in mouse renal tubular epithelial cells, thereby inhibiting pre-miR-873 recognition by DGCR8, attenuating the production of mature miR-873-5p and hindering activation of Keap1/Nrf2 pathway (Wang et al. 2019a). Thus, reversing METTL3 expression is a feasible approach to cope with colistin-induced oxidative stress. Anders et al. (2018) found that transcripts are additionally m^{6}A modified in their 5′-UTR as well as 5′ vicinity of CDS, and the m^{6}A reader YTHDF3 mediates triaging of these m^{6}A-containing transcripts to SGs upon oxidative stress, while the fate of these m^{6}A-modified mRNAs in SGs is yet to be determined. Di-(2-ethylhexyl) phthalate (DEHP) is also one of the common EDCs that may induce male reproductive disorder, but the underlying mechanisms remain unclear. A recent study conducted in rats showed that DEHP increases global levels of m^{6}A RNA modification and alters the expression of *FTO* and *YTHDC2*; meanwhile, m^{6}A modification of *NRF2* mRNA also increased which might be responsible for downregulation and hindered activation of NRF2 antioxidant pathway in prepubertal testes. These findings provided novel insight into testicular toxicity of DEHP from the perspective of m^{6}A modification and oxidative stress imbalance (Zhao et al. 2020). Taken together, the cellular response to oxidative stress is also regulated by m^{6}A modification, and its elaborate mechanism merits further investigation.

Accruing evidence suggests that stress exposure evokes various responses in the brain, affecting its functions such as hearing, learning, cognition, emotion, memory, and so forth (Pardon and Marsden 2008; Ma et al. 2015; Joëls et al. 2018). Mechanically, these stresses alter gene expression in neural cells to modulate synaptic development and connectivity. Thus, revealing the regulatory mechanisms of gene expression by various stressors is crucial for understanding how they affect brain functions. Recent studies revealed that m^{6}A modification on mRNAs was involved in the events like neuro-development, embryonic neural stem cell self-renewal, hippocampus-dependent learning memory, striatal function and learning (Koranda et al. 2018; Li et al. 2018; Shi et al. 2018; Wang et al. 2018; Yoon et al. 2018; Chen et al. 2019a;), suggesting the importance of m^{6}A in stress-response mechanism of neurons and brain. Engel et al. (2018) reported that acute restraint

stress or glucocorticoids (e.g., dexamethasone) altered mRNA m^6A modification in a brain region-, time-, and gene-specific manner. Knockout of m^6A writer METTL3 or silencing the m^6A eraser FTO in adult neurons affected m^6A landscape, changed transcriptome response to fear and synaptic plasticity to alter fear memory. Authors also demonstrated that major depressive disorder (MDD) patients present impaired regulation of m^6A following glucocorticoid stimulation, suggesting fine tuning m^6A machinery as a potential treatment for MDD patients (Engel et al. 2018). Other brain function abnormalities apart from MDD might also associate with damaged m^6A machinery, which merit further examination.

8 Discussion and Conclusion

Organisms evolved multiple response mechanisms to cope with various sources of stresses (internal and external). Epigenetic machineries such as DNA methylation and histone modifications are well-characterized regulators of gene expression in response to stresses (Dowen et al. 2012). Recently, RNA modifications including the m^6A are widely implicated in normal physiological processes and stress-response mechanisms (Engel et al. 2018). Owing to prompt response kinetics and potential for both rapid and chronic effects, regulation of mRNA function at post-transcriptional level is proven more and more crucial under stress conditions. While the research regarding m^6A and stress-response is still at early stage, future studies are needed to profoundly understand the role of m^6A in stress response, so as to pave a way for exploiting novel treatment options against stress-associated diseases and prevent stress-induced adverse health outcomes.

Considering the interaction between RNA modifications and various stresses, we found that although differences exist among modified-RNA-mediated responses to distinct stresses, there are certain general features too. Taking m^6A, for example, their common characteristics include a redistribution of m^6A-containing RNAs in cells, relocalization of m^6A sites on an individual transcript, enhancement of m^6A-dependent translation, etc. As to the differences, first and foremost, although m^6A is closely related to the relocalization of m^6A readers into SGs, distinct readers associate with m^6A-containg RNA in different stress conditions (Stöhr et al. 2006; Wächter et al. 2013; Knuckles et al. 2017; Huang et al. 2018; Gao et al. 2019; Ries et al. 2019; Liu et al. 2020). Second, although m^6A-dependent translations are enhanced upon several stresses, they are induced by different m^6A modifiers upon distinct cellular context. For instance, the increase of translation resulted from heat shock mainly occurred by a cap-independent manner, which is mediated by the recognition of m^6A marks deposited at the 5′-UTR region of transcripts by YTHDF2 (Zhou et al. 2015). In contrast, the arsenite treatment induced the interaction between YTHDF1 and m^6A-containing mRNAs to promote their translation, and these interactions occurred at multiple m^6A sites, such as 3′-UTR, 5′-UTR, and coding regions (Wang et al. 2015).

One thing should be noted. The antibody-based assays and methods are wieldy used in the m^6A field, but due to the technical limitations, certain so-called "m^6A" signals around 5′-UTR might be not from m^6A modification. At the 2′-hydroxyl position of the first, and sometimes the second, nucleotide adjacent to the m^7G cap (5′ cap) of mRNA, m^6A might bear additional 2′-O methylation as well, which is called N^6, 2′-O-dimethyladenosine (m^6A_m) and generally associated with increased stability of transcripts (Engel et al. 2018). Since both m^6A and m^6A_m could be recognized by m^6A antibody, the m^6A signals from antibody-based assays would contain m^6A_m modification. Thus, the antibody-independent m^6A analyses, such as direct RNA sequencing, could be an alternative way to investigate m^6A profiling.

Cell fate is determined by its adaptive capacity following exposure to numerous stresses, and the dynamic nature of RNA modification contributes the ability of organisms for responding to stressful stimuli and protects cells from various insults. Thus, better understanding of the role of RNA modifications in different stress-response mechanisms is crucial for coping with stress-associated adverse effects and providing novel insights for the treatment of stress-related disorders.

Acknowledgments This work was supported by the grant from the National Natural Science Foundation of China (31972883, 82000155). Meanwhile, we are very grateful to our lab member Mr. Jiebo Lin for optimizing the figures in this chapter.

References

Abakir A, Giles TC, Cristini A et al (2020) N6-methyladenosine regulates the stability of RNA: DNA hybrids in human cells. Nat Genet 52:48–55

Alriquet M, Calloni G, Martínez-Limón A et al (2020) The protective role of m1A during stress-induced granulation. J Mol Cell Biol 12(11):870–880. https://doi.org/10.1093/jmcb/mjaa023

Anckar J, Sistonen L (2011) Regulation of HSF1 function in the heat stress response: implications in aging and disease. Annu Rev Biochem 80:1089–1115

Anders M, Chelysheva I, Goebel I et al (2018) Dynamic m 6 a methylation facilitates mRNA triaging to stress granules. Life Sci Alliance 1:e201800113

Baquero-Perez B, Antanaviciute A, Yonchev ID et al (2019) The Tudor SND1 protein is an m6A RNA reader essential for replication of Kaposi's sarcoma-associated herpesvirus. elife 8:e47261

Batie M, Frost J, Frost M et al (2019) Hypoxia induces rapid changes to histone methylation and reprograms chromatin. Science 363:1222–1226

Bersten DC, Sullivan AE, Peet DJ et al (2013) bHLH-PAS proteins in cancer. Nat Rev Cancer 13:827–841

Buono R, Longo VD (2018) Starvation, stress resistance, and cancer. Trends Endocrinol Metab 29:271–280

Cadet J, Douki T (2018) Formation of UV-induced DNA damage contributing to skin cancer development. Photochem Photobiol Sci 17:1816–1841

Chantarachot T, Bailey-Serres J (2018) Polysomes, stress granules, and processing bodies: a dynamic triumvirate controlling cytoplasmic mRNA fate and function. Plant Physiol 176:254–269

Chao Y, Shang J, Ji W (2020) ALKBH5-m6A-FOXM1 signaling axis promotes proliferation and invasion of lung adenocarcinoma cells under intermittent hypoxia. Biochem Biophys Res Commun 521:499–506

Chen H, Zhao T, Sun D et al (2019a) Changes of RNA N6-methyladenosine in the hormesis effect induced by arsenite on human keratinocyte cells. Toxicol In Vitro 56:84–92
Chen X-Y, Zhang J, Zhu J-S (2019b) The role of m6A RNA methylation in human cancer. Mol Cancer 18:103
Chen H, Yang H, Zhu X et al (2020) m5C modification of mRNA serves a DNA damage code to promote homologous recombination. Nat Commun 11:2834
Courtney DG, Kennedy EM, Dumm RE et al (2017) Epitranscriptomic enhancement of influenza a virus gene expression and replication. Cell Host Microbe 22:377–386.e5
Courtney DG, Tsai K, Bogerd HP et al (2019) Epitranscriptomic addition of m5C to HIV-1 transcripts regulates viral gene expression. Cell Host Microbe 26:217–227.e6
Cui Q, Shi H, Ye P et al (2017) m6A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. Cell Rep 18:2622–2634
Darnell RB, Ke S, Darnell JE (2018) Pre-mRNA processing includes N6 methylation of adenosine residues that are retained in mRNA exons and the fallacy of "RNA epigenetics". RNA 24:262–267
Dave A, Craig JE, Skrzypiec K et al (2019) Epha2 genotype influences ultraviolet radiation induced cataract in mice. Exp Eye Res 188:107806
Deng X, Su R, Weng H et al (2018) RNA N6-methyladenosine modification in cancers: current status and perspectives. Cell Res 28:507–517
Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485:201–206
D'Orazio J, Jarrett S, Amaro-Ortiz A et al (2013) UV radiation and the skin. Int J Mol Sci 14:12222–12248
Dowen RH, Pelizzola M, Schmitz RJ et al (2012) Widespread dynamic DNA methylation in response to biotic stress. Proc Natl Acad Sci U S A 109:E2183–E2191
Engel M, Eggert C, Kaplick PM et al (2018) The role of m6A/m-RNA methylation in stress response regulation. Neuron 99(2):389–403.e9
Esteve-Puig R, Bueno-Costa A, Esteller M (2020) Writers, readers and erasers of RNA modifications in cancer. Cancer Lett 474:127–137
Feng Y, Hu Y, Hou Z et al (2020) Chronic corticosterone exposure induces liver inflammation and fibrosis in association with m6A-linked post-transcriptional suppression of heat shock proteins in chicken. Cell Stress Chaperones 25:47–56
Flügel D, Görlach A, Michiels C et al (2007) Glycogen synthase kinase 3 phosphorylates hypoxia-inducible factor 1alpha and mediates its destabilization in a VHL-independent manner. Mol Cell Biol 27:3253–3265
Fustin J-M, Doi M, Yamaguchi Y et al (2013) RNA-methylation-dependent RNA processing controls the speed of the circadian clock. Cell 155:793–806
Fry NJ, Law BA, Ilkayeva OR et al (2017) N6-methyladenosine is required for the hypoxic stabilization of specific mRNAs. RNA 23:1444–1455
Fry NJ, Law BA, Ilkayeva OR et al (2018) N6-methyladenosine contributes to cellular phenotype in a genetically-defined model of breast cancer progression. Oncotarget 9:31231–31243
Gao Y, Pei G, Li D et al (2019) Multivalent m6A motifs promote phase separation of YTHDF proteins. Cell Res 29:767–769
Genenncher B, Durdevic Z, Hanna K et al (2018) Mutations in Cytosine-5 tRNA methyltransferases impact Mobile element expression and genome stability at specific DNA repeats. Cell Rep 22:1861–1874
Gokhale NS, McIntyre ABR, Mattocks MD et al (2020) Altered m6A modification of specific cellular transcripts affects Flaviviridae infection. Mol Cell 77:542–555.e8
Gordon DE, Jang GM, Bouhaddou M et al (2020) A SARS-CoV-2-human protein-protein interaction map reveals drug targets and potential drug-repurposing. bioRxiv Preprint. 2020 Mar 22
Green NH, Galvan DL, Badal SS et al (2019) MTHFD2 links RNA methylation to metabolic reprogramming in renal cell carcinoma. Oncogene 38:6211–6225

Hao H, Hao S, Chen H et al (2019) N6-methyladenosine modification and METTL3 modulate enterovirus 71 replication. Nucleic Acids Res 47:362–374

Heng J, Tian M, Zhang W et al (2019) Maternal heat stress regulates the early fat deposition partly through modification of m6A RNA methylation in neonatal piglets. Cell Stress Chaperones 24:635–645

Hesser CR, Karijolich J, Dominissini D et al (2018) N6-methyladenosine modification and the YTHDF2 reader protein play cell type specific roles in lytic viral gene expression during Kaposi's sarcoma-associated herpesvirus infection. PLoS Pathog 14(4):e1006995

Hou J, Zhang H, Liu J et al (2019) YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. Mol Cancer 18:163

Huang H, Weng H, Sun W et al (2018) Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol 20:285–295

Imam H, Khan M, Gokhale NS et al (2018) N6-methyladenosine modification of hepatitis B virus RNA differentially regulates the viral life cycle. Proc Natl Acad Sci U S A 115:8829–8834

Imam H, Kim G-W, Mir SA et al (2020) Interferon-stimulated gene 20 (ISG20) selectively degrades N6-methyladenosine modified hepatitis B virus transcripts. PLoS Pathog 16:e1008338

Jaakkola P, Mole DR, Tian YM et al (2001) Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. Science 292:468–472

Jabs S, Biton A, Bécavin C et al (2020) Impact of the gut microbiota on the m6A epitranscriptome of mouse cecum and liver. Nat Commun 11:1344

Jia G, Fu Y, Zhao X et al (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887

Joëls M, Karst H, Sarabdjitsingh RA (2018) The stressed brain of humans and rodents. Acta Physiol (Oxf) 223:e13066

Jurczyszak D, Zhang W, Terry SN et al (2020) HIV protease cleaves the antiviral m6A reader protein YTHDF3 in the viral particle. PLoS Pathog 16:e1008305

Kennedy EM, Bogerd HP, Kornepati AVR et al (2016) Posttranscriptional m(6)a editing of HIV-1 mRNAs enhances viral gene expression. Cell Host Microbe 19:675–685

Kim G-W, Imam H, Khan M et al (2020a) N6-Methyladenosine modification of hepatitis B and C viral RNAs attenuates host innate immunity via RIG-I signaling. J Biol Chem 295:13123–13133

Kim D, Lee J-Y, Yang J-S et al (2020b) The architecture of SARS-CoV-2 transcriptome. Cell 181:914–921

Knuckles P, Carl SH, Musheev M et al (2017) RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. Nat Struct Mol Biol 24:561–569

Koranda JL, Dore L, Shi H et al (2018) Mettl14 is essential for Epitranscriptomic regulation of striatal function and learning. Neuron 99(2):283–292.e5

Kramer GF, Baker JC, Ames BN (1988) Near-UV stress in salmonella typhimurium: 4-thiouridine in tRNA, ppGpp, and ApppGpp as components of an adaptive response. J Bacteriol 170:2344–2351

Kretschmer J, Rao H, Hackert P et al (2018) The m6A reader protein YTHDC2 interacts with the small ribosomal subunit and the 5′-3′ exoribonuclease XRN1. RNA 24:1339–1350

Kuai Y, Gong X, Ding L et al (2018) Wilms' tumor 1-associating protein plays an aggressive role in diffuse large B-cell lymphoma and forms a complex with BCL6 via Hsp90. Cell Commun Signal 16:50

Kuhrt D, Wojchowski DM (2015) Emerging EPO and EPO receptor regulators and signal transducers. Blood 125:3536–3541

Li M, Zhao X, Wang W et al (2018) Ythdf2-mediated m6A mRNA clearance modulates neural development in mice. Genome Biol 19(1):69

Liang Y, Zhan G, Chang K-J et al (2020) The roles of m6A RNA modifiers in human cancer. J Chin Med Assoc 83:221–226

Lichinchi G, Zhao BS, Wu Y et al (2016a) Dynamics of human and viral RNA methylation during Zika virus infection. Cell Host Microbe 20(5):666–673

Lichinchi G, Gao S, Saletore Y et al (2016b) Dynamics of the human and viral m(6)a RNA methylomes during HIV-1 infection of T cells. Nat Microbiol 1:16011
Lin J, Zhu Q, Huang J et al (2020a) Hypoxia promotes vascular smooth muscle cell (VSMC) differentiation of adipose-derived stem cell (ADSC) by regulating Mettl3 and paracrine factors. Stem Cells Int 2020:2830565
Lin Z, Niu Y, Wan A et al (2020b) RNA m6 a methylation regulates sorafenib resistance in liver cancer through FOXO3-mediated autophagy. EMBO J 39:e103181
Liu J, Zhang X, Chen K et al (2019a) CCR7 chemokine receptor-inducible lnc-Dpf3 restrains dendritic cell migration by inhibiting HIF-1α-mediated glycolysis. Immunity 50:600–615
Liu Y, You Y, Lu Z et al (2019b) N6-methyladenosine RNA modification-mediated cellular metabolism rewiring inhibits viral replication. Science 365:1171–1176
Liu S-Y, Feng Y, Wu J-J et al (2020) m6 a facilitates YTHDF-independent phase separation. J Cell Mol Med 24:2070–2072
Lu W, Tirumuru N, St Gelais C et al (2018) N6-Methyladenosine-binding proteins suppress HIV-1 infectivity and viral production. J Biol Chem 293:12992–13005
Lu Z, Ma Y, Li Q et al (2019) The role of N6-methyladenosine RNA methylation in the heat stress response of sheep (Ovis aries). Cell Stress Chaperones 24:333–342
Lu M, Zhang Z, Xue M et al (2020) N6-methyladenosine modification enables viral RNA to escape recognition by RNA sensor RIG-I. Nat Microbiol 5:584–598
Luo G-Z, MacQueen A, Zheng G et al (2014) Unique features of the m6A methylome in Arabidopsis thaliana. Nat Commun 5:5630
Ma L, Zhang J, Yang P et al (2015) Acute restraint stress alters sound-evoked neural responses in the rat auditory cortex. Neuroscience 290:608–620
Magimaidas A, Madireddi P, Maifrede S et al (2016) Gadd45b deficiency promotes premature senescence and skin aging. Oncotarget 7:26935–26948
Maimaitiyiming Y, Zhu H-H, Yang C et al (2020) Biotransformation of arsenic trioxide by AS3MT favors eradication of acute promyelocytic leukemia: revealing the hidden facts. Drug Metab Rev 52:425–437
Malbec L, Zhang T, Chen Y-S et al (2019) Dynamic methylome of internal mRNA N7-methylguanosine and its regulatory role in translation. Cell Res 29:927–941
Martínez-Chantar ML, Latasa MU, Varela-Rey M et al (2003) L-methionine availability regulates expression of the methionine adenosyltransferase 2A gene in human hepatocarcinoma cells: role of S-adenosylmethionine. J Biol Chem 278:19885–19890
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149:1635–1646
Meyer KD, Patil DP, Zhou J et al (2015) 5' UTR m(6)a promotes cap-independent translation. Cell 163:999–1010
Miao W, Li L, Zhao Y et al (2019) HSP90 inhibitors stimulate DNAJB4 protein expression through a mechanism involving N6-methyladenosine. Nat Commun 10:3613
Moriwaki S, Takahashi Y (2008) Photoaging and DNA repair. J Dermatol Sci 50:169–176
Nicolas S, Abdellatef S, Haddad MA et al (2019) Hypoxia and EGF stimulation regulate VEGF expression in human glioblastoma Multiforme (GBM) cells by differential regulation of the PI3K/rho-GTPase and MAPK pathways. Cell 8(11):1397
Niu Y, Lin Z, Wan A et al (2019) RNA N6-methyladenosine demethylase FTO promotes breast tumor progression through inhibiting BNIP3. Mol Cancer 18:46
Oliva Trejo JA, Tanida I, Suzuki C et al (2020) Characterization of starvation-induced autophagy in cerebellar Purkinje cells of pHluorin-mKate2-human LC3B transgenic mice. Sci Rep 10:9643
Oyarzún-Arrau A, Alonso-Palomares L, Valiente-Echeverría F et al (2020) Crosstalk between RNA metabolism and cellular stress responses during Zika virus replication. Pathogens 9(3):158
Pan R, Xu YH, Xu L et al (2020) Methylation changes in response to hypoxic stress in wheat regulated by methyltransferases. Russ J Plant Physiol 67:323–333
Panneerdoss S, Eedunuri VK, Yadav P et al (2018) Cross-talk among writers, readers, and erasers of m6A regulates cancer growth and progression. Sci Adv 4:eaar8263

Pardon M-C, Marsden CA (2008) The long-term impact of stress on brain function: from adaptation to mental diseases. Neurosci Biobehav Rev 32:1071–1072
Patil DP, Chen C-K, Pickering BF et al (2016) M(6)a RNA methylation promotes XIST-mediated transcriptional repression. Nature 537:369–373
Pendleton KE, Chen B, Liu K et al (2017) The U6 snRNA m 6 a methyltransferase METTL16 regulates SAM Synthetase intron retention. Cell 169(5):824–835.e14
Ping X-L, Sun B-F, Wang L et al (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res 24:177–189
Raffaghello L, Lee C, Safdie FM et al (2008) Starvation-dependent differential stress resistance protects normal but not cancer cells against high-dose chemotherapy. Proc Natl Acad Sci U S A 105:8215–8220
Ravanan P, Srikumar IF, Talwar P (2017) Autophagy: the spotlight for cellular stress responses. Life Sci 188:53–67
Richter K, Haslbeck M, Buchner J (2010) The heat shock response: life on the verge of death. Mol Cell 40:253–266
Ries RJ, Zaccara S, Klein P et al (2019) m6A enhances the phase separation potential of mRNA. Nature 571:424–428
Robinson M, Shah P, Cui Y-H et al (2019) The role of dynamic m 6 a RNA methylation in photobiology. Photochem Photobiol 95:95–104
Roundtree IA, He C (2016) Nuclear m(6)a reader YTHDC1 regulates mRNA splicing. Trends Genet 32:320–321
Rubio RM, Depledge DP, Bianco C et al (2018) RNA m6 a modification enzymes shape innate responses to DNA by regulating interferon β. Genes Dev 32:1472–1484
Ruggieri A, Dazert E, Metz P et al (2012) Dynamic oscillation of translation and stress granule formation mark the cellular response to virus infection. Cell Host Microbe 12:71–85
Schofield CJ, Ratcliffe PJ (2004) Oxygen sensing by HIF hydroxylases. Nat Rev Mol Cell Biol 5:343–354
Schwartz S, Agarwala SD, Mumbach MR et al (2013) High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell 155:1409–1421
Schwartz S, Mumbach MR, Jovanovic M et al (2014) Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5′ sites. Cell Rep 8:284–296
Shi H, Wang X, Lu Z et al (2017) YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. Cell Res 27:315–328
Shi H, Zhang X, Weng Y-L et al (2018) m6A facilitates hippocampus-dependent learning and memory through YTHDF1. Nature 563:249–253
Shi Y, Felley-Bosco E, Marti TM et al (2012) Starvation-induced activation of ATM/Chk2/p53 signaling sensitizes cancer cells to cisplatin. BMC Cancer 12:571
Shi Y, Fan S, Wu M et al (2019) YTHDF1 links hypoxia adaptation and non-small cell lung cancer progression. Nat Commun 10:4892
Shima H, Matsumoto M, Ishigami Y et al (2017) S-Adenosylmethionine synthesis is regulated by selective N6-adenosine methylation and mRNA degradation involving METTL16 and YTHDC1. Cell Rep 21:3354–3363
Snider JW, Datta NR, Vujaskovic Z (2016) Hyperthermia and radiotherapy in bladder cancer. Int J Hyperth 32:398–406
Soares PI, Ferreira IM, Igreja RA et al (2012) Application of hyperthermia for cancer treatment: recent patents review. Recent Pat Anticancer Drug Discov 7(1):64–73
Song H, Feng X, Zhang H et al (2019) METTL3 and ALKBH5 oppositely regulate m6A modification of TFEB mRNA, which dictates the fate of hypoxia/reoxygenation-treated cardiomyocytes. Autophagy 15:1419–1437
Stöhr N, Lederer M, Reinke C et al (2006) ZBP1 regulates mRNA stability during cellular stress. J Cell Biol 175:527–534

Sun C, Jora M, Solivio B et al (2018) The effects of ultraviolet radiation on nucleoside modifications in RNA. ACS Chem Biol 13:567–572
Suragani RNVS, Zachariah RS, Velazquez JG et al (2012) Heme-regulated eIF2α kinase activated Atf4 signaling pathway in oxidative stress and erythropoiesis. Blood 119:5276–5284
Svobodová Kovaříková A, Stixová L, Kovařík A et al (2020) N6-adenosine methylation in RNA and a reduced m3G/TMG level in non-coding RNAs appear at microirradiation-induced DNA lesions. Cell 9(2):360
Tan B, Liu H, Zhang S et al (2018) Viral and cellular N6-methyladenosine and N6,2'-O-dimethyladenosine epitranscriptomes in the KSHV life cycle. Nat Microbiol 3:108–120
Tanabe A, Tanikawa K, Tsunetomi M et al (2016) RNA helicase YTHDC2 promotes cancer metastasis via the enhancement of the efficiency by which HIF-1α mRNA is translated. Cancer Lett 376:34–42
Thienpont B, Steinbacher J, Zhao H et al (2016) Tumour hypoxia causes DNA hypermethylation by reducing TET activity. Nature 537:63–68
Tirumuru N, Zhao BS, Lu W et al (2016) N(6)-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 gag protein expression. elife 5:e15528
Tirumuru N, Wu L (2019) HIV-1 envelope proteins up-regulate N6-methyladenosine levels of cellular RNA independently of viral replication. J Biol Chem 294:3249–3260
Tsai K, Courtney DG, Cullen BR (2018) Addition of m6A to SV40 late mRNAs enhances viral structural gene expression and replication. PLoS Pathog 14:e1006919
Vabulas RM, Raychaudhuri S, Hayer-Hartl M et al (2010) Protein folding in the cytoplasm and the heat shock response. Cold Spring Harb Perspect Biol 2:a004390
Wächter K, Köhn M, Stöhr N et al (2013) Subcellular localization and RNP formation of IGF2BPs (IGF2 mRNA-binding proteins) is modulated by distinct RNA-binding domains. Biol Chem 394:1077–1090
Wacker M, Holick MF (2013) Sunlight and vitamin D: a global perspective for health. Dermatoendocrinol 5:51–108
Wang X, Lu Z, Gomez A et al (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505:117–120
Wang X, Zhao BS, Roundtree IA et al (2015) N(6)-methyladenosine modulates messenger RNA translation efficiency. Cell 161:1388–1399
Wang X, Feng J, Xue Y et al (2016) Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. Nature 534:575–578
Wang Y, Li Y, Yue M et al (2018) N6-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications. Nat Neurosci 21(2):195–206
Wang J, Ishfaq M, Xu L et al (2019a) METTL3/m6A/miRNA-873-5p attenuated oxidative stress and apoptosis in Colistin-induced kidney injury by modulating Keap1/Nrf2 pathway. Front Pharmacol 10:517
Wang JX, Breaker RR (2008) Riboswitches that sense S-adenosylmethionine and S-adenosylhomocysteine. Biochem Cell Biol 86:157–168
Wang X, Li Y, Chen W et al (2019b) Transcriptome-wide reprogramming of N6-methyladenosine modification by the mouse microbiome. Cell Res 29:167–170
Wang X, Wu R, Liu Y et al (2020a) m6A mRNA methylation controls autophagy and adipogenesis by targeting Atg5 and Atg7. Autophagy 16:1221–1235
Wang J, Wang L, Diao J et al (2020b) Binding to m6A RNA promotes YTHDF2-mediated phase separation. Protein Cell 11:304–307
Wu C, Chen W, He J et al (2020) Interplay of m6A and H3K27 trimethylation restrains inflammation during bacterial infection. In: Sci Adv 6:eaba0647
Xiang Y, Laurent B, Hsu C-H et al (2017) RNA m6A methylation regulates the ultraviolet-induced DNA damage response. Nature 543:573–576
Xiao Y, Peng H, Hong C et al (2017) PDGF promotes the Warburg effect in pulmonary arterial smooth muscle cells via activation of the PI3K/AKT/mTOR/HIF-1α signaling pathway. Cell Physiol Biochem 42:1603–1613

Xue M, Zhao BS, Zhang Z et al (2019) Viral N6-methyladenosine upregulates replication and pathogenesis of human respiratory syncytial virus. Nat Commun 10:4595

Yang Y, Fan X, Mao M et al (2017) Extensive translation of circular RNAs driven by N6-methyladenosine. Cell Res 27:626–641

Yang L, Perrera V, Saplaoura E et al (2019a) m5C methylation guides systemic transport of messenger RNA over graft junctions in plants. Curr Biol 29:2465–2476.e5

Yang S, Wei J, Cui Y-H et al (2019b) m6A mRNA demethylase FTO regulates melanoma tumorigenicity and response to anti-PD-1 blockade. Nat Commun 10:2782

Yao M, Dong Y, Wang Y et al (2020) N6-methyladenosine modifications enhance enterovirus 71 ORF translation through METTL3 cytoplasmic distribution. Biochem Biophys Res Commun 527:297–304

Ye F, Chen ER, Nilsen TW (2017) Kaposi's sarcoma-associated herpesvirus utilizes and manipulates RNA N6-adenosine methylation to promote lytic replication. J Virol 91

Yoon K-J, Vissers C, Ming G-L et al (2018) Epigenetics and epitranscriptomics in temporal patterning of cortical neural progenitor competence. J Cell Biol 217(6):1901–1914

Yu J, Li Y, Wang T et al (2018) Modification of N6-methyladenosine RNA methylation on heat shock protein expression. PLoS One 13:e0198604

Yue Y, Liu J, Cui X et al (2018) VIRMA mediates preferential m6A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. Cell Discov 4:10

Zhang C, Samanta D, Lu H et al (2016a) Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m^{6}A-demethylation of NANOG mRNA. Proc Natl Acad Sci U S A 113:E2047–E2056

Zhang C, Zhi WI, Lu H et al (2016b) Hypoxia-inducible factors regulate pluripotency factor expression by ZNF217- and ALKBH5-mediated modulation of RNA methylation in breast cancer cells. Oncotarget 7:64527–64542

Zhang J (2017) Brothers in arms: emerging roles of RNA epigenetics in DNA damage repair. Cell Biosci 7:24

Zhang Q, Riddle RC, Yang Q et al (2019) The RNA demethylase FTO is required for maintenance of bone mass and functions to protect osteoblasts from genotoxic damage. Proc Natl Acad Sci U S A 116:17980–17989

Zhang C, Chen L, Di P et al (2020a) METTL3 and N6-Methyladenosine promote homologous recombination-mediated repair of DSBs by modulating DNA-RNA hybrid accumulation. Mol Cell 79:425–442

Zhang X, Wang F, Wang Z et al (2020b) ALKBH5 promotes the proliferation of renal cell carcinoma by regulating AURKB expression in an m6A-dependent manner. Ann Transl Med 8:646

Zhao BS, Wang X, Beadell AV et al (2017) m6A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. Nature 542:475–478

Zhao T, Li X, Sun D et al (2019) Oxidative stress: one potential factor for arsenite-induced increase of N6-methyladenosine in human keratinocytes. Environ Toxicol Pharmacol 69:95–103

Zhao T-X, Wang J-K, Shen L-J et al (2020) Increased m6A RNA modification is related to the inhibition of the Nrf2-mediated antioxidant response in di-(2-ethylhexyl) phthalate-induced prepubertal testicular injury. Environ Pollut 259:113911

Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29

Zheng Q, Hou J, Zhou Y et al (2017) The RNA helicase DDX46 inhibits innate immunity by entrapping m6A-demethylated antiviral transcripts in the nucleus. Nat Immunol 18:1094–1103

Zhong L, Liao D, Zhang M et al (2019) YTHDF2 suppresses cell proliferation and growth via destabilizing the EGFR mRNA in hepatocellular carcinoma. Cancer Lett 442:252–261

Zhou J, Wan J, Gao X et al (2015) dynamic m(6)a mRNA methylation directs translational control of heat shock response. Nature 526:591–594

Zhou J, Wan J, Shu XE et al (2018) N6-Methyladenosine guides mRNA alternative translation during integrated stress response. Mol Cell 69:636–647

Zhu Y, Zhao Y, Zou L et al (2019) The E3 ligase VHL promotes follicular helper T cell differentiation via glycolytic-epigenetic control. J Exp Med 216:1664–1681

Conventional and Advanced Techniques for N6-Methyladenosine Modification Mapping in Transcripts

Subhadra Kumari and Srinivasan Muthusamy

Contents

Abstract Until 2012, the genome-wide distribution of N6-methyladenosine (m^6A) was ill-defined, and its importance was of the least concern. Later, transcriptome-wide m^6A mapping in both human and mouse transcriptomes revealed that more than 7000 human genes with 12,000 m^6A sites. m^6A sites are significantly enhanced in the consensus motif RRACH, in which A is converted to m^6A. m^6A modifications are preferentially enriched in the 3′ end of the coding sequence (CDS) and near stop codons. This data emphasises the selective enrichment of m^6A in the transcripts and their role in RNA dynamics. The m^6A modification can be investigated by several methods with different intensity of resolution, transcriptome level global changes, transcript-specific m^6A identification and even nucleotide level identification. The techniques can be placed under two broad categories of conventional and advanced methods. The conventional methods employed in characterising m^6A modifications

S. Kumari · S. Muthusamy (✉)
Department of Life Science, National Institute of Technology, Rourkela, India

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_14

are chromatography-based, mass spectrometry, Sanger sequencing and blotting methods. The advanced techniques rely on antibody-based capture as well as reverse transcription-mediated stalling of cDNA synthesis followed by RNAseq for having a single-nucleotide resolution. Some of such advanced procedure are MeRIP-seq, PAm^{6}A-seq, DART-Seq, m^{6}A-REF-Seq, high-resolution melting analysis and SCARLET. The present book chapter will have an extensive review of these techniques.

Keywords RNA modifications · m^{6}A methylation · Epitranscriptomics · RNA regulation · mRNA methylation

1 Introduction

RNA is an unpretentious molecule with well-established chronicle documentation of reversible post-transcriptional chemical modifications just as DNA and histones (Helm and Motorin 2017). RNA modifications were initially illustrated in 1968, accompanying the exploration of RNA methylation in Hela cells (Christofi and Zaravinos 2019). The newly uncovered reversible RNA modifications have added layer in the field of epigenetics. Initially, the scientific community had the notion that most RNA species were passive, short-lived and rigid after their covalent attachment (Fu et al. 2014). According to MODOMICS database, as of now, RNA encompasses 163 distinct RNA modifications (Boccaletto et al. 2018). Inosine (I), pseudouridine (Ψ), m^{3}C, m^{5}C, m^{7}G, 2'-O-methylated nucleotides (Nm) and dihydrouridine (D) are some of the examples.

The epitranscriptomic modifications are a common phenomenon both in coding RNA (mRNA) and noncoding RNAs (tRNA, rRNA, lncRNA, miRNA, circRNA). Of these, tRNAs are extensively modified with immense chemical diversity, likely to 25% of its nucleotides modified (Helm and Motorin 2017). Second to tRNA, mRNA modifications have been revealed by the implementation of fast-growing high-throughput sequencing technology (HTS), for instance, methylation (e.g. m^{6}A, m^{1}A, m^{5}C, hm^{5}C, 2'OMe), pseudourylation (Ψ) and deamination (e.g. A-to-I RNA editing) (Schaefer et al. 2017). Modifications like 5′ cap and poly(A) tail at the 3′ of mRNA transcripts are well-known modifications. 5′ cap mediates important functions like RNA processing, translation and transcript stability, whereas poly (A) tail is concerned with functions like nuclear export, mRNA stability and translation initiation (Morena et al. 2018).

N6-methyladenosine (m^{6}A) RNA methylation is the most pervasive form of mRNA modification modulating varieties of biological functions like regulation of gene expression, cell self-renewal, differentiation, invasion and apoptosis; it also impacts RNA structure and metabolism. m^{6}A mRNA modification exerts its biological effects through "writers" (m^{6}A methyltransferases), "erasers" (demethylases) and "readers" (recognise m^{6}A-containing mRNA regulating mRNA stability and

translation controlling downstream effects) (Lin et al. 2019) suggesting their importance in the homeostasis of biological processes like metabolism, circadian rhythm, immune response, embryogenesis and cancer development. The emerging evidence suggests that the imbalance of RNA modifications resulting from abnormal expression of RNA modifiers have been reported to be associated with the progression of the malignant tumours reinforcing the importance of m^6A RNA modification. So far, several tools and technologies have been identified to figure out biological function and significance of these modifications, and all of them have come up with important information about types and allocation of RNA modifications over the transcriptome. All the existing wet lab methodologies and bioinformatics tools vary from one another based on their ability to detect m^6A post-transcriptional alteration within the RNA sequence (Limbach and Paulines 2017).

The conventional technique such as dot blot, chromatography and mass spectrometry is used to detect and quantify RNA modifications, and following advanced methodologies MeRIP-seq, PA-m^6A-Seq, DART-Seq, m^6A-REF-Seq, SCARLET and high-resolution melting analysis are used in the genome-wide single-nucleotide resolution of m^6A analysis. The SCARLET and HRM analysis are used in the targeted analysis of transcripts, while the other methods are used in the transcriptome-wide analysis. This chapter introduces the reader to the established wet lab methodologies and bioinformatics tools as well as recent breakthroughs in the detection and quantification of RNA modifications.

2 Conventional Methods

2.1 Reverse Transcription Sequencing

A key approach on which many high-throughput techniques were built is based on the identification of modification in reverse transcription (RT) signatures. Modifications in an RNA template may obstruct RT enzymes, forcing it to apprehend and to misincorporate a dNTP to an unaltered template. The abortive cDNA resulting from RT-arrest are observed to be terminating at or near the modification site, and the incidence of dNTP misincorporation exhibits mutations in the sequence of cDNA (Helm and Motorin 2017). Some RNA modifications like m^6A, m^5C pseudouridine and ribothymidine are RT-silent and don't exhibit any variation in cDNA sequence (Muthusamy 2020).

In the recent past, a new technique has accurately decoded m^6A allocation over unidentified regions, through an antibody-independent approach that yields abortive cDNA signatures at m^6A locus, having high resolution (Hong et al. 2018; Nigita et al. 2018). Reverse transcription of m^6A modified RNA in the presence of dATP, dCTP, dGTP and 4SedTTP (selenium at 4-position of deoxythymidine triphosphate) resulted in truncated cDNA. The A:T base pairing is stable enough to accommodate the selenium atoms between them while m^6A is unable to do so. The Se and methyl group interaction affect the base-pairing stability with reduced stacking stabilisation

and loss of hydrogen bond energy. Therefore, 4SeT cause RT truncation opposite the m^6A site because of the unfavourable stacking interaction between m^6A and 4SeT, resulting in a seemingly aborted cDNA product besides the full-length cDNA and making the "RT-silent" modification visible during the reverse transcription process, while unmodified A in RNA remains silent (Fig. 1). The truncation of synthesis in the presence of m^6A is influenced by incubation temperature, Mg^{2+}, 4SedTTP and reverse transcriptase concentration; a low concentration of these components decreases the specificity. This method can also detect the hemi-methylation status of RNA. In the future, this method could be proven propitious in mapping m^6A along with the transcriptome at a single-nucleotide resolution (Hong et al. 2018).

2.2 Chromatography-Based Methods

Thin-layer chromatography (TLC) is the classical, low-cost and straightforward approach for detecting RNA modifications, accompanied by ultraviolet (UV) spectrophotometric determinations. Differences in net charge, polarity and hydrophobicity between nucleotides form the basis of this technique, consequently leading to their chromatographic separation for qualitative and quantitative analysis. The technique is instrumental in resolving both types as well as the position of a modified nucleotide when applied to RNA of 50–150 nucleotide in length. Modified nucleosides or nucleotides are determined by comparing them against a well-recognised standard (Grosjean et al. 2004).

In contrast to one-dimension (1D) chromatography, two-dimensional thin-layer chromatography (2D-TLC) is proficient enough for determining most of the altered nucleotides existing in RNA (Bodi et al. 2010). The RNA samples are to be digested with ribonuclease T1 and RNase A and labelling the fragmented RNA with [γ-^{32}P] ATP using polynucleotide kinase enzyme. The labelled RNA has to be precipitated and again digested with P1 nuclease to generate individual nucleotides. 2D separation of the nucleotides over the cellulose plates depends on their charge and hydrophobicity in two different solvents, characteristics that are influenced by methylation (Mongan et al. 2019). 2D-TLC has been widely used to identify the position of numerous methylated nucleotides, and it allows the mapping of m^6A modification site to be assessed quantitatively (Bodi et al. 2010).

It has been reported that 3D-TLC have mapped the retention values of about 70 modified nucleotides. Like 2D-TLC, the modified nucleotides in RNA are often radiolabelled with [32P], [14C] or [3H] for detection. The application of [32P]-post labelling methods assists in the identification of even minute quantities of modified nucleotides resulting in increased sensitivity (Basturea 2013). On the other hand, the employment of [14C]- or [3H]-labelling process enables the quantification of nucleotides that might be assessed by different methods as well (Grosjean et al. 2004). Recently, a TLC-based method was developed to assay the stoichiometry of the methylation at a known location. When sequence information is known, site-specific cleavage and radioactive labelling followed by ligation-assisted extraction and

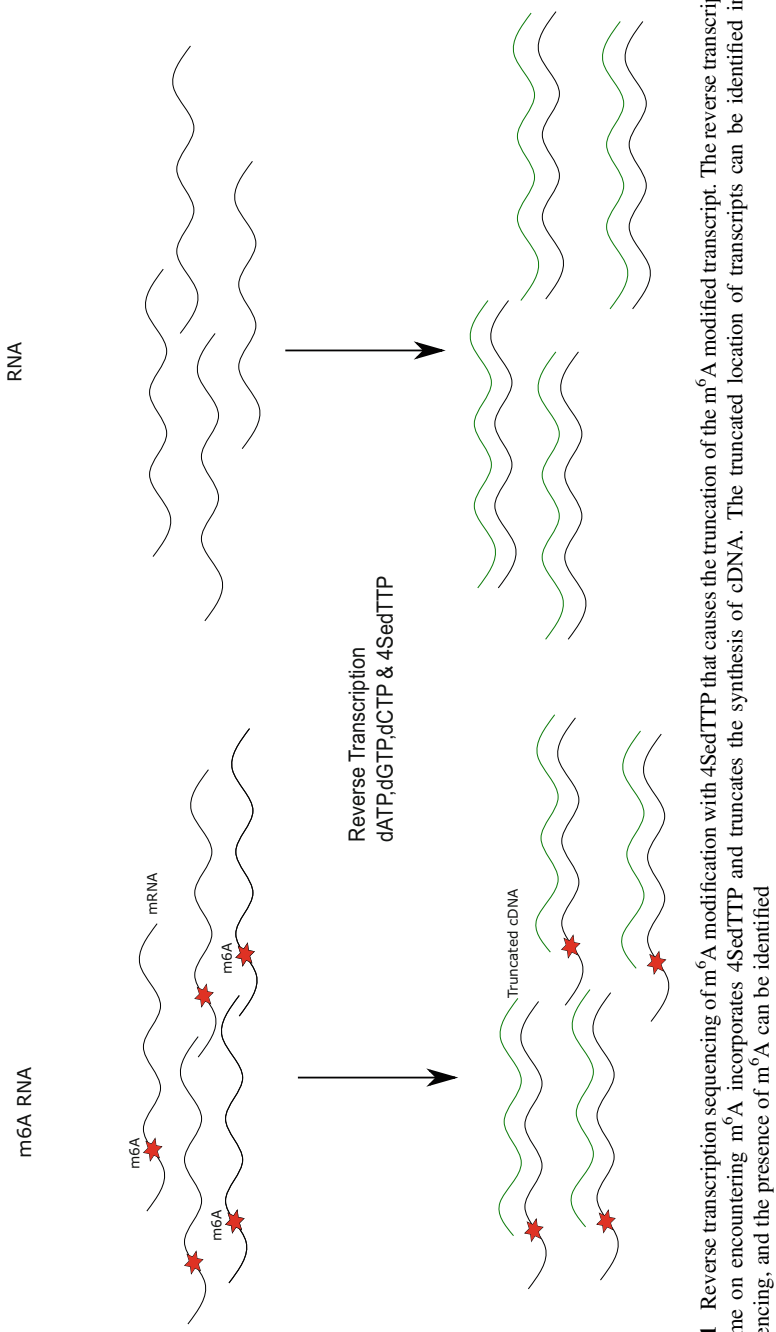


Fig. 1 Reverse transcription sequencing of m^6A modification with 4SedTTP that causes the truncation of the m^6A modified transcript. The reverse transcriptase enzyme on encountering m^6A incorporates 4SedTTP and truncates the synthesis of cDNA. The truncated location of transcripts can be identified in the sequencing, and the presence of m^6A can be identified

thin-layer chromatography (SCARLET) are used to identify and analyse modification levels within individual genes (Schwartz and Motorin 2017).

Though TLC-based analysis of RNA modifications is a low-cost method, without the need for sophisticated instrumentation, however, this analytical procedure is often avoided because of the use of radioactive labelling and susceptibility to degradation. It is a time-consuming procedure and provides shallow information into the transcriptome-wide methylation status and sequence specificity (Mongan et al. 2019).

2.3 Dot Blot

It is the simplest semi-quantitative method for detecting of m^6A in the total RNA or semi-purified RNA samples without the electrophoretic or chromatography separation. The sample is fragmented and enriched using immunoprecipitation method to enhance the low abundant m^6A signals. Fragmented RNA, after immunoprecipitation, is directly applied to the nitrocellulose membrane with circular templates that form a dot blot. The membrane is vacuum dried and exposed to UV to cross-link samples with the membrane. The membrane is then exposed to a blocking agent to prevent the non-specific binding of antibodies to the membrane. A primary anti-m^6A antibody is used to mark the RNA containing m^6A followed by using enzyme-conjugated secondary antibodies which deposits chromogenic products on the site of RNA and antibody interaction. The dot blot analysis provides transcriptome-wide semi-quantitative data (Li et al. 2017). However, the dot blot has poor sensitivity when samples are low in m^6A modifications which can be dealt with enrichment of m^6A RNA by immunoprecipitation before analysis. This method can be applied to a different type of RNA after their enrichment.

2.4 Mass Spectrometry-Based Methods

Mass spectrometry (MS) is a high-throughput and highly sensitive analytical tool for analysing RNA modification, and it is capable of providing information on both mass and structure of biomolecules. In amalgamation with separation based on biophysical characteristics (chromatography), it can determine nucleotides by the mass-to-charge ratio when compared with well-known standards. Mass spectrometry is established on a similar principle to chromatography-based technique except for the requirement for radioisotope labelling.

Over the past 5 years, mass spectrometry-based quantification methods have obtained progressive popularity for identifying and characterising modified RNAs both quantitatively and qualitatively, often, at the level of sequence specificity (Wetzel and Limbach 2016). The method allows the mapping of methylations on different types of RNA fragments. The procedure for analysis involves isolation of

RNA followed by digestion to nucleosides with a suitable endonuclease like RNase T1 before analysis, and the resulting products are separated using reverse phase chromatography using ammonium acetate buffers and C18 stationary phase, analysed using MS and UV light (254 nm) absorption (Limbach and Paulines 2017).

Though progress has been made in the error-free quantification, the technique does not furnish any sequence information (Helm and Motorin 2017). Except for pseudouridine (ψ) modification, all other modified nucleosides are mapped using direct MS detection. Even Cap2 structures are analysed using MS-based methods that are usually not accessible with 2D-TLC (Wetzel and Limbach 2016). Mass spectrometry (MS) offers advantages in biomolecule analysis because of its broad applicability and high accuracy (mass and structural information can be measured with high precision by MS). However, a significant drawback of MS-based methods is that MS equipment is very costly and highly specialised, demanding highly developed expertise for data mining. It needs a too high input of RNA, and it also requires prior information for generating informative data.

3 Advanced Techniques

RNA modification detection using conventional techniques relies on the physiochemical properties of the modified bases that are likely to be underestimated if the modifications are in low-abundance RNAs. The massively parallel sequencing technology has empowered us to detect the modification at a single-nucleotide resolution even in low-abundant RNAs (Schaefer et al. 2017). The exciting part is that massive parallel sequence techniques allow us to deduce the RNA sequence context to the modifications. Before RNA is subjected to parallel sequencing massively, a prerequisite is the enrichment of modified RNA using different enrichment techniques that differ from each other. The transcriptome-wide mapping of the internal m^6A methylated nucleotide first emerged in 2012 (Mongan et al. 2019). Antibodies are economically accessible for most of the methylated RNA residues, for instance, m^6A, m^1A and m^5C. Antibodies can equally be utilised in many conventional methods, for instance, dot blots; this is an economical method to observe changes in m^6A, but it provides less quantitative information about the exact position of m^6A locus (Mongan et al. 2019).

3.1 Low-Throughput Non-sequencing-Based Methods

Site-specific cleavage and radioactive labelling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET) and high-resolution melting analysis (HRM) are the two common methods for studying the m^6A modification with a single-nucleotide resolution. These two techniques can be used to study the targeted

RNA modification in the validation phase of any study as well as in clinical setup for targeted or precision medicine.

3.1.1 SCARLET

SCARLET, site-specific cleavage and radioactive labelling followed by ligation-assisted extraction and thin-layer chromatography, developed in 2013 enabled identification of m^6A modification site in a random mRNA or lncRNA from a total RNA pool to be examined quantitatively (Zhang et al. 2019). The approach is built by collaborating two previously established biochemical approaches, site-specific cleavage and splint ligation, to examine m^6A RNA modification status. In the total RNA pool, site-specific cleavage of candidate sequence hybridised with chimeric DNA with a 2′-OMe/2′-H modification is achieved by adding RNase H, succeeded by labelling of the cleaved 5′ target nucleotide with radioactive ^{32}P; chimeric oligonucleotides are then constructed by splint-assisted ligation to DNA oligonucleotides of 116mer to the target RNA followed by total RNAs digestion with RNases T1/A and purified for the chimeric oligos with terminal modified and unmodified adenine residues of RNA. Finally, nuclease P1 treatment leaves behind the nucleotide mixture that is separated using thin-layer chromatography (TLC), and the spots of radiolabelled modified and unmodified residues are quantified (Fig. 2). Even though SCARLET is based on the direct measurement of m^6A modification at single-nucleotide resolution, it is a low-throughput approach. As a whole, it is a prolonged, time-consuming approach and not suitable for global epitranscriptome study (Wei et al. 2017). Besides, SCARLET can identify the m^6A modification at non-consensus sequences in the RNA as well as the fraction of RNA carrying m^6A and unmodified RNA.

3.1.2 High-Resolution Melting Analysis

Modified nucleotides in nucleic acids can influence the base pairing or stacking interaction that changes the melting property of nucleic acids. HRM-based analysis is routinely used in studying RNA editing, DNA methylation and SNPs (Chateigner-Boutin and Small 2007; Vossen et al. 2009; Wojdacz and Dobrovic 2007). Here m^6A bases affect the stacking interaction of RNAs. The target RNA to be studied for the modification is hybridised with probes having quencher and fluorophore. The probes hybridise side by side, bringing the quencher and fluorophore next to each other and the target modified base in the border area flanked by the probes. The quencher-containing probes are shorter compared to the other probe. Presence of modified bases in the RNA decreases the melting temperature of the probe and the target RNA duplex. The recommended length of the quencher probe is 10 to 12 nucleotide long and more than 20 nucleotide length for fluorophore probe. The position of the target modified base and length of the quencher probe in the hybridisation reaction influence the melting of the RNA. The method can be used

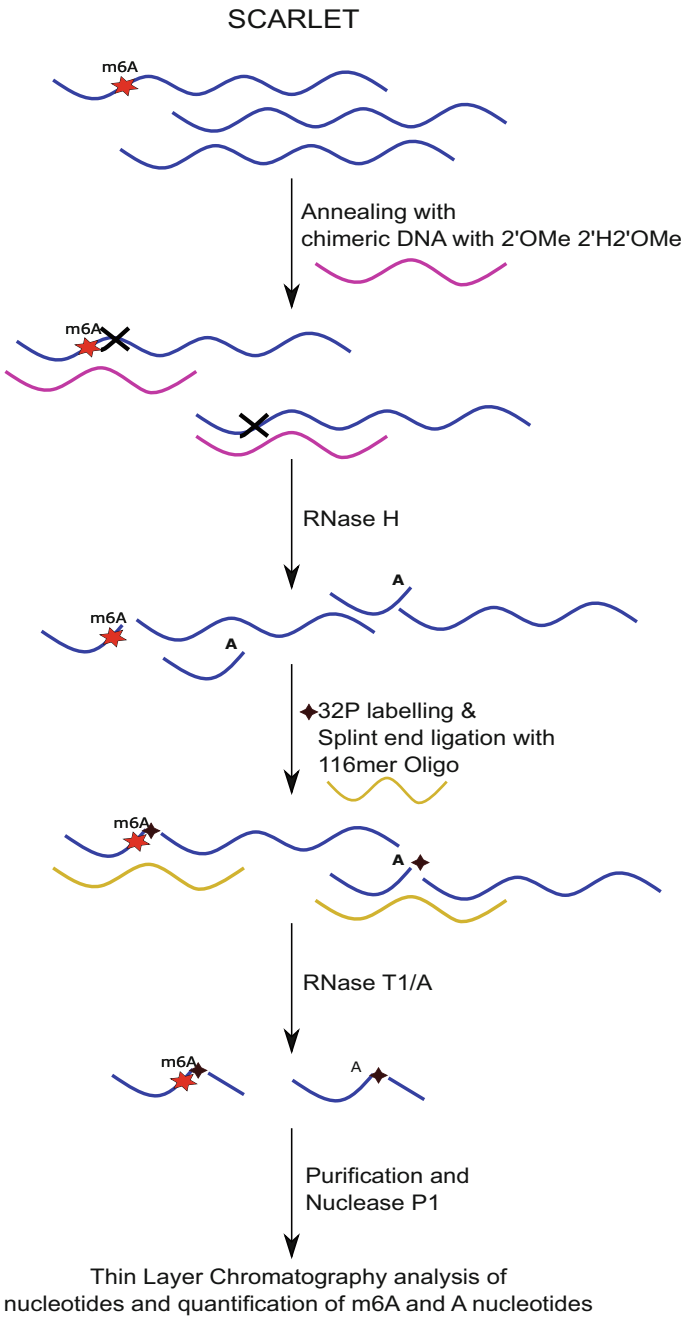


Fig. 2 SCARLET method of mapping and quantifying the m^6A modification in transcripts. A chimeric DNA with 2'OMe 2'H2'OMe modifications is annelated to the site of m^6A followed by RNase H digestion. The 3'ends are labelled ^{32}P and ligated with 116 oligos by splint end ligation. The hybrid RNA is digested with RNase T1/A, and remaining RNA is purified. The purified RNA is digested with nucleases followed by thin-layer chromatography analysis for the quantification of m^6A and A nucleotides

for studying the modified bases in tRNA, rRNA, snRNA and mRNA, but mRNA may require partial purification as it is less abundant in cells. This HRM-based method can be used for studying targeted m^6A modification in transcripts in large scale as well as few samples (Golovina et al. 2014).

3.2 High-Throughput Sequencing-Based Methods

3.2.1 Methylated RNA Immunoprecipitation Sequencing (MeRIP-seq)

It is a straightforward approach of characterising the RNA modifications with a resolution of 150 to 200 bases. To begin with the pulldown, poly-A RNA is to be purified from the total RNA fraction. The purified poly-A RNA is then fragmented into approximately 100–150 bps, and m^6A RNA pulldown is achieved by immunoprecipitation. The immunoprecipitated RNA, along with the fraction of control input RNA, without immunoprecipitation, is used in library preparation for the next-generation sequencing. The reads from NGS are aligned to the reference transcriptome to identify the areas that are enriched m^6A signals against the control input (Dominissini et al. 2012). MeRIP-seq peak detection tools are exomePeak, MeTPeak, MeTDiff, and bespoke scripts and MACS2 (McIntyre et al. 2020). These tools allow us to call peaks respective to the targets and show the m^6A enrichment along with the targets.

The MeRIP-seq allows to map the location of m^6A along the gene but is unable to pinpoint the precise location of m^6A base in sequenced fragments, and suitable for the transcriptome-wide mapping of methylated mRNAs. Still, the authenticity of such antibodies is yet to be answered. Also, they provide low resolution of ~100–200 nucleotides and demand significant amounts of input RNA. Additionally, in case antibody manifests off-target binding, then wrong conclusions can be drawn (Fig. 3a) (Ovcharenko and Rentmeister 2018).

3.2.2 Photo-Crosslinking-Assisted m^6A Sequencing Strategy (PA-m^6A-Seq)

It is inspired by the PAR-CLIP approach used in studying RNA-protein interaction. Here before the RNA harvest, the cells are fed with 4-thiouridine (4-SU) that is readily incorporated into the cellular RNA. 4SU has the 4′ oxygen substituted by sulfur that forms a thioketone structure significantly decreasing the bond dissociation energy leading to the T-to-C transition, and then the base-pair changes in PCR step (Dietz and Koch 1987). The total RNA from the cells is immunoprecipitated with anti-m^6A antibody. To only study mRNA, poly-A RNA can be purified before immunoprecipitation. The immunoprecipitated RNA is cross-linked by exposing them to UV of 365 nm following which the RNA is fragmented for library construction with RNase T1. The crossed-linked RNA is cleared of proteins using

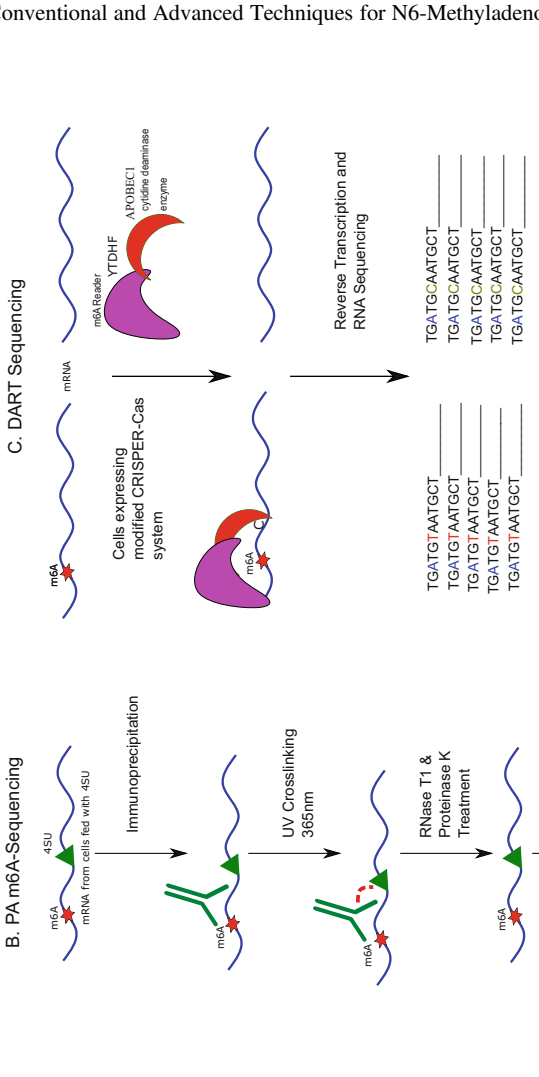


Fig. 3 High-throughput method of mapping m^6A modification in transcripts. (**a**) MeRIP-seq—Cellular total RNA is fragmented and immunoprecipitated with m^6A antibodies. The precipitated fragments are used for RNAseq and mapped to the transcripts to identify the enriched region. (**b**) PA-m^6A-Sequencing—Before RNA harvest, the cells were treated with 4SU. The total RNA was enriched for mRNA and immunoprecipitated with anti-m^6A antibodies. Crosslinking of antibody to the nearby 4SU nucleotide is induced by UV exposure. RNase T1 and proteinase K treated RNA is subjected to RNA sequencing following reverse transcription. The crossing induces the T to C conversion if m^6A modification is present while in the absence of m^6A, no such mutation occurs. (**c**) DART Sequencing—A fusion cas protein having YTDHF reader and APOBEC1 (cytidine deaminase) is constructed and transfected into the cells. The hybrid cas protein recognises the m^6A and deaminates nearby cytosine to uracil. The deaminated uracil is identified in the sequencing reads by mapping to the database

proteinase K and precipitation with Trizol reagent. The sequencing library is constructed for deducing the base modification next to the m^6A site due to the crosslinking of anti-m^6A antibody to the 4-SU uptake nucleotide base. The crosslinking of 4-SU treated uracil with antibody is misread for adenine resulting in T to C transition mutation in the sequencing reads (Fig. 3b). Therefore the uracil base next to the m^6A sites is mutated to T allowing us to likely map the presence of m^6A, by aligning the sequence to the reference transcriptome. This strategy remarkably enabled detection of m^6A motifs to particular sites in RNA fragments and helped in sequencing m^6A and that additionally enhanced m^6A resolution to ~20 nucleotides (Chen et al. 2015a, b, c; Limbach and Paulines 2017).

3.2.3 m^6A-Crosslinked Immunoprecipitation (m^6A-CLIP)

The method differs from PA-m^6A-Seq by not using photoactivatable ribonucleoside modifier before UV crosslinking. The cellular RNA enriched for mRNA is treated with m^6A antibody at 4 °C for more than 2 hours, followed by crosslinking of the incubated mRNA with UV light of 254 nm. Fragmented RNA molecules, to a length of 20 to 80 nucleotides using RNase T1, is antibody crosslinked to the nucleotides close to the m^6A sites. The protein A beads are used to immunoprecipitate cross-linked anti-m^6A antibody using a magnetic stand. The bases that are cross-linked to the antibody undergo mutation during the reverse transcription. Such mutational changes can be observed on performing high-throughput sequencing (NGS). The most common type of mutations is C to T conversion or insertion or deletion of a base. Excess of antibodies must be used to accelerate change in all the transcripts. During data analysis, the adenine next to the mutation sites is considered to be a methylated base (Hsu and He 2019).

m^6A-CLIP has a resolution of approximately 100 nucleotides with an input of as low as 1μg of poly(A)-enriched mRNA (Dietz and Koch 1987). The UV crosslinking provides improved resolution than m^6A-seq, while using significantly less starting material than PA-m6A-Seq and miCLIP. In contrast to PA-m^6A-Seq, m6A-CLIP-seq can be performed on RNA extracted from fresh tissue samples, whereas PA-m^6A-Seq can only be done in cells grown in culture medium containing 4-SU.

3.2.4 Deamination Adjacent to RNA Modification Targets (DART) Sequencing

Here, the RNA modifications are detected by modifying and detecting the nucleotide base next to the m^6A signal. The base next to the m^6A nucleotide is most likely to be "C", as the m^6A signals are enriched in the RRACH motif (Dominissini et al. 2012; Meyer et al. 2012). APOBEC1 is a cytidine deaminase enzyme that converts C nucleotide to U nucleotide in nucleic acids. The targeted editing of C next to m^6A using CRISPER-CAS system is done by constructing a fusion protein of YTDHF, m^6A reader, and APOBEC1, cytidine deaminase. The YTDHF reader recognises the

m^6A residue with high specificity and APOBEC1 proximal to YTDHF deaminates the "C" base to "U". The YTDHF and APOBEC1 upon expression in the cells can cause deamination of C as the cell grows. The conversion of C and the adjacent m^6A can be detected via sequencing techniques like Sangers sequencing or RNAseq (Fig. 3c). DART-seq can identify thousands of m^6A sites even in low input RNA samples as well as in the single-cell m^6A detection. It was demonstrated with input RNA as low as 10 ng, and 79% of the DART-seq edited mRNAs, from a high input DART-seq experiment, were identified. The m^6A deposition in cells over time can be characterised. Besides, the use of long-read NGS methodologies can provide information on m^6A distribution along the length of transcripts (Meyer 2019).

The limitation of antibody-based methods is their inability to distinguish m^6A from N6, 2-O-dimethyladenosine (m^6Am). The use of m^6A reader, YTDHF, in the CRISPER-CAS system can distinguish m^6A from m^6Am. With the help of in vitro deamination assay, the DART sequencing output was found to have a positive correlation between m^6A abundance and editing efficiency. This findings demonstrated that DART-Seq can measure m^6A abundance in individual transcripts.

3.2.5 Next-Generation Sequencing Library Preparation Methods

The library preparation involves the attachment of adaptor to the end of target nucleic acids to be sequenced and amplified. However, in the case of RNA library preparation, an additional reverse transcription step to convert RNA into a double-stranded DNA is essential. The adaptors serve the purpose of enrichment of the ligated target sequence via a polymerase chain reaction, using primers annealing to the adaptors. The adaptor primed enrichment reduces the biasness of enrichment of any specific RNA fragments. Another exciting feature of NGS is the use of index or barcode sequences in the primers that allow us to do multiplexing the samples together in a single sequencing run. The barcode/index sequence marks are distinguished during the analysis pipeline, and data of an individual sample can be extracted from RAW data.

4 Bioinformatics and Data Mining for Epitranscriptomics

The data generated by the methodologies elucidated here are complicated, resulting in the development of robust bioinformatics tools. Bioinformatics approaches provide an integrative analysis of modified nucleotides from high-throughput sequencing data. Several databases in combination with NGS-based modification mapping technologies have been employed for removing experimental noise from RNA modified signatures and thus providing important information about the chemical modifications in RNAs and also the biological consequences of such modifications. Some of the databases for RNA modifications are described below:

4.1 MODOMICS

MODOMICS is the leading global database resource for the biology of RNA modification. It is a database of RNA modifications with complete information regarding chemical structures of modified ribonucleosides, type of reaction (e.g. methylation, thiolation, deamination, etc.), the biosynthetic pathways of RNA modifications, the position of particular modified nucleosides in RNA sequences and functionally characterised RNA modifying enzymes (Boccaletto et al. 2018).

According to MODOMICS database, as of now, RNA encompasses 163 distinct post-transcriptional modifications of RNA that establish a functional diversity, and about 340 functionally characterised proteins have been identified that are involved in RNA modifications. Additionally, the MODOMICS database hosts data of liquid chromatography-mass spectrometry (LC-MS) for modified nucleosides and simplified molecular-input line-entry system (SMILES) for the chemical structures of modified nucleosides that are represented by 3D structures and their occurrence in Protein Data Bank (PDB) (Morena et al. 2018).

4.2 RNAMDB

Since its establishment in 1994, it acts as a reference database of RNA modifications. RNAMDB reports detailed data on the 109 currently known RNA modifications. This database provides a user-friendly, searchable interface that addresses the user to comprehensive information on chemical structures, common name, symbols elemental composition and molecular weight (Cantara et al. 2010). This database also provides a collection of tools of immense advantage in mass spectrometry-based quantification methods and classification of natural or modified RNAs. The current record of the database, at present situated at the RNA Institute at the State University of New York at Albany, comprises of all naturally occurring, chemically modified nucleotide residues for which the chemical structures are known (Rozenski et al. 1999). Commencing from an RNA sequence, it is feasible to calculate the molecular mass, electrospray series, CID fragments, base losses and fragment digestions (Morena et al. 2018).

4.3 RMBase (RNA Modification Base)

The RMBase is a comprehensive database for decoding the post-transcriptional modifications of RNAs analysed from high-throughput sequencing data (MeRIP-seq, m^6A-seq, miCLIP, m^6A-CLIP, Pseudo-seq, Ψ-seq, CeU-seq, Aza-IP, RiboMeth-seq) (Sun et al. 2016). Since its establishment, RMBase provides a variety of interfaces and graphic visualisations to show relationships between RNA modification sites and miRNA. It is also used to illustrate the disease-related single-

nucleotide polymorphisms (SNPs) residing in the modification sites/regions and RNA-binding proteins (RBPs), thus facilitating identification of the massive modification sites in normal tissues and cancer cells (Morena et al. 2018).

At present, the RMBase v2.0 accommodates ~1,373,000 m^6A, ~5400 m1A, and ~5100 20-O-methylations, ~9600 pseudouridine modifications, ~1000 m5C modifications and ~ 2800 other types of RNA modifications. Additionally, a new module called "Motif" has been constructed that provides the visualised logos and position weight matrices (PWMs) of the modification motifs. Another web-based module called modTool is built to annotate, visualise and query the relationships between RNA modifications and RBPs. RMBase identified an abundance of RNA modifications positioned within mRNAs, regulatory ncRNAs (e.g. lncRNAs, miRNAs, pseudogenes, circ-RNAs, snoRNAs, tRNAs), miRNA and disease-related SNPs (Xuan et al. 2018).

4.4 MeT-DB (MethylTranscriptome DataBase)

MeT-DB originated in 2014, and it is the first comprehensive database focussing on m^6A methylation in the mammalian transcriptome. MeT-DB encompasses ~ 300 k m^6A methylation sites recognised in 74 MeRIP-Seq samples from various experimental conditions and estimated by exomePeak and MACS2 algorithms. It has been reported that Met-DB v2.0, a significantly upgraded class of Met-DB, has been redesigned to explore rich information on context-specific m^6A methylation peaks and single-base sites under different conditions (Liu et al. 2015).

The renovated Met-DB v2.0 web interface and genome browser yields more beneficial, robust and informative ways to examine and visualise the data. Especially, MeT-DB v2.0 provides premier series of tools precisely designed for understanding m^6A functions. Besides this, MeT-DB encompasses the binding site data of microRNA, splicing factor and RNA binding proteins for comparison with m^6A modification sites and for investigating the potential m^6A functions (Liu et al. 2018).

5 Bioinformatic Tools for Predicting m^6A RNA Modifications

In recent years, the development of innovative tools helps us to predict post-transcriptional modification sites from sequence data. These in silico approaches for high-resolution (base-pair resolution) mapping of m^6A in the mammalian transcriptome are required to manipulate and follow the fate of RNA modifications to characterise these pathways functionally. Some of the user-friendly online computational or bioinformatics tools for predicting the RNA modification sites are summarised in Table 1.

Table 1 List of bioinformatics tools available for analysing the m^6A modification using different data types

Tools	Source	Description	Reference
HAMR	http://www.lisanwanglab.org/hamr/	HAMR (high-throughput annotation of modified ribonucleotides) is a tool that allows identification of RNA modification at single-nucleotide resolution from RNA-seq data	(Kuksa et al. 2017)
iRNA-3typeA	http://lin-group.cn/server/iRNA-3typeA/	Identify the occurrence sites of m1A, m^6A and A-to-I modifications at RNA's adenosine sites	(Chen et al. 2018)
iRNA-PseColl	http://lin.uestc.edu.cn/server/iRNA-PseColl/	Identifying the occurrence sites of different RNA modifications by incorporating collective effects of nucleotides into PseKNC	(Feng et al. 2017)
iRNA-methyl	http://lin.uestc.edu.cn/server/iRNA-Methyl/	Identification of N6-methyladenosine sites using pseudo nucleotide composition across the genome	(Chen et al. 2015a, b, c)
m^6Apred	http://lin.uestc.edu.cn/server/m^6Apred/	A sequence-based predictor for identifying N6-methyladenosine sites in *Saccharomyces cerevisiae* transcriptome	(Chen et al. 2015a, b, c)
MethyRNA	http://lin.uestc.edu.cn/server/methyrna/	Web server for identification of N6-methyladenosine sites	(Chen et al. 2017a, b)
SRAMP	http://www.cuilab.cn/sramp/	SRAMP (sequence-based RNA adenosine methylation site predictor), a tool to predict mammalian N6-methyladenosine (m^6A) sites based on sequence-derived features	(Zhou et al. 2016)
RAM-ESVM	http://server.malab.cn/RAM-ESVM/	Computational predictors for identification of RNA N6-methyladenosine sites in *S. cerevisiae* transcriptomes	(Chen et al. 2017a, b)
WHISTLE	http://whistle-epitranscriptome.com	High-accuracy tool for mapping the human N6-methyladenosine (m^6A) epitranscriptome predicted using a machine learning approach	(Chen et al. 2019)
DeepM6APred	http://server.malab.cn/DeepM^{6}APred	A sequence-based predictor for identifying N6- methyladenosine sites using deep and handcrafted features	(Chen et al. 2015a, b, c)

6 Summary

Understanding m^6A modification is uncovering an additional layer of regulation in epitranscriptomics, and methods to map modifications are crucial for inquiring its biological function. The techniques that are discussed in the present chapter covers both conventional and advanced methods with varying level of sensitivity and amount of input RNA. The conventional methods like RT-Sequencing, chromatography, dot blot, and mass spectrometry helped to uncover different RNA modifications, but they failed to provide the sequence-specific information on the transcript. The advanced techniques have the potential to uncover more mysteries about the m^6A distribution on the transcripts and global distribution. The data that are emerging from the advanced techniques are so huge, which demand the bioinformatic approach to analyse the data with ease. It is to be noted that bioinformatics approaches are indispensable for advanced techniques, which involve the next-generation sequencing for the data generation. However, the efficiency of these advanced techniques requires extensive validation. As of now, each of the methods provides valuable information, and the required information on m6A must govern the choice.

References

Basturea GN (2013) Research methods for detection and quantitation of RNA modifications. Mater Methods 3:1–10

Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307

Bodi Z, Button JD, Grierson D et al (2010) Yeast targets for mRNA methylation. Nucleic Acids Res 38:5327–5335

Cantara WA, Crain PF, Rozenski J et al (2010) The RNA modification database, RNAMDB: 2011 update. Nucleic Acids Res 39:D195–D201

Chateigner-Boutin AL, Small I (2007) A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. Nucleic Acids Res 35:e114

Chen W, Feng P, Ding H et al (2015a) iRNA-methyl: identifying N(6)-methyladenosine sites using pseudo nucleotide composition. Anal Biochem 490:26–33

Chen W, Feng P, Yang H et al (2018) iRNA-3typeA: identifying three types of modification at RNA's adenosine sites. Mol Ther Nucleic Acids 11:468–474

Chen K, Lu Z, Wang X et al (2015c) High-resolution N(6) -methyladenosine (m(6) a) map using photo-crosslinking-assisted m(6) a sequencing. Angew Chem Int Ed Engl 54:1587–1590

Chen W, Tang H, Lin H (2017a) MethyRNA: a web server for identification of N6-methyladenosine sites. J Biomol Struct Dyn 35:683–687

Chen W, Tran H, Liang Z et al (2015b) Identification and analysis of the N(6)-methyladenosine in the Saccharomyces cerevisiae transcriptome. Sci Rep 5:13859

Chen K, Wei Z, Zhang Q et al (2019) WHISTLE: a high-accuracy map of the human N6-methyladenosine (m6A) epitranscriptome predicted using a machine learning approach. Nucleic Acids Res 47:e41

Chen W, Xing P, Zou Q (2017b) Detecting N(6)-methyladenosine sites from RNA transcriptomes using ensemble support vector machines. Sci Rep 7:40242
Christofi T, Zaravinos A (2019) RNA editing in the forefront of epitranscriptomics and human health. J Transl Med 17:1–15
Dietz TM, Koch TH (1987) Photochemical coupling of 5-bromouracil to tryptophan, tyrosine and histidine, peptide-like derivatives in aqueous fluid solution. Photochem Photobiol 46:971–978
Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485:201–206
Feng P, Ding H, Yang H et al (2017) iRNA-PseColl: identifying the occurrence sites of different RNA modifications by incorporating collective effects of nucleotides into PseKNC. Mol Ther Nucleic Acids 7:155–163
Fu Y, Dominissini D, Rechavi G et al (2014) Gene expression regulation mediated through reversible m 6 a RNA methylation. Nat Rev Genet 15:293–306
Golovina AY, Dzama MM, Petriukov KS et al (2014) Method for site-specific detection of m6A nucleoside presence in RNA based on high-resolution melting (HRM) analysis. Nucleic Acids Res 42:e27
Grosjean H, Keith G, Droogmans L (2004) Detection and quantification of modified nucleotides in RNA using thin-layer chromatography. RNA Interference, Editing, and Modification, Springer: 357–391
Helm M, Motorin Y (2017) Detecting RNA modifications in the epitranscriptome: predict and validate. Nat Rev Genet 18:275–291
Hong T, Yuan Y, Chen Z et al (2018) Precise antibody-independent m6A identification via 4SedTTP-involved and FTO-assisted strategy at single-nucleotide resolution. J Am Chem Soc 140:5886–5889
Hsu PJ, He C (2019) High-resolution mapping of N (6)-Methyladenosine using m(6)a crosslinking immunoprecipitation sequencing (m(6)A-CLIP-Seq). Methods Mol Biol 1870:69–79
Kuksa PP, Leung YY, Vandivier LE et al (2017) In silico identification of RNA modifications from high-throughput sequencing data using HAMR. RNA Methylation, Springer: 211–229
Li Z, Weng H, Su R et al (2017) FTO plays an oncogenic role in acute myeloid leukemia as a N(6)-Methyladenosine RNA demethylase. Cancer Cell 31:127–141
Limbach PA, Paulines MJ (2017) Going global: the new era of mapping modifications in RNA. Wiley Interdisciplinary Reviews: RNA 8:e1367
Lin X, Chai G, Wu Y et al (2019) RNA m 6 a methylation regulates the epithelial mesenchymal transition of cancer cells and translation of snail. Nat Commun 10:1–13
Liu H, Flores MA, Meng J et al (2015) MeT-DB: a database of transcriptome methylation in mammalian cells. Nucleic Acids Res 43:D197–D203
Liu H, Wang H, Wei Z et al (2018) MeT-DB V2. 0: elucidating context-specific functions of N 6-methyl-adenosine methyltranscriptome. Nucleic Acids Res 46:D281–D287
McIntyre ABR, Gokhale NS, Cerchietti L et al (2020) Limits in the detection of m(6)a changes using MeRIP/m(6)A-seq. Sci Rep 10:6590
Meyer KD (2019) DART-seq: an antibody-free method for global m(6)a detection. Nat Methods 16:1275–1280
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149:1635–1646
Mongan NP, Emes RD, Archer N (2019) Detection and analysis of RNA methylation [version 1; peer. peer review: 2 approved] https://f1000research.com/articles/8-559
Morena F, Argentati C, Bazzucchi M et al (2018) Above the epitranscriptome: RNA modifications and stem cell identity. Genes 9:329
Muthusamy S (2020) m6A mRNA methylation: a pleiotropic regulator of cancer. Gene 736:144415
Nigita G, Acunzo M, Cho WCS et al (2018) Epitranscriptomics: the novel RNA frontier. Front Bioeng Biotechnol 6:191
Ovcharenko A, Rentmeister A (2018) Emerging approaches for detection of methylation sites in RNA. Open Biol 8:180121

Rozenski J, Crain PF, McCloskey JA (1999) The RNA modification database: 1999 update. Nucleic Acids Res 27:196–197

Schaefer M, Kapoor U, Jantsch MF (2017) Understanding RNA modifications: the promises and technological bottlenecks of the 'epitranscriptome'. Open Biol 7:170077

Schwartz S, Motorin Y (2017) Next-generation sequencing technologies for detection of modified nucleotides in RNAs. RNA Biol 14:1124–1137

Sun W-J, Li J-H, Liu S et al (2016) RMBase: a resource for decoding the landscape of RNA modifications from high-throughput sequencing data. Nucleic Acids Res 44:D259–D265

Vossen RH, Aten E, Roos A et al (2009) High-resolution melting analysis (HRMA): more than just sequence variant screening. Hum Mutat 30:860–866

Wei W, Ji X, Guo X et al (2017) Regulatory role of N6-methyladenosine (m6A) methylation in RNA processing and human diseases. J Cell Biochem 118:2534–2543

Wetzel C, Limbach PA (2016) Mass spectrometry of modified RNAs: recent developments. Analyst 141:16–23

Wojdacz TK, Dobrovic A (2007) Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. Nucleic Acids Res 35: e41

Xuan J-J, Sun W-J, Lin P-H et al (2018) RMBase v2. 0: deciphering the map of RNA modifications from epitranscriptome sequencing data. Nucleic Acids Res 46:D327–D334

Zhang C, Fu J, Zhou Y (2019) A review in research progress concerning m6A methylation and immunoregulation. Front Immunol 10:922

Zhou Y, Zeng P, Li YH et al (2016) SRAMP: prediction of mammalian N6-methyladenosine (m6A) sites based on sequence-derived features. Nucleic Acids Res 44:e91

*N*6-Methyladenosine and G-Quadruplex in Bacterial Messenger RNA

Yingpeng Xie and Xin Deng

Contents

Abstract The modification and secondary structure of messenger RNA (mRNA) are tightly associated with its functions to control many intracellular events and biological processes. In the past decade, rapid development of high-resolution sequencing techniques enables the detection of chemical modifications and structures onto various types of RNA, especially the low-abundance mRNA. Unlike the rapid progress made in eukaryotes, studies on the modifications of prokaryotic mRNA remain limited. Here, we review bacterial *N*6-methyladenosine (m^{6}A) and RNA G-quadruplex (rG4), which are two recently studied mRNA modifications or structures in bacterial mRNA. This chapter includes their occurrences, topological patterns, and potential regulatory functions in bacterial mRNA. Although the role of m^{6}A in bacterial mRNA is still illusive, experiments have shown that rG4 can regulate gene expression at the post-transcriptional level. We also discuss the different characteristics of m^{6}A and rG4 between eukaryotic cells and bacteria.

Keywords m^{6}A · rG4 · mRNA modification · mRNA structure · Bacteria · Post-transcriptional regulation

Y. Xie · X. Deng (✉)
Department of Biomedical Sciences, City University of Hong Kong, Hong Kong SAR, China
e-mail: xindeng@cityu.edu.hk

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_15

1 Introduction

Although more than 170 types of RNA modifications have been identified so far, various modifications on mRNA have not been explored due to the lack of sufficiently sensitive detection technology until recent years (Frye et al. 2018). In eukaryotic mRNA, a group of chemical modifications and secondary structures [such as *N*1-methyladenosine (m^1A), *N*7-methyladenosine (m^7G), 2′-*O*-methylation (Nm), m^6A and rG4] have been detected with high-resolution methods. In contrast, RNA epigenetics is still a largely uncharted territory in prokaryotes. Recently, by using strategies similar to those in eukaryotes, the m^6A modification and rG4 structure are detected and explored in bacterial mRNA (Shao et al. 2020; Deng et al. 2015), which suggests roles of mRNA modifications in prokaryotes.

In mammalian cells, m^6A is the most abundant internal modification in mRNA and plays important regulatory functions in gene expression. The modification on human m^6A is a reversible and dynamic process, which is mediated by a methyltransferase complex and two demethylases (Fig. 1). The methyltransferase complex is composed of three "writer" proteins including methyltransferase-like 3 (METTL3), methyltransferase-like 14 (METTL14), and Wilms tumor 1 associated protein (WTAP). METTL3 and METTL14 proteins form a heterodimer, which WTAP interacts and guides to specific RNA loci (Roundtree et al. 2017). Two demethylases, fat mass and obesity-associated protein (FTO) and AlkB homolog 5 (ALKBH5), act as "eraser" proteins to catalyze m^6A demethylation (Jia et al. 2011; Zheng et al. 2013). The YTH-domain family proteins are identified as the "readers" for m^6A modification that bind and regulate the metabolism of methylated RNAs (Fig. 1). For instance, YTHDF2 binds and promotes the degradation of mRNA by binding the m^6A sites, while YTHDF1 and YTHDF3 contribute to the translation of m^6A-containing mRNAs (Roundtree et al. 2017; Wang et al. 2014, 2015; Meyer et al. 2015). Antibody-based m^6A-seq identifies that m^6A sites are preferentially enriched around stop codons and in long internal exons in mammals (Dominissini et al. 2012). In mammals, m^6A modification in mRNA regulates embryonic development and growth of tumor cells (Roundtree et al. 2017; Weng et al. 2018; Vu et al. 2017). A study in *Arabidopsis thaliana* demonstrates that m^6A sites distribute in plant mRNA around three landmarks: stop codon, start codon, and 3′-untranslated regions (3′-UTRs) (Luo et al. 2014). Recently, several antibody-independent approaches have been developed to identify single-base m^6A maps in eukaryotic cells. The application of m^6A-sensitive RNA-endoribonuclease-facilitated sequencing (m^6A-REF-seq) to the different tissues of human and rats reveals the conservation of m^6A site in the base level (Zhang et al. 2019). Another atom-specific strategy allows precise identification of two closely gathered m^6A sites with the help of m^6A demethylase FTO (Hong et al. 2018). The newly developed m^6A-crosslinking-exonuclease-sequencing (m^6ACE-seq) quantitatively maps both m^6A and *N*6, 2′-O-dimethyladenosine (m^6A_m) in human (Koh et al. 2019). These precise antibody-independent methods provide efficient and convenient tools for m^6A studies in the future.

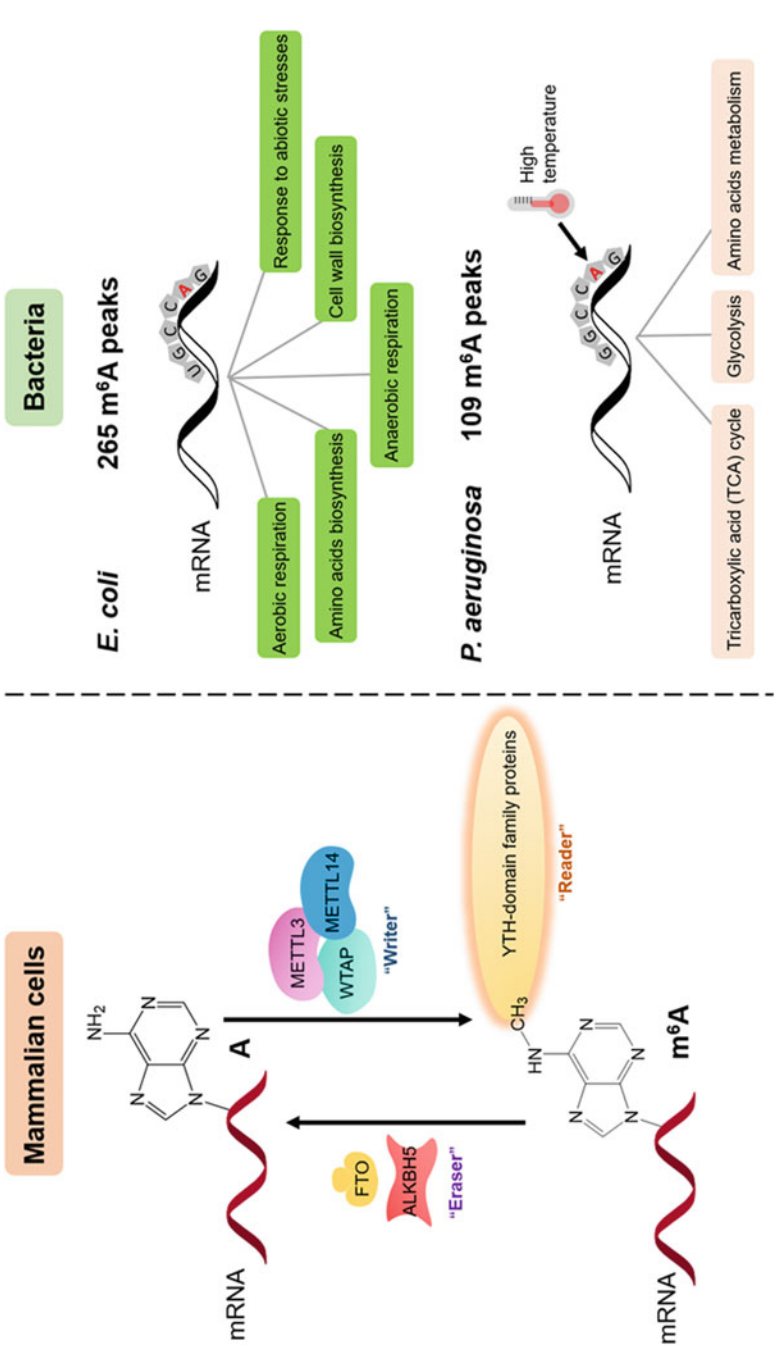


Fig. 1 A model for m^{6}A modification in mammalian and bacterial mRNA. In mammalian cells, m^{6}A is modified by a methyltransferase complex composed of METTL3, METTL14, and WTAP. The m^{6}A is removed by two demethylases, FTO and ALKBH5. YTHDF1, YTHDF2, and YTHDF3 bind and regulate the metabolism of methylated RNAs. In bacteria, 265 m^{6}A peaks are identified in *E. coli*, while 109 peaks are found in *P. aeruginosa*. Culture under high temperature impairs the formation of m^{6}A in *P. aeruginosa*. In *E. coli*, m^{6}A-containing transcripts encode proteins involved in aerobic or anaerobic respiration,

The m^6A sites in bacterial rRNA have been studied for years. In *Escherichia coli*, two methyltransferases RlmF and RlmJ are identified to methylate the A1618 and A2030 sites of 23S rRNA, respectively (Golovina et al. 2012; Sergiev et al. 2008). The RlmF is indispensable for cell growth and fitness, while the *rlmJ* mutant displays mild phenotypes in different culture media (Golovina et al. 2012; Sergiev et al. 2008). The highly conserved KsgA protein works as the dimethyltransferase (catalyze the biosynthesis of *N*6, *N*6-dimethladenosine) of 16S rRNA, which plays a role in antibiotic resistance (O'Farrell et al. 2004). However, the homologs of eukaryotic m^6A methyltransferases (WTAP, METTL3, and METTL14) are absent in the Bacteria Kingdom, suggesting that bacteria have unique m^6A modification patterns.

Similar with the G-quadruplexes (G4) formed in DNA, guanine-rich sequences in RNA can fold into complex structures termed RNA G-quadruplex (rG4). A typical rG4 structure is composed by several layers of stacked G-quartets, which are connected by nucleic acid loops (Kwok et al. 2018; Kwok and Merrick 2017). The rG4 structures could be further stabilized in vitro in the presence of monovalent cations, especially potassium ions (K^+) or sodium ions (Na^+) (Fig. 2) (Fay et al. 2017; Huppert et al. 2008). The rG4 structures are divided into both canonical and noncanonical types. The canonical structures share a G_3L_{1-7} (L_{1-7} represent the loop length 1–7 nt) consensus sequence, while the noncanonical types include long loops (loop length $>$ 7 nt), bulges, and 2-quartet (Fig. 2) (Kwok et al. 2016). Specifically, rG4 sequences that contain $\geq$40% G content but fail to be included into the four previous categories are classified as G $\geq$ 40% (G-rich sequences) (Kwok et al. 2016). In human cells, rG4 structures are involved in transcription, RNA splicing, translation processes, or even many human diseases (Fay et al. 2017; Simone et al. 2015; Millevoi et al. 2012; Cammas and Millevoi 2017). In plants, rG4 structures are important regulators of translation and plant development (Yang et al. 2020). The rG4s are located in both coding sequence (CDS) or 3′/5′-UTRs to impede translation (Demetriades et al. 1992; Arora and Suess 2011; Endoh et al. 2013). However, rG4s in CDS also contribute to ribosomal frameshifting and pre-mRNA splicing, while rG4s in the 3′-UTR regulate miRNA binding, polyadenylation, and mRNA shortening (Weldon et al. 2018; Yu et al. 2014; Rouleau et al. 2017; Beaudoin and Perreault 2013). Thus, rG4 structures display crucial roles in nucleic acid metabolism and post-transcriptional regulation in living eukaryotic cells.

In bacteria, although the presence and critical roles in nucleic acid metabolism of G4 are identified, the study of rG4 in bacterial mRNA is still in its infancy (Saranathan and Vivekanandan 2019; Holder and Hartig 2014; Beaume et al. 2013; Perrone et al. 2017). Although very few rG4s are detected in bacterial transcriptomes previously, our recent study verified the presence of rG4s in multiple strains by using different analyses and biochemical experiments (Guo and Bartel 2016; Shao et al. 2020). Interestingly, bacteria can form hybrid DNA:RNA

Fig. 1 (continued) cell wall organization, amino acids metabolism, and response to stresses. In *P. aeruginosa*, m^6A peaks are located in the coding regions of many housekeeping genes

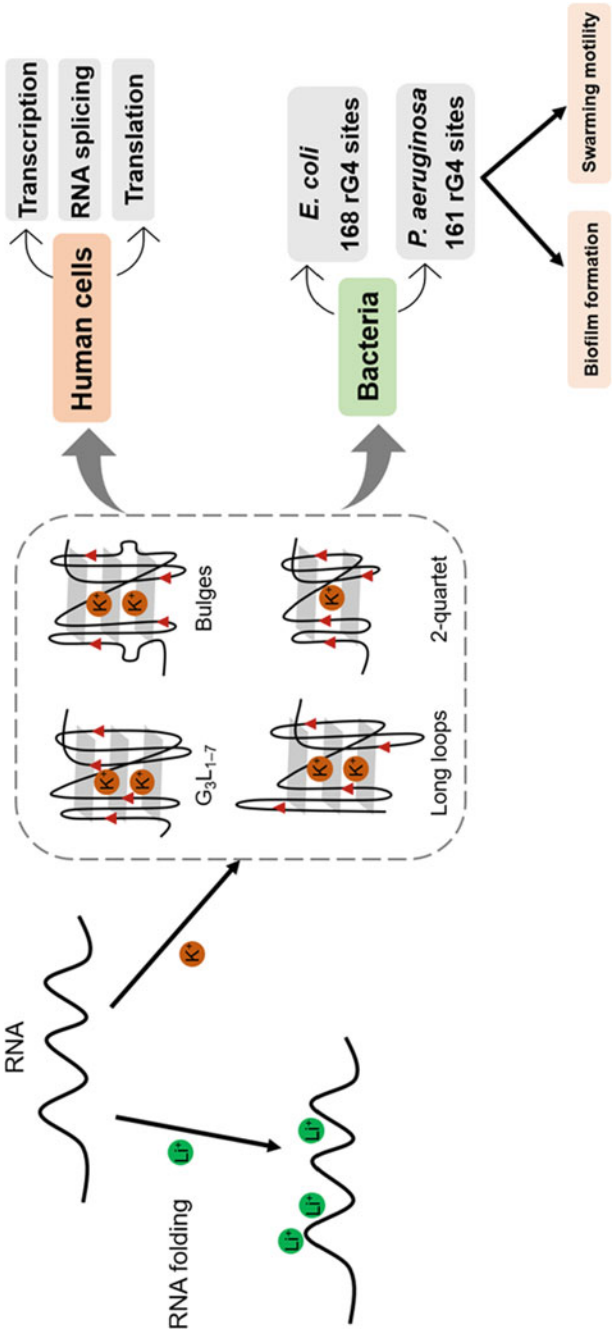


Fig. 2 A model for RNA G-quadruplex in human and bacterial cells. The rG4 structure is preferentially stabilized by K^+ but not Li^+. Four different rG4 structural subclasses are shown. In human cells, rG4 structures are involved in transcription, RNA splicing, and translation processes. In *P. aeruginosa*, rG4s regulate both biofilm formation and swarming motility

G-quadruplexes, which modulate the premature termination of mRNA and transcriptional regulation (Wu et al. 2015).

2 m^6A Modification in Bacteria mRNA

2.1 *Occurrence of m^6A Modification in Bacterial mRNA*

Our recent work uses an ultra-high pressure liquid chromatography coupled with triple-quadrupole tandem mass spectrometry (UHPLC-QQQ-MS/MS) analysis to detect the presence of m^6A, which illustrates that m^6A is prevalent in a wide range of bacterial mRNA with the ratio of m^6A/A varying from 0.02% to 0.28% (Deng et al. 2015). Among the tested bacteria, *Escherichia coli*, *Pseudomonas aeruginosa*, and *Pseudomonas syringae* show >0.2% m^6A/A ratio, while *Staphylococcus aureus* and *Bacillus subtilis* show < 0.08% m^6A/A ratio. The high m^6A/A ratio in Gram-negative bacterial mRNA suggest that m^6A is an important modification in bacteria. It is obvious that the Gram-negative bacteria have a higher m^6A/A ratio than that in the Gram-positive bacteria, which show trace amount of m^6A in mRNA. Notably, two Gram-negative cyanobacteria strains, *Anabaena* sp. PCC 7120 and *Synechocystis* sp. PCC 6803, also show low m^6A/A ratios (<0.04%). These results suggest the presence of m^6A methylase in Gram-positive bacteria, but not in Gram-negative bacteria.

Strikingly, the m^6A/A ratio in *P. aeruginosa* is significantly reduced by increasing culture temperature. When cultured at 45 °C, the m^6A modifications are almost diminished (Fig. 1). However, the m^6A/A ratio of *P. aeruginosa* is constant in different growth media, presence of antibiotics or even oxidative stresses. In contrast, all different stress conditions fail to alter m^6A/A ratio in *E. coli.* For the same bacterial species, different strains have a constant m^6A/A ratio (Deng et al. 2015). For example, three strains of *E. coli* (K-12, 5α, and XL-blue) show m^6A/A ratio about 0.3%. All of six strains of *S. aureus* (Newman, USA100, USA400, USA700, RN4220, and COL) show low m^6A/A ratio about 0.04%. The m^6A/A ratios in both *rlmF* and *rlmJ* mutants are not significantly lower than the wild-type strain, suggesting that these two rRNA methyltransferases do not function on mRNA. In sum, m^6A modification is widely distributed in bacterial mRNA.

2.2 *Topology of m^6A Modification in Bacterial mRNA*

By performing newly developed photo-crosslinking-assisted m^6A-seq (PA-m^6A-seq), the m^6A methylomes of both *E. coli* and *P. aeruginosa* are determined. PA-m^6A-seq capable of improving the resolution of m^6A peaks to around 23 nt (Chen et al. 2015). As the result, 265 m^6A peaks are located in 213 genes of *E. coli.* The enriched m^6A peaks distribute across the entire transcript, especially

inside the coding regions, which is distinct from that found in plants and mammals (Luo et al. 2014; Dominissini et al. 2012; Deng et al. 2015). The metagene profiles of *E. coli* m^6A methylome reveal that 75% of m^6A peaks are enriched inside open reading frame (ORF). Peaks located in the start of gene or the end of gene account for 15% or 13% of all m^6A peaks, respectively. In addition, the identified 41.2% of m^6A peaks in *E. coli* share a conserved sequence (UGCCAG) that is different from eukaryotes (RRACU, R = A/G) (Fig. 1).

In *P. aeruginosa*, the PA-m^6A-seq identifies 109 m^6A peaks representing the mRNA of 68 genes, which is less than that in *E. coli*. The *P. aeruginosa* m^6A methylome shows similar distribution features and m^6A consensus sequence (GGCCAG) as *E. coli* (Fig. 1). The metagene profiles reveal that m^6A peaks in *P. aeruginosa* are also enriched in ORF (77%), which is followed by the start (15%) and the end (8%) of gene regions. The shared topology patterns in both *E. coli* and *P. aeruginosa* suggest that Gram-negative bacteria may have similar m^6A characteristics in mRNA. In addition, the m^6A consensus sequence in mRNA (UGCCAG) is different from the two known methylation sites in rRNA (GUGAAGA and CACAGGU, in which the m^6A sites are highlighted with underline) (Sergiev et al. 2008; Golovina et al. 2012). Overall, the unique patterns of m^6A modification in bacterial mRNA suggest the unique functions of m^6A in bacteria.

2.3 *Potential Functions of mRNA m^6A Modification in Bacteria*

Bacterial m^6A-containing genes encode products involved in various biological pathways. In *E. coli*, gene ontology (GO) identified m^6A-containing transcripts encode proteins involved in aerobic/anaerobic respiration, cell wall organization, amino acids metabolism, and response to stresses, suggesting their roles in regulating these biological pathways (Deng et al. 2015). For example, multiple m^6A peaks are identified in the *hyaABCD* operon, which encodes hydrogenase 1 to regulate hydrogen uptake and transport. The ORFs of *gabD* and *gabT* genes, products of which encode succinate-semialdehyde dehydrogenase and 4-aminobutyrate aminotransferase, also contain several m^6A peaks. Another example is *lacI*, which encodes the negative regulator of the *lacZYA* operon that is required for lactose metabolism. In addition, the m^6A peaks are also identified in 15 small RNAs.

In *P. aeruginosa*, m^6A peaks locate in the coding regions of many housekeeping genes, which encode proteins involved in energy production and central metabolism. For instance, several m^6A peaks are found in the ORF of *ldh*, which encodes leucine dehydrogenase that contributes to amino acid metabolism. m^6A peaks are also found in three adjacent genes PA3415-PA3417 that encode enzymes involved in glycolysis and tricarboxylic acid cycle. Notably, several virulence-related genes (such as *rhlAB* and *rsmYZ*) of *P. aeruginosa* also contain m^6A peaks, indicating that m^6A modification potentially modulates bacterial pathogenesis. These newly identified m^6A

marks in bacteria mRNA may provide new perspectives in regulatory mechanism study.

3 rG4 Structures in Bacteria mRNA

3.1 Distribution of rG4 Structures in Bacterial mRNA

Our recent study focuses on the occurrence and function of rG4 structures among ten diverse model bacterial species (*Acinetobacter* strain ATCC 25922, *Bacillus cereus*, *E. coli*, *Enterococcus faecalis*, *Klebsiella pneumoniae* ATCC 13883, *P. aeruginosa*, *P. syringae*, *S. aureus*, *Salmonella enterica* serovar Typhimurium PY1, and *Vibrio parahaemolyticus* VP001) (Shao et al. 2020). By treating the purified mRNA from these bacterial species with (*E*)-2-(2-(7-(diethylamino)-2-oxo-2*H*-chromen-3-yl) vinyl)-6-fluoro-1-methyl-7-(4-methylpiperazin-1-yl) quinolin-1-ium iodide (also known as QUMA-1), a rG4-specifc fluorescent probe, the rG4 structures are identified in both Gram-positive and Gram-negative strains. However, the occurrence of rG4 in bacteria is species-specific. For example, *Klebsiella pneumoniae* (Gram-negative) and three Gram-positive strains (*S. aureus*, *Bacillus cereus*, and *Enterococcus faecium*) contain more abundant rG4 structures than other tested strains. In addition, after cultured in the LB (Lysogeny broth) medium to mid-log phase, rG4s elicit red fluorescence in live *E. coli* and *P. aeruginosa* cells by labeling with QUMA-1. In sum, these results confirm the presence of rG4 across a wide range of bacterial species.

3.2 Topology of rG4 Structures in Bacterial mRNA

In the presence of K^+, the RNA G-quadruplex sequencing (rG4-seq) can detect reverse transcriptase stalling (RTS) that is specifically incited by rG4 structures, thus determining the sequence context of rG4 (Kwok et al. 2016). To map rG4 locations in bacterial transcriptomes, our recent study performs the rG4-seq for the mRNA samples from *E. coli* and *P. aeruginosa* (Shao et al. 2020). As the result, the rG4-seq detects 168 and 161 rG4 sites in *E. coli* and *P. aeruginosa*, respectively (Fig. 2). Among identified rG4 sites in *E. coli*, 91.67% are 2-quartet type; 7.14% are G-rich sequences; 0.6% are bulges; and 0.6% are long loops. However, the G_3L_{1-7} type is not detected in *E. coli*. The metagene profiles show that almost all *E. coli* rG4 sites are located in the CDS regions, with most of these sites enriched near the start or the end of CDS regions. In *P. aeruginosa*, all five types of rG4 structures are identified in the transcriptome (86.96% as 2-quartet type, 9.94% as G-rich sequences, 1.86% as bulges, 0.62% as long loops, and 0.62% as G_3L_{1-7}), and the uncovered rG4 structures are predominantly located in the CDS regions. However, the rG4 structures in *P. aeruginosa* are enriched at the end of the CDS.

These bacterial topology characteristics are quite different from that of human in at least three aspects (Kwok et al. 2016). First, the most abundant canonical rG4 structure in human transcriptome (G_3L_{1-7}) is in trace amount in bacteria. Second, the majority of rG4 structures in bacteria are 2-quartet type (91.67% in *E. coli* and 86.96% in *P. aeruginosa*), which is only a quarter in human rG4s. Third, rG4s are significantly enriched in CDS regions, but not UTRs of mRNA in human. In conclusion, bacterial mRNA shows distinct topological features of rG4 structures from eukaryotes.

3.3 Potential Functions of mRNA rG4 Structures in Bacteria

In *E. coli*, the rG4-associated genes are involved in multiple pathways, such as envelope biosynthesis, nucleic acid metabolism, transportation, and drug resistance (Shao et al. 2020). For instance, a 2-quartet type rG4 region is found in the *bamA* gene (encodes the outer membrane protein assembly factor BamA). Another G $\geq$ 40% rG4 region is found in the *katG* gene (encodes a catalase-peroxidase). The 2-quartet rG4 peaks are also detected in the coding regions of the *deaD* gene (encodes ATP-dependent RNA helicase DeaD), the *lplA* gene (encodes a lipoate-protein ligase A), the *deoB* gene (encodes a phosphopentomutase), the *pepD* gene (encodes a cytosol non-specific dipeptidase), and the *mreB* gene (encodes a cell shape-determining protein). The rG4 sites in gene coding regions can regulate bacterial gene expression. For example, in the coding region of *hemL* gene (encodes a glutamate-1-semialdehyde 2,1-aminomutase), a 21-bp rG4 region is found to positively regulate the expression of *hemL*.

In *P. aeruginosa*, rG4 sites are associated with metabolic pathways, gene regulation, and several virulence-related processes such as motility, type VI secretion system (T6SS), and quorum sensing (QS). For instance, rG4s are found in the coding regions of *exoS* (encode a type III secretion system effector), *pilU/J* (involved in twitching motility), *tssA1/G1* (involved in T6SS), and *pqsD* (encode an anthraniloyl-CoA anthraniloyltransferase involved in quorum sensing). Notably, an rG4 structure regulates several virulence-related phenotypes in *P. aeruginosa*. Deletion of the rG4 region in *bswR* CDS reduces its own translation, which causes attenuated biofilm formation but enhanced swarming motility of *P. aeruginosa* (Fig. 2). Taken together, rG4s potentially play crucial regulatory roles in bacteria.

4 Conclusion and Perspectives

Although several types of RNA modifications (such as Nm, m^7G, m^1A, 5-methylcytosine, and pseudouridine) have been found in bacteria, their distribution and biological significance in mRNA are rarely studied. Here, we review both m^6A modification and rG4 structure in bacterial mRNA. m^6A is prevalent in mRNA of

many bacterial strains, most of which are Gram-negative. Although the consensus sequence of m^6A is different from that of eukaryotic cells, it is proposed to be conserved in different bacteria (such as *P. aeruginosa* and *E. coli*). In addition, the level of m^6A in *P. aeruginosa* reduces with the increase of temperature, which indicates that this modification is also regulated by external environments. The distribution of rG4 varies from strain to strain, and the predominant rG4 structure is quite different from that of eukaryotic cells. rG4 can regulate gene expression at the post-transcriptional level, thereby affecting certain phenotypes of bacteria. For example, in *P. aeruginosa*, rG4 structure contributes to the biofilm formation by promoting the expression of *bswR*.

The exciting progress in chemical biology approaches allows us to profile many RNA modifications in Bacteria Kingdom. We expect an increasing number of RNA epigenetic maps will be reported for prokaryotic mRNA in the near future. More importantly, the following key questions need to be answered in bacteria: (1) Are RNA modification/structure processes reversible? (2) Which enzymes are the "writers," "erasers," and "readers" for different RNA modifications/structures? (3) What are the biological functions of the RNA modifications/structures? Solving these key issues will significantly improve our understanding of bacterial RNA epigenetics.

Acknowledgments This work was supported by the General Research Fund of Hong Kong (11102720, 21103018, and 11101619 to X. D.), National Natural Science Foundation of China (31670127 and 31870116 to X. D.), Innovation Technology Fund of Hong Kong (ITS/195/18 to X. D.), and City University of Hong Kong Applied Research Grant (9667172 to X. D.).

References

Arora A, Suess B (2011) An RNA G-quadruplex in the 3' UTR of the proto-oncogene PIM1 represses translation. RNA Biol 8:802–805

Beaudoin JD, Perreault JP (2013) Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening. Nucleic Acids Res 41:5898–5911

Beaume N, Pathak R, Yadav VK et al (2013) Genome-wide study predicts promoter-G4 DNA motifs regulate selective functions in bacteria: radioresistance of *D. radiodurans* involves G4 DNA-mediated regulation. Nucleic Acids Res 41:76–89

Cammas A, Millevoi S (2017) RNA G-quadruplexes: emerging mechanisms in disease. Nucleic Acids Res 45:1584–1595

Chen K, Lu Z, Wang X et al (2015) High-resolution *N*6-methyladenosine (m6A) map using photo-crosslinking-assisted m^6A sequencing. Angew Chem 127:1607–1610

Demetriades D, Hager J, Nikolaides N et al (1992) Proteus syndrome: musculoskeletal manifestations and management: a report of two cases. J Pediatr Orthop 12:106–113

Deng X, Chen K, Luo GZ et al (2015) Widespread occurrence of *N*6-methyladenosine in bacterial mRNA. Nucleic Acids Res 43:6557–6567

Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. Nature 485:201–206

Endoh T, Kawasaki Y, Sugimoto N (2013) Suppression of gene expression by G-quadruplexes in open reading frames depends on G-quadruplex stability. Angew Chem Int Ed Engl 52:5522–5526

Fay MM, Lyons SM, Ivanov P (2017) RNA G-Quadruplexes in biology: principles and molecular mechanisms. J Mol Biol 429:2127–2147

Frye M, Harada BT, Behm M et al (2018) RNA modifications modulate gene expression during development. Science 361:1346–1349

Golovina AY, Dzama MM, Osterman IA et al (2012) The last rRNA methyltransferase of *E. coli* revealed: the *yhiR* gene encodes adenine-*N*6 methyltransferase specific for modification of A2030 of 23S ribosomal RNA. RNA 18:1725–1734

Guo JU, Bartel DP (2016) RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. Science 353(6306):aaf5371–aaf5371

Holder IT, Hartig JS (2014) A matter of location: influence of G-quadruplexes on Escherichia coli gene expression. Chem Biol 21:1511–1521

Hong T, Yuan Y, Chen Z et al (2018) Precise antibody-independent m^6A identification via 4SedTTP-involved and FTO-assisted strategy at single-nucleotide resolution. J Am Chem Soc 140:5886–5889

Huppert JL, Bugaut A, Kumari S et al (2008) G-quadruplexes: the beginning and end of UTRs. Nucleic Acids Res 36:6260–6268

Jia G, Fu Y, Zhao X et al (2011) *N*6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887

Koh CW, Goh YT, Goh WS (2019) Atlas of quantitative single-base-resolution *N*6-methyl-adenine methylomes. Nat Commun 10:1–15

Kwok CK, Marsico G, Balasubramanian S (2018) Detecting RNA G-Quadruplexes (rG4s) in the transcriptome. Cold Spring Harb Perspect Biol 10(7):a032284

Kwok CK, Marsico G, Sahakyan AB et al (2016) rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. Nat Methods 13:841–844

Kwok CK, Merrick CJ (2017) G-Quadruplexes: prediction, characterization, and biological application. Trends Biotechnol 35:997–1013

Luo GZ, MacQueen A, Zheng G et al (2014) Unique features of the m^6A methylome in Arabidopsis thaliana. Nat Commun 5:5630

Meyer KD, Patil DP, Zhou J et al (2015) 5′ UTR m^6A promotes cap-independent translation. Cell 163:999–1010

Millevoi S, Moine H, Vagner S (2012) G-quadruplexes in RNA biology. Wiley Interdiscip Rev RNA 3:495–507

O'Farrell HC, Scarsdale JN, Rife JP (2004) Crystal structure of KsgA, a universally conserved rRNA adenine dimethyltransferase in *Escherichia coli*. J Mol Biol 339:337–353

Perrone R, Lavezzo E, Riello E et al (2017) Mapping and characterization of G-quadruplexes in *Mycobacterium tuberculosis* gene promoter regions. Sci Rep 7:5743

Rouleau S, Glouzon JS, Brumwell A et al (2017) 3' UTR G-quadruplexes regulate miRNA binding. RNA 23:1172–1179

Roundtree IA, Evans ME, Pan T et al (2017) Dynamic RNA modifications in gene expression regulation. Cell 169:1187–1200

Saranathan N, Vivekanandan P (2019) G-Quadruplexes: more than just a kink in microbial genomes. Trends Microbiol 27:148–163

Sergiev PV, Serebryakova MV, Bogdanov AA et al (2008) The *ybiN* gene of *Escherichia coli* encodes adenine-*N*6 methyltransferase specific for modification of A1618 of 23 S ribosomal RNA, a methylated residue located close to the ribosomal exit tunnel. J Mol Biol 375:291–300

Shao X, Zhang W, Umar MI et al (2020) RNA G-Quadruplex Structures Mediate Gene Regulation in Bacteria. MBio 11(1):e02926–e02919

Simone R, Fratta P, Neidle S et al (2015) G-quadruplexes: emerging roles in neurodegenerative diseases and the non-coding transcriptome. FEBS Lett 589:1653–1668

Vu LP, Pickering BF, Cheng Y et al (2017) The *N*6-methyladenosine (m^6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. Nat Med 23:1369–1376

Wang X, Lu Z, Gomez A et al (2014) *N*6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505:117–120

Wang X, Zhao BS, Roundtree IA et al (2015) *N*6-methyladenosine modulates messenger RNA translation efficiency. Cell 161:1388–1399

Weldon C, Dacanay JG, Gokhale V et al (2018) Specific G-quadruplex ligands modulate the alternative splicing of Bcl-X. Nucleic Acids Res 46:886–896

Weng H, Huang H, Wu H et al (2018) METTL14 inhibits hematopoietic stem/progenitor differentiation and promotes Leukemogenesis via mRNA m^6A modification. Cell Stem Cell 22:191–205 e199

Wu RY, Zheng KW, Zhang JY et al (2015) Formation of DNA:RNA hybrid G-quadruplex in bacterial cells and its dominance over the intramolecular DNA G-quadruplex in mediating transcription termination. Angew Chem Int Ed Engl 54:2447–2451

Yang X, Cheema J, Zhang Y et al (2020) RNA G-quadruplex structures exist and function *in vivo* in plants. Genome Biol 21:226

Yu CH, Teulade-Fichou MP, Olsthoorn RC (2014) Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. Nucleic Acids Res 42:1887–1892

Zhang Z, Chen L-Q, Zhao Y-L et al (2019) Single-base mapping of m^6A by an antibody-independent method. In: Science Adv 5:eaax0250

Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29

Regulation of RNA Methylation by TET Enzymes

Kelly M. Banks and Todd Evans

Contents

Abstract The ten-eleven-translocation (TET1/2/3) proteins are capable of oxidizing 5-methylcytosine (m^5C) to 5-hydroxymethylcytosine (hm^5C) and further oxidized derivatives in the process of active demethylation to cytosine. This process has been investigated almost exclusively in the context of DNA demethylation and epigenetics. However, a handful of recent studies have demonstrated that m^5C and TET protein-mediated oxidation are also important in the context of the epitranscriptome and mRNA, leading to modified RNA binding protein profiles and ultimately altering translation efficiency and mRNA stability. Here we summarize the current state of knowledge on the presence and function of m^5C in eukaryotic mRNA and its oxidation.

Keywords Epitranscriptomics · TET · m^5C · Demethylation · mRNA

K. M. Banks · T. Evans (✉)
Department of Surgery, Weill Cornell Medicine, Cornell University, New York, NY, USA
e-mail: tre2003@med.cornell.edu

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_16

1 Introduction

Over 160 modified nucleosides have been described in RNA since their discovery over half a century ago (Boccaletto et al. 2018). The earliest studies were focused on methylated residues in the more abundant RNA species, primarily rRNA and tRNA where base modifications are consistent and well characterized (Motorin and Helm 2010). In the 1970s, mRNA enrichment allowed for the demonstration of modified nucleotides in these species as well (Dubin and Taylor 1975; Adams and Cory 1975). These studies utilized DEAE cellulose chromatography to show various fractions of modified nucleosides, in mRNA the most common being the 5′ 7-methylguanosine (m^7G) cap as well as methyl-6-adenosine (m^6A) (Adams and Cory 1975). As this field has progressed, it has become clear that much as the ability of DNA base modifications can impact gene expression, mRNA base modifications also regulate gene expression, except in this case through control of mRNA metabolism.

5-methylcytosine (m^5C) has been studied extensively in the context of DNA to the point where it has been given the moniker "the fifth base,"(Lister and Ecker 2009). Interestingly, m^5C is also present in tRNA and rRNA species, and was first described in mRNA in the mid-1970s using BHK-21 hamster cells (Dubin and Taylor 1975; Adams and Cory 1975).

2 Profiling 5-Methylcytosine in mRNA

Using chromatography and radiolabeling, it was estimated that 0.18% of residues in mRNA were methylated, 50% as m^7G, 40% as m^6A, and 10% as m^5C (Dubin and Taylor 1975). However, the function of m^5C in mRNA remained elusive as a higher-resolution view of these modifications on specific mRNAs was not feasible, and this line of research was largely abandoned until it was discovered that methyl-CpG-binding protein 2 (MECP2) binds to RNA to regulate splicing (Jeffery and Nakielny 2004; Young et al. 2005). Furthermore, modified m^5C-containing RNA encoding the Yamanaka factors was found to be more stable and could be used to reprogram pluripotent stem cells more efficiently (Warren et al. 2010).

Finally, nearly 40 years after its initial discovery, the advent of chemically modified high throughput sequencing and immunoprecipitation techniques provided nucleotide level resolution of RNA modifications (Squires et al. 2012; Amort and Lusser 2017; Hussain et al. 2013; Delatte et al. 2016). These findings are summarized in Table 1. The first transcriptome-wide screen for m^5C was completed using RNA from HeLa cells and repurposed for RNA the bisulfite sequencing protocol used in the study of DNA methylation, adding a reverse transcription step after bisulfite treatment (Squires et al. 2012). In this study previously reported methylated residues were identified in tRNA and rRNA, as well as a significant enrichment of m^5C in noncoding RNA including pseudogenes compared to coding mRNA. Within

Table 1 Summary of the prevalence and distribution of m^5C and hm^5C in eukaryotic mRNA from transcriptome profiling studies

Base	Method	Model	Prevalence	Distribution	Reference
m^5C	Cellulose chromatography	Hamster cells	0.02% of nucleosides		Dubin and Taylor (1975)
m^5C	Bisulfite sequencing	HeLa cells	9177 sites	–5′ and 3′ UTRs –Argonaute binding sites	Squires et al. (2012)
m^5C	Bisulfite sequencing	Mouse ESCs and neurons	7541 sites (ESCs) 2075 sites (neurons)	–Translation start site, 3′ UTR –Upf1 binding	Amort and Lusser (2017)
m^5C hm^5C	LC-MS	Human tissue	0.5–2.2% of C 0.02% of m^5C		Fu et al. (2014)
hm^5C	hMeRIP-seq	Drosophila S2 cells	1597 sites	–Coding sequence –UC-rich regions	Delatte et al. (2016)
m^5C hm^5C	Bisulfite sequencing	HeLa cells Mouse tissues	5399 sites 2540–4371 sites	–Coding sequence –100 bp downstream of translation start site -CG-rich regions –3′ UTR –Argonaute binding sites	Yang et al. (2017)
m^5C	Bisulfite sequencing	Zebrafish embryos	2902–7521 sites	–Coding sequence	Yang et al. (2019)
hm^5C	hMeRIP-seq	hESCs	1633 sites	–Introns –UC-rich regions –Chromatin-associated mRNA –Mid-level expressed mRNA	Lan et al. (2020)

mRNA an enrichment was found in the untranslated regions at both the 5′ and 3′ end, with enrichment for RNA binding protein sites close to these methylated residues, especially for Argonaute RNA silencing proteins. Through siRNA knockdown in the HeLa cells, it was shown that the methyltransferase NSUN2 is responsible for methylation in mRNA, rRNA, and tRNA, whereas DNMT2 was only active in tRNA methylation. Other studies have confirmed that NSUN2 is the sole cytosine methyltransferase acting on mRNA, with other NSUN family members and DNMT2 involved in mRNA, rRNA, and tRNA methylation (Bohnsack et al. 2019; Yang et al. 2017).

Another study used RNA bisulfite sequencing to profile m^5C in the transcriptome, in this case using murine embryonic stem cells (ESCs) and neurons (Amort and Lusser 2017). The comparison between cell types in this study demonstrated that the vast majority of m^5C sites were lineage specific and that m^5C was far more prevalent

in ESCs (7541 sites compared to 2075, present in 11% compared to 3% of expressed genes). Interestingly, there were roughly twice as many m^5C sites when the nuclear fraction was isolated. In this study, sites were enriched at the translation start site in both cell types and in the 3′ UTR in neurons. In contrast to the previous work done using RNA from HeLa cells, correlation of m^5C sites with Argonaute binding proteins was not found; rather there was a strong correlation with the nonsense-mediated RNA decay protein Upf1 and various splicing factors. Finally, genes with m^5C sites were enriched for highly expressed transcripts and functions specific to each cell type, for instance, early development in ESCs and neuronal development and synaptic plasticity in neurons.

Profiles from mouse tissues with bisulfite sequencing confirmed tissue-specific patterning of m^5C in mRNA, again with uniquely methylated transcripts enriched in functionally relevant GO terms for each tissue type (Yang et al. 2017). In this study, m^5C was found mostly in the coding region, with a clear peak 100 base pairs downstream of the translation start site and was enriched in CG-rich regions. Interestingly, the mean methylation level was only 20.5% at methylated sites, and it did not occur exclusively in the CpG context (55% CG, 28% CHG, 17% CHH, H = A, C, U). In total, 5399 m^5C sites were found in RNA from HeLa cells and 2540–4371 sites in the various murine tissues. In a study of early zebrafish development at 1–6 hours post fertilization, bisulfite sequencing revealed 2902–7521 m^5C sites within 1300–3741 mRNAs, again with enrichment in the coding region but in this case evenly distributed with no predilection for CG-rich regions (Yang et al. 2019).

3 Tet Proteins Oxidize 5-Methyl-Cytosine to 5-Hydroxymethylcytosine and Further Derivatives in mRNA

A major finding in this line of research was that TET1 and TET2 bind mRNA in human embryonic stem cells (hESCs) (He et al. 2016). Schematic representation of the TET proteins is shown in Fig. 1. This original screen failed to identify TET3, but this is likely due to low expression of the TET3 isoform in hESCs (He et al. 2016). Four years later, 5-hydroxymethylcytosine (hm^5C), the oxidized derivative of m^5C generated through action of the Tet enzymes, was demonstrated to also be present in mRNA, thus raising the possibility that m^5C is dynamically regulated in mRNA as occurs in DNA (Fu et al. 2014). Using liquid chromatography mass spectrometry (LCMS) of various human tissues, hm^5C was detected at 3.9 hm^5C modifications per million nucleotides. This study also quantified m^5C using LCMS in these same samples, and it was calculated that roughly 0.5%–2.2% of cytosine residues in RNA were methylated, and therefore roughly 0.02% of all methylated cytosines in RNA were hm^5C in adult human tissues (Fu et al. 2014). It was hypothesized that hm^5C levels are so low because it is possibly an unstable or transient intermediate in m^5C

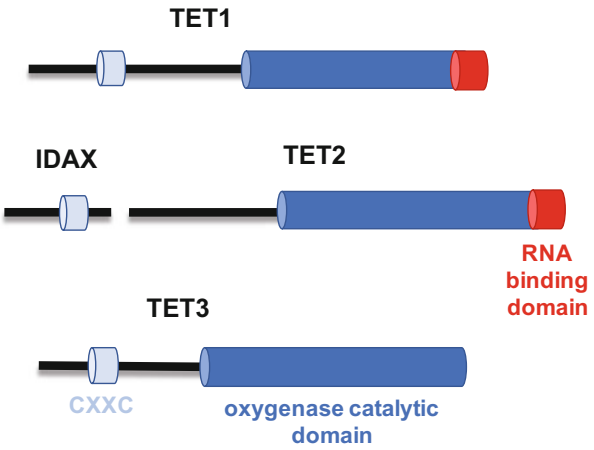


Fig. 1 Schematic representation of TET1, TET2, and TET3 enzyme structure. The CXXC DNA-interacting motif is shown in light blue, the oxygenase-catalytic domain is in dark blue, and the RNA-binding domain is in red

decay. Others have hypothesized that hm^5C is further broken down through the action of the AlkB family of enzymes, as these have been shown to be active in both tRNA and DNA (Aas et al. 2003; Zhang et al. 2018). Through combining recombinant Tet1 and methylated RNA ex vivo followed by high-performance liquid chromatography (HPLC) and LCMS, TET proteins were shown to generate hm^5C on RNA. But interestingly, when overexpressing each of the TET isoforms in HEK293T cells, only TET3 was able to significantly increase hm^5C in RNA although all three isoforms could increase hm^5C in DNA (Fu et al. 2014). Of note, TET1 and TET2 were only located in the nucleus, whereas TET3 was found in both the nucleus and cytoplasm (Fu et al. 2014).

Immunoprecipitation was carried out with an anti-hm^5C antibody followed by sequencing (hMeRIP-seq) using RNA from *Drosophila* S2 cells to more precisely map hm^5C (Delatte et al. 2016). The hm^5C marks were enriched in coding sequences, UC-rich regions, and mRNA actively being translated. Furthermore, m^5C decreased in vitro translation and hm^5C restored translation efficiency to control levels. The *Drosophila* TET orthologue, *dtet*, was knocked out in S2 cells as well as in vivo, and hm^5C was substantially decreased, again implicating TET proteins in hm^5C generation (Delatte et al. 2016). Based on this dataset, a machine learning algorithm was created to predict hm^5C sites in RNA using only sequence data (Liu et al. 2020).

A study of HEK293T cells transfected with vectors to express TET enzymes also demonstrated the presence of m^5C and hm^5C in RNA of human cells (Xu et al. 2016). In this study, hm^5C was found to be enriched in mRNA over total RNA and was present at a level of 7 bases per million bases within mRNA as measured by LCMS. This work further supported the role of TET enzymes for generating hm^5C in mRNA, as the mark was absent either when catalytically dead TET enzymes were expressed or if TET enzymes were co-expressed with dominant negative IDH1 or IDH2 that inhibit the catalytic activity of TET (Xu et al. 2016).

A recent study using mouse embryonic stem cells focused more on distribution and function of hm^5C in pluripotency and differentiation (Lan et al. 2020). Using hMeRIP-seq, 1633 peaks were mapped in 795 transcripts, 110 of which are related to maintaining pluripotency. The levels of hm^5C were found to decrease during differentiation to early embryoid bodies (EBs) in 80% of the transcripts, with 72 pluripotency factors having less hm^5C. The hm^5C peaks were enriched in introns, UC-rich regions, and chromatin-associated nascent mRNAs (Lan et al. 2020). The authors went on to show that TET1 and TET2 bind to many additional sites by performing RIP-seq using a flag-tag knocked into the endogenous gene loci; TET3 was not studied as it is expressed at very low levels at this stage. TET1 bound to 7798 targets and TET2 bound to 6659, with substantial overlap. Therefore, there are roughly 3–4 times the number of TET-bound sites than hm^5C sites. Interestingly, mRNAs bound by TET in wild-type cells, regardless of hm^5C status, were significantly differentially expressed in TET-mutant cells (Lan et al. 2020). While this study goes on to report on the functional significance of hm^5C, this finding raises the possibility that TET is also performing other functions, be it recruiting other RNA-binding factors or competitive binding. Alternatively, some of this data may just represent noise in the sequencing assays. The Tet2 RNA binding domain (RBD) previously described to be located near the catalytic domain (He et al. 2016) is necessary for catalysis with roughly 70% of hm^5C sites requiring this domain but less so for RNA binding with only 30% of TET2 binding sites dependent on the RBD (Lan et al. 2020).

Finally, it has been shown that both 5-formylcytosine (f^5C) and 5-carboxylcytosine (ca^5C), the further oxidized moieties of hm^5C found in DNA, also exist in mRNA in vivo (Huang et al. 2016; Zhang et al. 2016). The ability of Tet1 to catalyze the conversion of f^5C to ca^5C was also verified in vitro (Basanta-Sanchez et al. 2017).

4 The Function of Cytosine Methylation in mRNA

Descriptive studies of m^5C and hm^5C sites have shown that these marked bases are present in transcripts that are actively translated and enriched in cell functions critical to specific cell types, indicating that RNA cytosine methylation also has a role in regulation of gene expression (Fu et al. 2014; Delatte et al. 2016; Amort and Lusser 2017). Furthermore, the m^5C profile changes over the course of testis development on stage-specific transcripts and hm^5C is reduced in cancer tissues, suggesting a possible role of RNA demethylation in regulating normal development and cancer (Huang et al. 2016; Fu et al. 2014). Currently there is excitement about a possible role for RNA cytosine methylation in brain development and function owing to the fact that hm^5C levels are highest in brain mRNA and that NSUN2 loss of function has been implicated in human neurodevelopmental disorders (Dang et al. 2011). Further supporting this hypothesis is the neurodevelopmental phenotype observed in *dtet* knockout flies, a species in which DNA methylation is non-existent, suggesting

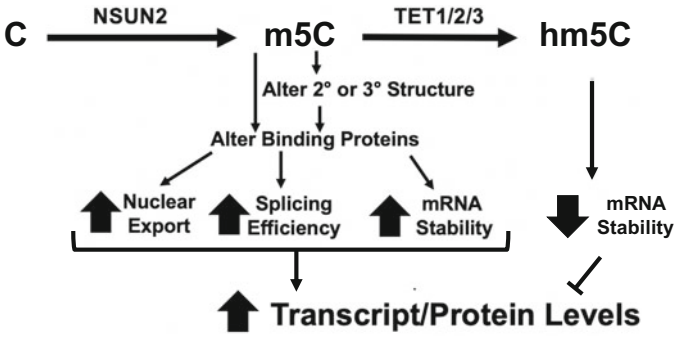


Fig. 2 Summary of studies investigating the functional role of m^5C and hm^5C in mRNA

in this case that TET function appears devoted to RNA demethylation. The *dtet* null flies have disorganized and smaller brains and ultimately die during the pupal stage (Delatte et al. 2016).

Some clues about the function of m^5C in mRNA can also be garnered from basic chemical studies of the behavior of m^5C in tRNA, where methylation promotes base stacking and increases the strength of the hydrogen bond with guanosine, both of which lead to an increase in structural stability (Motorin and Helm 2010; Hayrapetyan et al. 2009). We can therefore hypothesize that m^5C stabilizes structures in mRNA as well, allowing for differential binding by RNA binding proteins that can then regulate stability, localization, splicing, or translation efficiency.

Indeed, while only a handful of direct investigations have been performed to probe specific mechanisms for gene regulation by m^5C in mRNA, the results confirm a role for m^5C in increasing transcript or protein levels. Results of these studies are summarized in Fig. 2. The first showed that NSUN2 works together with ALYREF to promote methylated mRNA export from the nucleus in HeLa cells (Yang et al. 2017). A screen was performed for m^5C RNA binding proteins by cross-linking and pulling down synthesized oligonucleotides containing m^5C compared to unmethylated cytosine, then performing mass spectrometry to find binding proteins. ALYREF, the mRNA nuclear export adapter, was identified as one of the most highly enriched proteins, and this was confirmed by co-IP western blotting. Furthermore, knockdown of either NSUN2 or ALYREF decreased the cytoplasmic to nuclear mRNA ratio, consistent with decreased nuclear export. These phenotypes could be rescued by wild-type but not mutant proteins lacking the m^5C catalytic or binding motifs (Yang et al. 2017).

The mechanism of reduced emergency myelopoiesis in *Tet2* mutant mice was investigated leading to the discovery that TET2 is required to initiate myeloid development after infection (Zhang et al. 2018). Intriguingly, in this case it was demethylation of the 3′ UTR of *Socs3* mRNA, rather than DNA, within mast cells that was key to the expansion of these cells in a murine infection model. In *Tet2* mutant mice there was a transcriptome-wide increase in mRNA cytosine methylation according to RNA bisulfite sequencing, with hyper-methylated transcripts being

generally increased in the *Tet2* mutants and 60% of hyper-methylated transcripts also observed to bind to TET2 via cross-linking immuno-precipitation sequencing (CLIP-seq). *Socs3* RNA, encoding a member of the suppression of cytokine signaling family, showed both hyper-methylation and TET2 binding. There was also loss of A-to-I editing near the hyper-methylated region, indicating loss of interaction with the RNA adenosine deaminase ADAR1 in the absence of TET2. The authors went on to show that demethylation of *Socs3* mRNA allows a hairpin to form and ADAR1 to bind leading to degradation of the transcript. In the context of infection, TET2 normally demethylates *Socs3* RNA leading to its degradation and de-repression of JAK/STAT cytokine signaling to increase emergency myelopoiesis.

The function of Tet proteins in regulating neutrophil development and function via the methylation status of *socs3b* mRNA was recently revealed using a zebrafish model (Banks et al. 2021). This was the first study to demonstrate a role for TET proteins in regulating neutrophil development. A defect was found in the maturation of neutrophils in $tet3^{-/-}$ and $tet2^{-/-}tet3^{-/-}$ ($tet2/3^{DM}$) embryos and in $tet2^{-/-}tet3^{+/-}$ ($tet2/3^{MH}$) adults, such that these cells are blocked for maturation early in the granulation process. The relative importance of Tet2 and Tet3 in embryonic compared to adult neutrophils was interesting, with Tet3 being more important in the embryonic system and Tet2 in adult neutrophils, although at both timepoints there is some compensation among the two Tet isoforms. In light of the earlier study of murine mast cells, after recognizing a similar increase in *socs3b* transcript levels without DNA methylation changes in *tet*-mutant neutrophils, it was shown that there is a significant increase in methylation of the *socs3b* transcript in $tet2/3^{DM}$ neutrophils compared with wild-type neutrophils. However, in contrast to what was found in murine mast cells, most of the methylated cytosines were observed within the zebrafish *socs3b* gene body (coding sequence), with only one differentially methylated base in the 3′ UTR. Also, only half of the modified bases were in the CpG context, consistent with an earlier report characterizing m^5C in mRNA (Yang et al. 2017). The *socs3b* gene is downregulated between 35 hpf and 48 hpf in wild-type neutrophils, at precisely the time when neutrophils undergo granulation in this system. CRISPR-mediated knockout of *socs3b* rescued the granulation defect in $tet2/3^{DM}$ embryos, demonstrating the importance of reducing *socs3b* expression levels for neutrophil maturation. Finally, by mutating differentially methylated cytosines at wobble positions in *socs3b* mRNA, reducing methylation of the transcript was shown to decrease stability of the mutant mRNA compared to the wild-type RNA when injected into wild-type embryos. Again, methylation status of the transcript is related to its stability, with increased m^5C levels conferring an increase in half-life of the transcript (Banks et al. 2021).

Another recent report characterized RNA methylation in very early zebrafish development (0 to 6 hpf), revealing its role in regulating the maternal to zygotic transition through stabilization of transcripts. At 6 hpf, the number of m^5C sites was dramatically reduced (2902 sites) compared to 0 hpf (6500 sites) (Yang et al. 2019). Methylated cytosines on maternal mRNA increased the stability of these transcripts as measured by comparing RNA-seq profiles before and after transcription inhibition with RNA polymerase II inhibitor *a*-amanitin. The study identified *ybx1*, a known

mediator of RNA stability, as another RNA-m^5C binding protein through RNA affinity chromatography and mass spectrometry (MS) analyses. Further analyses showed that *ybx1* is required for gastrulation and acts by recruiting *papbc1a* to stabilize transcripts (Yang et al. 2019).

Finally, comparison of wild-type and *Tet1/2/3* triple knockout (TKO) mESCs demonstrated that the presence of hm^5C decreases transcript stability (Lan et al. 2020). By comparing hm^5C-containing transcripts identified by hMeRIP to publicly available databases relating mRNA half-life in mESCs, it was found that hm^5C-containing transcripts had a significantly shorter half-life. Through *a*-amantin transcriptional arrest experiments, there was found to be an increase in mRNA half-life in *Tet* TKO mESCs compared to wild-type, with far fewer destabilized transcripts in the TKO cells. Finally, following transfection of mESCs with synthesized mRNA containing either unmodified cytosine or hm^5C, it was observed that hm^5C-containing transcripts were roughly four times less stable than unmodified mRNAs (Lan et al. 2020). This study also indicates the importance of proper mRNA methylation in splicing, in addition to stability. The hm^5C marks were enriched in introns and splicing factor binding sites in mESC RNA. Indeed, *Tet* TKO mESCs had a higher ratio of unspliced to spliced transcripts compared to wild-type cells (Lan et al. 2020).

5 Conclusions and Future Prospects

The presence of m^5C in mRNA has been known for decades, although its function has remained largely a mystery. It was only in the last few years that sequencing methodologies have allowed the fine mapping of this modification necessary to begin to address these questions. As more transcriptome-wide m^5C and hm^5C sequencing studies are performed, it is becoming clear and certainly confusing that the studies are discordant in terms of the number, location, and distribution of methylation in mRNA (Table 1). One confounder is that these studies were performed using samples from different species and cell types and that the pattern of methylation is highly dependent on cell type. This hypothesis is supported by the fact that gene ontology for methylated transcripts generally fits with the type of cell being studied. If indeed lineage-specific differences exist in mRNA cytosine methylation and demethylation, it leaves open the question of how specific transcripts are being differentially methylated in various cell types. Perhaps there is some co-factor basis for sequence specificity, as is hypothesized to occur via transcription factor interactions in DNA, which will be an interesting avenue of future study in this field.

However, it seems likely that limitations in our current sequencing techniques also play a role in these discordant results. Bisulfite sequencing is a notoriously noisy protocol, which is likely more problematic in studies of RNA than DNA given the inherent instability of RNA as a molecule. Immunoprecipitation-based sequencing assays are also generally quite noisy and dependent on antibody specificity, which unfortunately is quite low both for m^5C and hm^5C antibodies. Perhaps with the next

iteration of transcriptome m^5C profiling, a clearer picture will emerge. In this regard, nanopore sequencing is quite promising, as it allows the direct sequencing of mRNA and all base modifications. This has the added benefit of simultaneously identifying all base modifications which will also allow a better understanding of how different marks are interacting. Unfortunately, base calling using nanopore sequencing has in the past been quite error prone, and better data may rely on further technology development.

It also remains to be shown precisely what role TETs are playing in the mRNA context. Given the rapid turnover of mRNA, it seems unlikely that mRNA transcripts are actively demethylated back to cytosine as occurs in DNA. Perhaps the conversion of m^5C to hm^5C and to a lesser extent f^5C and ca^5C are the end of this process, which could either alter mRNA binding protein profiles or directly lead to chemical instability and degradation of these transcripts to impact mRNA metabolism. Alternatively, TET enzymes may function to block methylation of RNA by NSUN methyltransferases through competitive binding as sometimes occurs with DNMT3 in DNA (Charlton et al. 2020). One finding that is difficult to explain in RNA m^5C methylation is that there is relatively much lower hm^5C compared to m^5C, which could be due either to instability of mRNAs containing oxidized bases or that TET functions mostly to block methylation rather than oxidize m^5C in this context.

Finally, this compilation of data leaves open the question of why such a complex and energetically costly system for epitranscriptomic regulation of gene expression would arise. Given the impact of mRNA methylation on stability, this system is well suited to allow fine-tuning of transcripts that require rapid turnover in certain contexts such as specific stages of development, for example, underlying pluripotency or immune cell activation. Another possibility is that these modifications on RNA are being made in conjunction with epigenetic changes on the same genes in DNA. Given that some of the players seem to be the same (TET proteins and MECP2), it leaves the possibility that these changes are occurring via one large complex. This would allow simultaneous fast tuning of gene expression via post-transcriptional control of mRNA while slower but longer-lived epigenetic signals are also being laid down. Future study could focus on how DNA and RNA methylation profiles change together over development or disease states.

References

Aas PA, Otterlei M, Falnes PO et al (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature 421:859–863

Adams JM, Cory S (1975) Modified nucleosides and bizarre 5′-termini in mouse myeloma mRNA. Nature 255:28–33

Amort T, Lusser A (2017) Detection of 5-Methylcytosine in specific poly(a) RNAs by bisulfite sequencing. Methods Mol Biol 1562:107–121

Banks KM, Lan Y, Evans T (2021) Tet proteins regulate neutrophil granulation in zebrafish through demethylation of socs3b mRNA. Cell Rep 34:108632

Basanta-Sanchez M, Wang R, Liu Z et al (2017) TET1-mediated oxidation of 5-Formylcytosine (5fC) to 5-Carboxycytosine (5caC) in RNA. Chembiochem 18:72–76
Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307
Bohnsack KE, Höbartner C, Bohnsack MT (2019) Eukaryotic 5-methylcytosine (m^5C) RNA methyltransferases: mechanisms, cellular functions, and links to disease. Genes (Basel) 10
Charlton J, Jung EJ, Mattei AL et al (2020) TETs compete with DNMT3 activity in pluripotent cells at thousands of methylated somatic enhancers. Nat Genet 52:819–827
Dang EV, Barbi J, Yang HY et al (2011) Control of T(H)17/T(reg) balance by hypoxia-inducible factor 1. Cell 146:772–784
Delatte B, Wang F, Ngoc LV et al (2016) RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. Science 351:282–285
Dubin D, Taylor R (1975) The methylation state of poly A-containing messenger RNA from cultured hamster cells. Nucleic Acid Res 2:1653–1668
Fu L, Guerrero CR, Zhong N et al (2014) Tet-mediated formation of 5-hydroxymethylcytosine in RNA. J Am Chem Soc 136:11582–11585
Hayrapetyan A, Grosjean H, Helm M (2009) Effect of a quaternary pentamine on RNA stabilization and enzymatic methylation. Biol Chem 390:851–861
He C, Sidoli S, Warneford-Thomson R et al (2016) High-resolution mapping of RNA-binding regions in the nuclear proteome of embryonic stem cells. Mol Cell 64:416–430
Huang W, Lan MD, Qi CB et al (2016) Formation and determination of the oxidation products of 5-methylcytosine in RNA. Chem Sci 7:5495–5502
Hussain S, Sajini AA, Blanco S et al (2013) NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep 4:255–261
Jeffery L, Nakielny S (2004) Components of the DNA methylation system of chromatin control are RNA-binding proteins. J Biol Chem 279:49479–49487
Lan J, Rajan N, Bizet M et al (2020) Functional role of Tet-mediated RNA hydroxymethylcytosine in mouse ES cells and during differentiation. Nat Commun 11:4956
Lister R, Ecker JR (2009) Finding the fifth base: genome-wide sequencing of cytosine methylation. Genome Res 19:959–966
Liu Y, Chen D, Su R et al (2020) iRNA5hmC: the first predictor to identify RNA 5-Hydroxymethylcytosine modifications using machine learning. Front Bioeng Biotechnol 8:227
Motorin Y, Helm M (2010) tRNA stabilization by modified nucleotides. Biochemistry 49:4934–4944
Squires JE, Patel HR, Nousch M et al (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40:5023–5033
Warren L, Manos PD, Ahfeldt T et al (2010) Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell 7:618–630
Xu Q, Wang K, Wang L et al (2016) IDH1/2 mutants inhibit TET-promoted oxidation of RNA 5mC to 5hmC. PLoS One 11:e0161261
Yang Y, Wang L, Han X et al (2019) RNA 5-Methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA decay. Mol Cell 75 e11:1188–1202
Yang X, Yang Y, Sun BF et al (2017) 5-methylcytosine promotes mRNA export–NSUN2 as the methyltransferase and ALYREF as an m5C reader. Cell Res 27:606–625
Young JI, Hong EP, Castle JC et al (2005) Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. Proc Natl Acad Sci USA 102:17551–17558
Zhang Q, Shi Y, Shi Q et al (2018) Tet2 promotes pathogen infection-induced myelopoiesis through mRNA oxidation. Nature 554:123–127
Zhang HY, Xiong J, Qi BL et al (2016) The existence of 5-hydroxymethylcytosine and 5-formylcytosine in both DNA and RNA in mammals. Chem Commun (Camb) 52:737–740

Discovery, Processing, and Potential Role of Noncanonical Caps in RNA

Hao Hu, Nora Flynn, and Xuemei Chen

Contents

Hao Hu and Nora Flynn contributed equally.

H. Hu
Department of Botany and Plant Sciences, Institute of Integrative Genome Biology, University of California, Riverside, CA, USA

Key Laboratory for Biology of Horticultural Plants, Ministry of Education, College of Horticulture & Forestry Sciences, Huazhong Agricultural University, Wuhan, China
e-mail: haohu@ucr.edu

N. Flynn · X. Chen (✉)
Department of Botany and Plant Sciences, Institute of Integrative Genome Biology, University of California, Riverside, CA, USA
e-mail: nflyn002@ucr.edu; xuemei.chen@ucr.edu

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_17

Abstract The modification of RNA species has been well documented for decades. In mRNA, both internal and 5′ end modifications can occur. Specifically, modification of the 5′ end is known as capping. The 5′-5′ triphosphate linked N7-methyl guanosine (m^7G) structure is involved in a myriad of RNA processes and was long presumed to be the only functional cap. In recent years, this view was overturned by reports that nicotinamide adenine dinucleotide (NAD^+), a critical redox cofactor, can serve as an RNA cap in bacteria. Subsequently, yeast, mammalian, and plant RNA species were also found to harbor the NAD^+cap. Apart from NAD^+, other noncanonical nucleotide analogs, including NADH, FAD, dpCoA, UDP-Glucose, and UDP-N-acetylglucosamine, were found to be caps in endogenous RNA, suggesting that a wide repertoire of RNA caps may be present. However, the functions of noncanonical capping remain mostly unknown. This chapter describes the detection methods for noncanonical RNA caps, their mode of capping and decapping, and their potential molecular and biological functions. The discovery of noncanonical caps represents a revolution in research on RNA modifications and prompts future efforts to delve into novel epitranscriptomic processes, which may link cellular metabolism with gene expression.

Keywords Noncanonical caps · LC-MS analysis · NAD^+ captureSeq · NAD^+ tagSeq · Capping mechanism · Decapping enzymes · Nudix · DXO · RNA stability · Translation initiation

1 Introduction

Beyond the base genetic code provided by DNA and RNA, various kinds of chemical or “epigenetic” modifications to these structures provide another layer of information. Functionally, the addition of these chemical marks can be recognized by specific proteins, leading to versatile gene expression without changing the genetic sequence itself (Roundtree et al. 2017; Shi et al. 2019; Boo and Kim 2020). The critical need to understand RNA epigenetic modifications set off the research field known as epitranscriptomics (Fu and He 2012; Chen et al. 2020).

While both RNA and DNA can be modified, RNA modifications play a more direct role in dynamically tuning transcript output, such as by affecting stability and translatability (Chen et al. 2016). Since the first modified nucleotide in RNA was discovered in the 1960s (Cohn 1960), more than 170 different RNA modifications have been identified in coding and noncoding RNA (Boccaletto et al. 2018; Nachtergaele and He 2018). tRNA and rRNA contain the most modifications, including 2’-O-methylation and pseudouridylation (Roundtree et al. 2017). On the

other hand, the known modifications on mRNA are less diverse, but contribute more to shaping the cellular transcriptome. mRNA modifications include internal modifications, such as N6-methyladenosine (m^6A), N6, 2′-O-dimethyladenosine (m^6Am), N1-methyladenosine (m^1A), pseudouridine (Ψ), 5-methylcytidine (m^5C), 5-hydroxymethylcytosine (hm^5C), and N4-acetylcytidine (ac^4C), as well as modifications of the 3′ end, known as the poly(A) tail and oligo(U) tail, and modification of the 5′ end, known as the caps (Boo and Kim 2020).

In particular, 5′ end capping is a critical determinant of the fate of an RNA. The 5′ cap is known to play a pivotal role in numerous RNA metabolic processes, such as polyadenylation (and possibly oligouridylation), pre-mRNA splicing, mRNA export, transcript stability, and translation initiation. Thus, this structure is mechanistically involved in every stage of the mRNA lifecycle (Ramanathan et al. 2016; Galloway and Cowling 2019). The predominant 5′ cap of mRNA is the 7-methylguanosine moiety linked via a 5′ to 5′ triphosphate chain to the first transcribed nucleotide, which is abbreviated as m^7GpppN and known as cap 0. Incorporation of the m^7GpppN cap is accomplished through characterized enzymatic activities (Ramanathan et al. 2016). In addition, the first and second transcribed nucleotides can be methylated on the ribose 2'-O position, resulting in m^7GpppNm or m^7GpppNmNm structures referred to as cap 1 and cap 2, respectively (Werner et al. 2011). In cap 1, when the first nucleotide is adenosine, another N6-methylation may also be observed at a ratio that reaches up to 20% in human cells (Mauer et al. 2017). These m^7G-related cap structures, or canonical caps, have been observed at varied levels in specific tissues and cells and could be differentially regulated in specific biological processes (Wetzel and Limbach 2016; Sikorski et al. 2020). However, despite the important roles of the m^7GpppN cap, this cap is found only in eukaryotes (Galloway and Cowling 2019).

Until recently, in bacteria, such as *Escherichia coli* (*E. coli*), it was assumed that the 5′ end of RNA consisted only of a 5′ triphosphate. This was overturned when nicotinamide adenine dinucleotide (NAD^+) was identified as a cap in *E. coli* RNA (Chen et al. 2009). NAD^+ is a pyridine dinucleotide and is an electron carrier involved in oxidation-reduction reactions, making it a key component of cellular signaling (Gakière et al. 2018). As NAD^+ contains an adenosine moiety, it may be recognized by RNA polymerase and incorporated into the 5′ end of RNA. After the identification of NAD^+-capped RNA in *E. coli*, yeast, mammalian, and plant RNA species were also found to harbor the noncanonical NAD^+ cap (Cahová et al. 2015; Jiao et al. 2017; Walters et al. 2017; Kiledjian 2018; Julius and Yuzenkova 2019; Wang et al. 2019b; Zhang et al. 2019a). Additionally, other adenosine-containing metabolites, such as dpCoA and FAD, were also found to initiate transcription in vitro (Huang 2003).

To date, many noncanonical nucleotides have been reported to prime RNA transcription by RNA polymerases from different organisms (Fig. 1) (Wang et al. 2019a; Doamekpor et al. 2020a; Hudeček et al. 2020). In this chapter, we aim to summarize the various types of noncanonical caps in different organisms, the detection methods used to identify these structures, the mechanism of incorporation

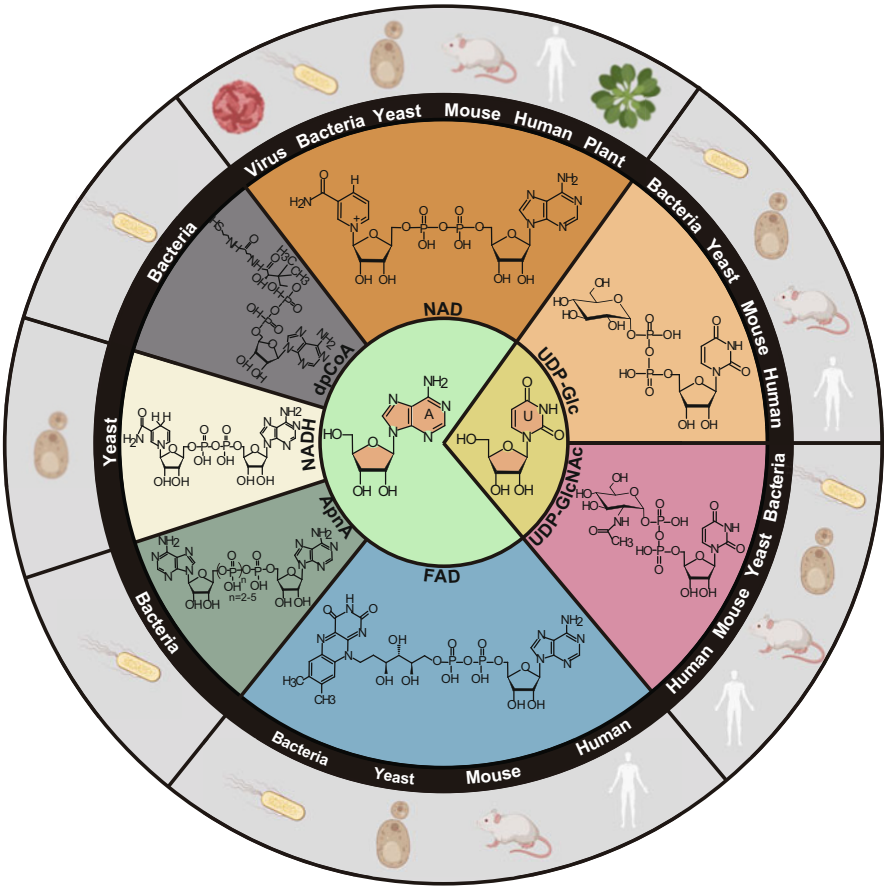


Fig. 1 Noncanonical RNA caps discovered in different organisms. RNA cap structures that have been discovered to date are classified into adenosine-containing nucleotides and uridine-containing nucleotides in the center of the circle. In the middle ring, the structure of each cap is displayed. The outer ring indicates the organisms reported to contain each RNA cap. The cartoons of the organisms were created with BioRender.com. dpCoA, dephospho-coenzyme A; GlcNAc, N-acetylglucosamine

into a transcript, and the possible regulation and biological functions of noncanonical caps, with an emphasis on the NAD^+ cap.

2 Discovery and Detection of RNA Modifications in Cells

RNA modifications can be detected, mapped, and quantified through various methods, although there are many challenges associated with the characterization of mRNA modifications in particular (Helm and Motorin 2017). For example, unlike

non-coding RNAs that have relatively abundant modifications, mRNAs have low levels of modified nucleotides. The most abundant one, m^6A, was estimated to only apply to 0.2% of total adenosine in cellular mRNA, equivalent to 2–3 nucleotides per transcript (Meyer et al. 2012). An additional challenge arises due to the different chemical properties of each modified residue. The need for specific detection and mapping methods for diverse RNA modifications resulted in the rise of an assortment of techniques.

The long-established method for the global detection and quantification of RNA modifications is thin layer chromatography (TLC), which relies on radioactive ^{32}P labeling for sensitivity (Grosjean et al. 2007; Kellner et al. 2010). TLC was later supplemented by high-performance liquid chromatography (HPLC) coupled to mass-spectrometry (MS) (Thüring et al. 2016; Wetzel and Limbach 2016). In these methods, the modified nucleotides are released from mRNA by complete chemical or enzymatic digestion and identified according to their chromatographic retention times and fragmentation patterns. However, while these methods have opened the door for the detection of various RNA modifications, they do not provide information on the exact localization of the modifications.

Recently, next-generation sequencing (NGS) technologies have paved the way for mapping mRNA modifications. It was observed that some modified RNA nucleotides can naturally block primer extension or cause misincorporations during reverse transcription (RT), thus leaving a signature mark at modified sites in cDNA sequences (Ryvkin et al. 2013). However, naturally occurring abortive RT events due to RNA modifications are limited and do not apply to the majority of modifications (Helm and Motorin 2017; Schwartz and Motorin 2017). To expose further RNA modifications, mRNA can be treated with chemical reagents that react with specific modifications to change or enhance the RT signature. This type of methodology has been used to uncover Ψ, internal m^7G, and m^5C in specific RNAs (David et al. 2017; Zhang et al. 2019b).

Additionally, affinity-based enrichment of RNA modifications before high-throughput sequencing is highly beneficial for detection due to the low levels of modified residues in cellular RNA. Modified RNA can be selectively recognized by specific antibodies (Dominissini et al. 2012; Mishima et al. 2015; Li et al. 2016) or by clickable chemical reactions depending on the functional structure (Cahová et al. 2015). Recognition of specific, modified RNAs is followed by library preparation and sequencing, yielding information on the location and abundance of the modification. Despite the various methods to detect RNA modifications, challenges remain and information on many of these modifications is limited. Therefore, to reinforce these methods, currently used techniques to investigate RNA noncanonical capping are rapidly evolving.

2.1 Global Discovery and Detection of Noncanonical Capping in RNA

2.1.1 HPLC and MS Coupling

LC-MS has been extensively used in the detection and quantification of novel RNA modifications. In general, RNA is cleaved and fragments are subsequently analyzed, such as by fragmentation pattern or comparison to the calculated mass of the unmodified residue. This method is responsible for the discovery of NAD^+-linked RNA in *E. coli* and *Streptomyces venezuelae* in 2009 (Chen et al. 2009).

In this 2009 study, a group led by David Liu employed a workflow with key treatments to detect noncanonical nucleotides in RNA (Fig. 2a). First, cellular RNA is separated through size-exclusion chromatography, and the macromolecular fraction is retained. This macromolecular fraction is further treated with the nuclease P1, an endonuclease that generates mononucleotides with a 3′ hydroxyl group and a 5′-phosphate. The treated sample is then subjected to LC-MS. Using this method, 24 and 28 unknown small molecule-RNA conjugates were significantly enriched compared to untreated samples in *E. coli* and *S. venezuelae*, respectively (Chen et al. 2009). These candidate small molecules were shown to be cleaved from cellular RNA, as the detected amount decreased if samples were pretreated with RNase.

In both species, two molecules that were found to be highly enriched after P1 digestion were NAD^+ and dpCoA, alongside their derivatives. Both structures were present in shorter RNAs of lengths below ~200 nucleotides. The former showed a higher abundance at 3000 copies per cell, while the latter only displayed 100 copies per cell. Further repeated assays using isotopically labeled water revealed that NAD^+ and dpCoA in cellular RNA are attached to the 5′ terminus. This novel finding showed that adenosine-based noncanonical metabolites could serve as a cap structure in bacteria. Following detection in prokaryotes, the NAD^+ cap was detected using LC-MS techniques in eukaryotes such as yeast, mammalian cells, and the plant, *Arabidopsis* (Wang et al. 2019a; Zhang et al. 2019a).

The discovery of NADylated RNA prompted research efforts dedicated to the understanding of noncanonical RNA caps. Rapidly, more metabolites, all of which shared a nucleotide-containing structure, were found capable of being incorporated into the 5′ end of RNA by in vitro transcription. For example, FAD, a coenzyme involved in redox reactions, dinucleoside polyphosphate (Np_nN), a potential alarmone, and UDP-glucose and UDP-N-acetylglucosamine (UDP-GlcNAc), cell wall precursors, could all be incorporated as cap structures in vitro (Huang 2003; Julius et al. 2018; Hudeček et al. 2020). The attachment of such a range of substrates to the 5′ end of RNA suggested that more noncanonical caps could exist in vivo. However, untargeted LC-MS analyses have not detected these caps in vivo, perhaps due to the lack of sensitivity.

To detect and quantify more of the RNA capping landscape in vivo, targeted LC-MS analyses combine off-line HPLC enrichment of cap nucleotides with triple-quadrupole mass spectrometry to enable absolute quantification of a given RNA cap structure (Wang et al. 2019a). In targeted analyses, filter parameters can be pre-set to the mass size and retention time of specific chemicals via synthesized standards to

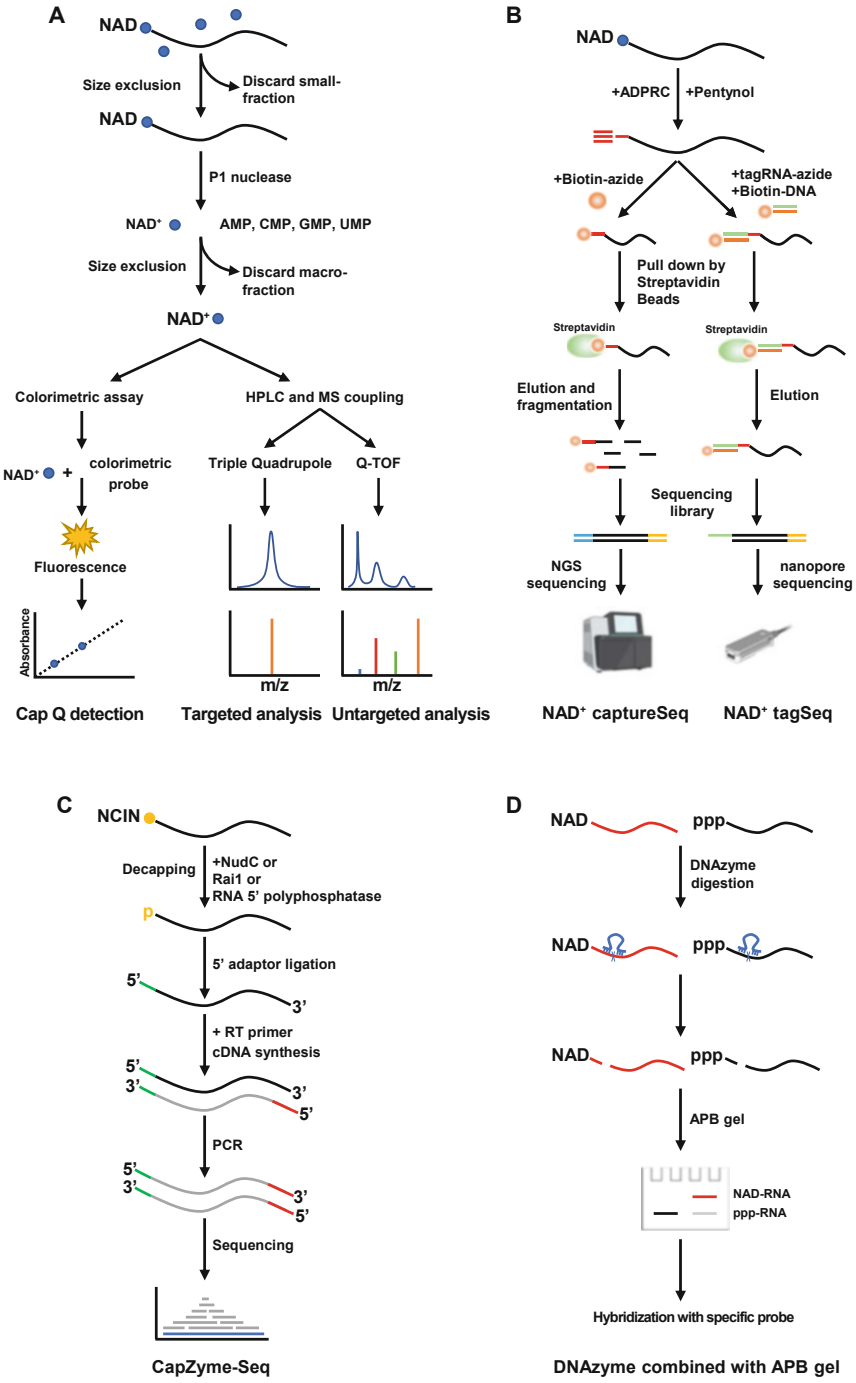


Fig. 2 Strategies for the detection of noncanonical RNA caps in vivo. (**a**) Global detection and quantification of noncanonical caps. Firstly, isolated RNA is separated from the small molecular weight fraction and treated with nuclease P1 or a decapping enzyme. After a secondary size

efficiently detect target molecules released from cellular RNA. In addition, highly accurate quantification can be achieved by combining isotope-labeled internal standards and a series of unlabeled external standards to generate a calibration curve. Using this technique, three novel metabolite caps (FAD, UDP-glucose, and UDP-GlcNAc) were discovered and quantified in virus, *E. coli*, yeast, mouse, and human cellular RNA. FAD and UDP-glucose caps accounted for <5 fmol/μg RNA (Wang et al. 2019a), but surprisingly, the UDP-GlcNAc cap was more abundant, reaching up to 28 fmol/μg RNA, higher than the NAD^+ cap and consistent with the relative abundance of such cellular metabolites in cells (Yang et al. 2007; Namboori and Graham 2008; Julius et al. 2018). Using the same LC-MS-based methodology, in *E. coli*, dinucleoside polyphosphates (Np_nN) were also detected as noncanonical caps in a short RNA fraction (Hudeček et al. 2020). The amount of Np_nN caps (Ap_3A, Ap_3G, Ap_5A) in small RNA (sRNA) was comparable to that of dpCoA (~ 75 fmol/μg sRNA) and much lower than that of NAD^+ (1900 fmol/μg sRNA). LC-MS experimentation further revealed that some of the Np_nN caps contained multiple methyl groups in the nucleotides (e.g. m^7Gp_4Gm, m^6Ap_3A), which maintained cap stability (Hudeček et al. 2020).

Untargeted LC-MS analysis provided an approach to discover novel RNA cap modifications but is hampered by the limited sensitivity of the MS detector and sample purity (Limbach and Paulines 2017). Conversely, targeted LC-MS analysis displays high accuracy and sensitivity, though it requires a synthesized standard and therefore can only be applied to previously identified structures. Weighing the strengths and weaknesses of each experimental approach is necessary to effectively address the desired research question.

2.1.2 CapQ Quantification

Cap detection and quantitation, known as CapQ, is another method used for RNA-cap identification that is both time-efficient and easily performed in the

Fig. 2 (continued) exclusion step, collected fractions are analyzed by colorimetric assays or coupled HPLC and MS. Quality or quantity is determined by comparison with a standard. (**b**) NAD^+-capped RNA capture and sequencing technologies. The nicotinamide moiety of NAD^+ is exchanged for an alkyne group by ADPRC, and the alkyne group undergoes a copper-catalyzed azide-alkyne cycloaddition reaction to link to a biotin moiety (in NAD^+ captureSeq) or a tagRNA that can be hybridized with a biotinylated DNA probe (in NAD^+ tagSeq). Biotinylated RNA is eluted and enriched by streptavidin beads. The profile of NAD^+-capped RNAs can be analyzed by high-throughput RNA sequencing. The sequencing machine cartoons were created with BioRender.com. (**c**) CapZyme-Seq workflow. Noncanonically capped RNA is first processed by decapping enzymes to yield a 5′ monophosphate end and then ligated with single-stranded oligonucleotide adaptors. Finally, 5′ end sequences are analyzed by high-throughput sequencing. (**d**) Validation technologies for individual RNA containing noncanonical caps. A specific RNA candidate is cleaved by a DNAzyme to yield short RNA 5′ end fragments. Capped RNA is distinguished from uncapped RNA in acrylaminophenyl boronic acid electrophoresis (APB) and then hybridization with a specific probe for candidate RNA transcripts can be performed

average laboratory (Fig. 2a). In general, the first step of CapQ is the same as LC-MS detection, where the intact 5′ end cap structure is released from RNA by enzymatic treatment, such as by nuclease P1. This step is followed by a colorimetric assay that affords measurement of the amount of released molecules based on an enzymatic cycling reaction.

Specifically, for detecting the NAD^+ cap, the released NAD^+ is reduced to NADH, which then reacts with a colorimetric probe to produce a colored product that can be measured at 450 nm. The intensity of the product color is proportional to the amount of NAD^+ in the test sample. Using this method, the extent of NAD^+ capping was determined to be ~120 fmol/μg RNA in *E. coli*, which is similar to previous estimates using an LC-MS approach (Chen et al. 2009; Grudzien-Nogalska et al. 2018). In other organisms, the level of NAD^+ capping is lower than in *E. coli* (80 fmol/μg in *S. cerevisiae*, 20 fmol/μg in HEK293T cell, 12 fmol/μg in *Arabidopsis*) (Grudzien-Nogalska et al. 2018; Wang et al. 2019b). The lower NAD^+ capping is reasonable as there is likely a dominant preference for the eukaryotic m^7G cap.

Similar to NAD^+, FAD can also be measured by a specific colorimetric assay. The recently developed FAD CapQ revealed that there is ~1 fmol FAD/μg of short RNAs in human cells. This is a comparable concentration to that measured by targeted LC-MS (Wang et al. 2019a; Doamekpor et al. 2020a).

While there are certainly benefits to the usage of CapQ methodology, this technique also has some shortcomings. For instance, so far, CapQ application is restricted to NAD^+ and FAD caps and relies on commercially available colorimetric assay kits. Nonetheless, compared to LC-MS detection, the CapQ method is highly suitable for comparisons of NAD^+ and FAD cap contents from different samples.

2.2 Next-Generation Sequencing Technologies for Use in the Study of Noncanonical Capping in RNA

The methods outlined above for the global quantification of noncanonical RNA caps cannot provide any information on the sequences harboring, or localizations of, these structures. Sequencing technologies have revolutionized epitranscriptomics research, affording the ability to map RNA modifications to specific transcripts and aiding in the illumination of the function of noncanonical caps.

2.2.1 NAD^+ captureSeq

After the discovery that RNA potentially possessed NAD^+ caps (Fig. 2b) (Chen et al. 2009), the precise transcripts that contained these caps were not profiled until 2015, when a next-generation sequencing technique known as NAD^+ captureSeq was established in *E. coli* (Cahová et al. 2015). NAD^+ captureSeq utilizes a

chemoenzymatic reaction to detect and identify NAD^+-capped RNA. In this reaction, adenosine diphosphate ribosylcyclase (ADPRC) removes the nicotinamide moiety from NAD^+-capped RNA. This step is followed by transglycosylation with an alkyne (such as pentynol) that reacts with the remaining 5′ end of RNA and subsequently click-chemistry-mediated biotinylation (Rostovtsev et al. 2002; Cahová et al. 2015). Thus, NAD^+-capped RNA is converted to biotinylated RNA, which can be captured and enriched by streptavidin beads and processed for high-throughput sequencing. Transcripts gleaned from this pipeline must be compared to a control background library without ADPRC treatment (ADPRC-) or to total RNA sequencing (RNA-seq) data. The transcripts that are significantly enriched in the ADPRC-treated sample are deemed to be NAD^+-capped (Cahová et al. 2015; Kwasnik et al. 2019).

Since its development, the NAD^+ captureSeq method has been widely utilized in many prokaryotes and eukaryotes, exposing new information on NAD^+-capped RNA. In *E. coli*, it was observed that the identified NAD^+-capped RNAs were mainly sRNAs involved in stress responses and mRNAs encoding enzymes involved in metabolism. The most abundantly NAD^+-capped sRNA was RNAI, which had 13% of its transcripts containing an NAD^+ cap (Cahová et al. 2015). Other than *E. coli*, the bacterium *B. subtilis* also exhibited NAD^+-capped RNA, but at a level 14-fold less than *E. coli*. In *B. subtilis*, NAD^+-capped transcripts were predominantly full-length mRNA, different from *E. coli*'s predisposition for NAD^+-capped sRNAs (Frindert et al. 2018). Through comparing the common sequence features of identified NAD^+-capped RNA, it was observed that most of the enriched RNA reads started with an adenosine, implying that NAD^+ caps are incorporated into RNA during transcription initiation (Bird et al. 2016). Interestingly, neither species of bacteria displayed ribosomal RNAs or transfer RNAs that were enriched for NAD^+-capped transcripts.

Following prokaryotes, NAD^+ captureSeq was applied to eukaryotes. In *Saccharomyces cerevisiae*, 1–5% of mRNA transcripts were shown to be modified by NAD^+ caps. Most of these transcripts were short RNAs involved in mitochondrial function and the translational machinery (Walters et al. 2017; Zhang et al. 2020). In human cells, NAD^+-capped mRNAs were detected, and the noncoding transcripts found to be preferentially capped included small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) (Jiao et al. 2017). Finally, in plants, NAD^+-capped RNAs were widespread throughout the transcriptome, except in chloroplast RNA, and these transcripts were found to be related to photosynthesis, protein synthesis, and stress responses (Wang et al. 2019b; Zhang et al. 2019a). NAD^+-capped RNAs were spliced and polyadenylated in both human cells and plants.

2.2.2 NAD^+ tagSeq

Based on the technique demonstrated in NAD^+ captureSeq, a modified approach called NAD^+ tagSeq allows for the full-length sequences of NAD^+-capped transcripts to be delineated by using single-molecule RNA sequencing (Fig. 2b). Similar

to NAD^+ captureSeq, ADPRC removes the nicotinamide of NAD^+-capped RNA, and subsequently an alkyne is introduced to the 5′ end of the RNA. However, instead of biotinylation, the copper-catalyzed azide-alkyne cycloaddition (CuAAC) reaction attaches a synthetic RNA, or tagRNA, that contains an azide group. The desired NAD^+-capped RNA, now linked to this RNA tag, is isolated by a DNA probe and sequenced using Oxford Nanopore sequencing technology. Sequencing starts from the polyA tail and ends with the 5′ end of transcripts. All sequence reads containing the tagged RNA are thus NAD^+-capped RNA. Through this method, features of NAD^+-capped RNA can be analyzed, revealing that the 5′ end of many NAD^+-capped RNAs are located around 30 to 400 bases downstream of canonical transcription start sites (TSS) in *Arabidopsis*. Therefore, NAD^+-capped RNAs tend to have shorter 5′ UTRs than m^7G-capped RNAs. NAD^+ tagSeq provides more accurate and broader information about NAD^+-capped RNA sequences than NAD^+ captureSeq but loses the capability to analyze very short (<100 nt) RNAs due to the use of nanopore sequencing (Zhang et al. 2019a).

Despite the genome-level NAD^+-capped RNA analysis offered by both NAD^+ captureSeq and NAD^+ tagSeq, there remain downfalls in using these techniques. One drawback revolves around the introduction of copper ions during the click chemistry CuAAC reaction. The introduction of copper ions is prone to causing RNA degradation, resulting in a bias toward the 5′ end (Liu et al. 2020). The density of reads at the 5′ end is increased through enrichment by streptavidin beads irrespective of the 3′ end. In addition, the alkyne moiety added during the first step seems capable of reacting with some other modified units in RNA in the absence of ADPRC, leading to nonspecific signals. For example, in *Arabidopsis* chloroplasts, the transcript level was comparably high in both ADPRC+ and ADPRC- samples, the latter presumably due to a false signal stemming from some other modification (s) in the RNA (Wang et al. 2019b).

2.2.3 CapZyme-Seq

NAD^+ captureSeq does not provide single nucleotide resolution of 5′ ends. Although NAD^+ tagSeq afforded the observation of full-length sequences, it still failed to determine the exact 5′ end sequence of NAD^+-capped RNAs due to inability to call bases at the junction between the tagRNA and the 5′ end of NAD^+-capped RNA (Cahová et al. 2015; Zhang et al. 2019a). Exact 5′ end high-throughput sequencing relies on adaptor ligation to RNA 5′ ends with a 5′ monophosphate. For RNA with noncanonical caps, a few decapping enzymes in various organisms were discovered that enable the removal of noncanonical caps, such as NAD^+, NADH, dpCoA, or FAD, resulting in a monophosphate at the 5′ end of the RNA (Jiao et al. 2017; Doamekpor et al. 2020a). A method that takes advantage of these decapping enzymes is CapZyme-Seq, which was established to identify the exact 5′ end sequence of RNA, as well as quantify the relative amount of noncanonically capped RNA or uncapped RNA (Fig. 2c).

CapZyme-Seq combines enzymatic removal of noncanonical caps or 5′ triphosphates with high-throughput sequencing. By performing CapZyme-Seq in *E. coli*, it was revealed that NAD^+-mediated initiation significantly preferred an adenosine at the TSS, while the capping efficiency for diverse promoter sequences varied. One sRNA with an A:T pair at the TSS position displayed a level of noncanonical capping of 22.4% compared to uncapped RNA. However, like the previous methods, this method also has limitations. For one, decapping enzymes may be unable to distinguish NAD^+ caps from other noncanonical caps. In addition, the different decapping enzymes used in CapZyme-Seq may exhibit various efficiencies for specific cap types or have differing specificities toward different RNAs with the same type of noncanonical cap, which could influence the results (Vvedenskaya et al. 2018).

The methodologies described above are most useful to study NAD^+-capped RNA. Unfortunately, other noncanonical caps still lack a robust sequencing technique to further explore their properties. Options for future studies on these caps could include methodologies using specific antibodies, affinity tagging through chemical reactions, or selective recognition by unique RNA or protein structures (Breaker 2012; Mishima et al. 2015).

2.3 *In Vitro Research and Validation Technologies for RNA with Noncanonical Caps*

Previous sections have detailed the powerful techniques used for global noncanonical cap detection in vivo. There are also simple tools available for studies of noncanonical capping in vitro. Most commonly, ^{32}P radioactively labelled capped RNA is analyzed by TLC. This method is usually used to examine the incorporation of a cap and the efficiency of sequence extension during in vitro transcription (Julius et al. 2018). In addition, acrylaminophenyl boronic acid electrophoresis (APB) provides a visual, user-friendly technique that allows distinction of the less-mobile, capped RNA containing a vicinal-diol moiety, such as m^7G, NAD^+, NADH, FAD, or Np_nN, from uncapped RNAs (Nübel et al. 2017; Luciano et al. 2019). For individual transcripts in vivo, APB gels also serve as a powerful validation tool for identification of capped RNAs. Combined with defined, specific oligodeoxynucleotide-mediated RNA cleavage (DNAzyme), which processes RNA to yield short 5′ end-containing fragments (Joyce 2001), APB gels can identify the noncanonical capping of RNAs, as well as distinguish between different capped species by comparing to synthetic noncanonically capped RNA standards (Fig. 2d) (Bird et al. 2018).

3 Mechanism of Noncanonical Capping

The mechanism involved in canonical m^7G capping has been clearly defined. After transcription initiation, addition of the m^7G cap is accomplished by a capping complex that interacts with the nascent RNA of ~20–25 nt (Shuman 2015). On the other hand, the mechanism of noncanonical capping in vivo requires further research, as current data is conflicting. After the discovery of NAD^+- and dpCoA-capped RNAs in 2009, in vitro experimentation failed to incorporate these caps into RNA, suggesting that noncanonical cap addition depended on post-transcriptional processes in vivo (Chen et al. 2009; Kowtoniuk et al. 2009). In contrast, earlier research had used *E. coli* RNA polymerase (RNAP) to successfully synthesize short transcripts initiated with NAD^+ or FAD (Malygin and Shemyakin 1979). More recently, evidence has accumulated that supports the incorporation of noncanonical caps by RNA polymerase during transcription initiation. Firstly, it has been demonstrated that eukaryotic RNAPs can use different noncanonical caps to initiate transcription (Bird et al. 2016; Julius and Yuzenkova 2017). Additionally, in vivo, NAD^+-capped RNA displays similar levels of enrichment on pre-mRNAs as on mRNAs, suggesting that NAD^+ is added cotranscriptionally (Walters et al. 2017; Bird et al. 2018; Sharma et al. 2020). Overall, there appear to be several mechanisms to achieve noncanonical capping of RNA in vivo, and these mechanisms can be affected by multiple factors.

3.1 RNA Polymerase

RNAPs are key enzymes in the delivery of genetic information from DNA to RNA through transcription. Usually, RNAP uses four NTPs (ATP, CTP, GTP, UTP) as substrates to initiate and extend RNA sequences. However, noncanonical substrates besides NTPs, such as coenzymes and long oligoribonucleotides, can also prime transcription by RNAP in vitro and in vivo (Fig. 3). For example, the bacteriophage T7 RNAP can use adenosine-containing NAD^+, FAD, dpCoA, and Np_nN to initiate transcription in vitro (Huang 2003; Hudeček et al. 2020). Structural research shows that such substrates can be accommodated in the space provided by the nucleotide-binding pocket of the T7 RNAP (Durniak et al. 2008). One exception is that T7 RNAP inefficiently incorporates NADP into transcripts, possibly owing to the additional phosphate group that causes steric hindrance in the pocket (Julius and Yuzenkova 2017).

The bacterial RNAP (both in *E. coli* and *B. subtilis*) and eukaryotic RNAP II can also incorporate NAD^+, NADH, dpCoA, and Np_4A into RNA during transcription initiation and can extend the sequence length from 2 nt up to 75 nt in vitro depending on the promoter (Bird et al. 2016; Frindert et al. 2018; Luciano and Belasco 2020). In

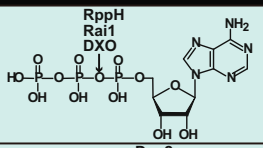
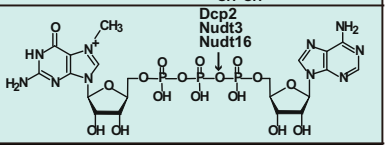
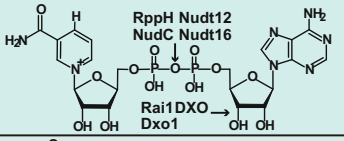
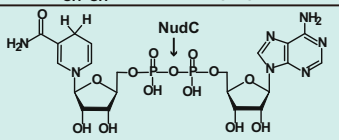
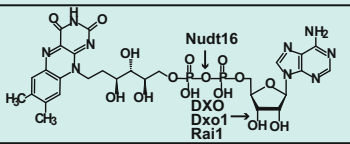
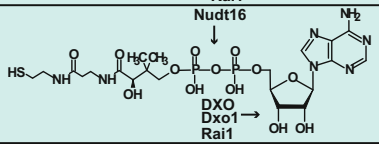
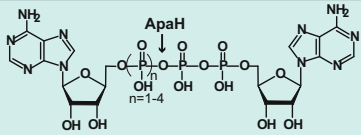
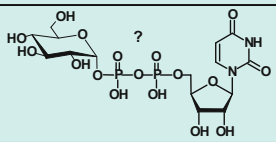
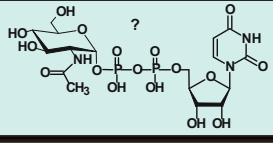
Caps	Capping enzyme	Decapping enzyme
Canonical 5' end		
ATP	*E. coli* RNAP Eukaryotic RNAP II Mitochondrial RNAP Chloroplast RNAP	
m^{7}G	Phosphatase Guanylyltransferase Methyltransferase	
Noncanonical 5' end		
NAD$^+$	*E. coli* RNAP Eukaryotic RNAP II Mitochondrial RNAP	
NADH	*E. coli* RNAP Eukaryotic RNAP II Mitochondrial RNAP	
FAD	*E. coli* RNAP Eukaryotic RNAP II Mitochondrial RNAP	
dpCOA	*E. coli* RNAP Eukaryotic RNAP II Mitochondrial RNAP	
APnA	*E. coli* RNAP Eukaryotic RNAP II	
UDP-Glc	*E. coli* RNAP	
UDP-GlcNAc	*E. coli* RNAP	

Fig. 3 The capping and decapping enzymes for canonical and noncanonical 5′ caps of RNA. The reported capping enzymes for each metabolite that can incorporate into RNA 5′ ends and the main decapping enzymes for each cap structure are shown. Decapping enzyme cleavage sites are

particular, *E. coli* RNAP is 60 times more efficient when incorporating Np_4A than NAD^+, which may be due to the presence of only two bridging phosphates between the two nucleosides in NAD^+. This is consistent with data showing that the incorporation efficiency of ADP only reaches up to 20% of that of ATP (Luciano and Belasco 2020). The cell wall synthesis precursors UDP-Glc and UDP-GlcNAc could also be incorporated into RNA by *E. coli* RNAP as pyrimidine-containing initial nucleotides, and even have higher extension efficiencies than UTP. The Km values of these noncanonical substrates (NAD^+ ~ 0.36 mM; UDP-Glc and UDP-GlcNAc ~0.3 mM) during transcription initiation by *E. coli* RNAP are much lower than their cellular concentrations, revealing the efficiency of incorporation as an RNA cap in vivo (Julius and Yuzenkova 2017).

However, the nuclear RNAP may not be the only polymerase responsible for incorporation of noncanonical caps. Other than nuclear RNAs, up to 15% and 60% of NAD^+-capped RNA in human and yeast cells, respectively, were attributed to mitochondrial transcripts (Bird et al. 2018), indicating that the mitochondrial RNAP is likely also responsible for the addition of noncanonical caps. In vitro transcription assays showed that yeast mitochondrial RNAP can use NAD^+ and NADH as initial substrates, and that human mitochondrial RNAP can also use other noncanonical substrates, such as FAD and dpCoA, to initiate transcription (Bird et al. 2018; Julius et al. 2018). The efficiency of transcription when initiating with NAD^+ is 40%–60% as efficient as initiating with ATP for yeast and human mitochondrial RNAP. Initiation with NAD^+ by mitochondrial RNAP is about 10- to 40-fold more efficient than that by *E. coli* RNAP and *S. cerevisiae* RNAP II (Bird et al. 2018). This difference in efficiency may be due to differences in the sequences and structures of nuclear and mitochondrial RNAPs. Additionally, the mitochondrial and T7 RNAPs are single-subunit RNAPs, while the *E. coli* RNAP and *S. cerevisiae* RNAP II are multi-subunit, which may lead to quantitative differences in the efficiency of noncanonical capping (Ringel et al. 2011; Bird et al. 2018; Hillen et al. 2018).

Other polymerases that could be involved in the addition of noncanonical caps are plastid polymerases in organisms such as plants. For example, in plants, plastids contain two types of RNAPs, the nuclear-encoded single subunit RNAP (NEP) and the plastid-encoded multi-subunit RNAP (PEP) (Gray and Lang 1998). However, no study has reported their ability to initiate noncanonical caps on transcripts. The failure to detect NAD^+-capped RNA in *Arabidopsis* chloroplasts may imply that chloroplast RNAPs are incapable of incorporating noncanonical caps (Wang et al. 2019b).

Fig. 3 (continued) displayed. RppH for NAD^+ decapping is specifically from *B. subtilis*. NudC for NAD^+ decapping is specifically from *E. coli*. ApaH for Ap_nA decapping is specifically from *E. coli*. Some decapping enzymes have homologs in various organisms. There are no reports of decapping enzymes for UDP-Glc and UDP-GlcNAc

3.2 σ Factors and RNAP Structure

RNAP-dependent transcription initiation requires association with σ factors to recognize template sequences. The number of σ factors varies, altering selection of the gene targets of RNAP (Paget 2015; Barvík et al. 2017), and potentially playing a role in capping with noncanonical caps. For example, in *E. coli*, the RNAP holoenzyme with the σ^s or housekeeping σ^{70} factors produces most of the transcripts in the stationary or exponential phases, respectively. NAD^+-capped transcript levels differ during these two phases, implying that certain RNAP factors may be involved in the specific capping of transcripts. However, no differences in capping efficiency were found in vitro between RNAP with these two σ factors when using the substrates ATP, NAD^+, NADH, and FAD, suggesting that they do not have a preference for cellular substrates (Julius and Yuzenkova 2017).

One σ factor region has demonstrated some impact on the noncanonical capping of RNA. The region 3.2 of σ^{70} has been shown to protrude into the catalytic site of RNAP and affect nucleotide incorporation at the 5' end of transcripts (Kulbachinskiy and Mustaev 2006). Mutation of region 3.2 of σ^{70} did not influence the incorporation of some noncanonical caps; however, intriguingly, RNAP acquired the ability to incorporate a complex cell wall precursor, UDP-MurNAc-pentapeptide. This suggests that region 3.2 may serve as protection against the incorporation of nucleotides with long side chains (Julius and Yuzenkova 2017). Other than this example, no σ factors have effectively been demonstrated to alter the transcription initiation of RNA with noncanonical caps. Nonetheless, other alternative σ factors may have the potential to affect incorporation of noncanonical caps. For example, the *E. coli* gene GlmY, which produces NAD^+-capped transcripts, contains the recognition sequences for σ^{54}, implying that this σ factor may be involved in noncanonical capping (Göpel et al. 2011; Cahová et al. 2015).

Additionally, the Rif pocket of RNAP is an important structural determinant for noncanonical capping, as the nascent transcripts both make contact with and pass through the Rif pocket. Crystal structures of the *E. coli* RNAP complex show the nicotinamide moiety of the NAD^+ nucleotide interacting with residues D516 and H1237 of this pocket (Bird et al. 2016). Mutation of D516 indeed strongly decreased the NAD^+ utilization efficiency (Julius and Yuzenkova 2017). However, in *B. subtilis*, no altered efficiency of NAD^+ capping was observed when the *E. coli* homologous site for the Rif pocket was mutated (Frindert et al. 2018). The cell wall precursors, UDP-Glc and UDP-GlcNAc, were also not affected by the amino acid substitutions in the Rifampicin binding pocket. This could be because they may not make specific contact with the amino acids of the Rif pocket (Julius and Yuzenkova 2017). Finally, while the addition of Rifampicin to the transcription reaction inhibited the extension of ATP-initiated transcripts due to its ability to block transcription elongation (Campbell et al. 2001), NAD^+-capped short RNAs were not affected, suggesting that the 5′ NAD^+ prevents Rifampicin binding to RNAP and thus stabilizes these short transcripts (Julius and Yuzenkova 2017). Collectively, the influence of the Rifampicin pocket as a determinant for capping might depend on

different RNAPs in various organisms, as well as the noncanonical substrates themselves.

3.3 Promoter Sequence

Another determinant of incorporation of a noncanonical cap is the promoter sequence. Experimentation using in vitro transcription suggested that noncanonical cap initiation only occurs from template DNA containing A:T at the transcription start site (+1) (Bird et al. 2016). In the case of *E. coli*, the RNA polymerase selects a position not far downstream (ranges from 7 to 10 nt) of the promoter −10 element as TSS. Normally, TSS selection for NAD^+-mediated initiation differs from that of NTPs due to this strong preference for an A:T base pair at the TSS position (Vvedenskaya et al. 2018). To put the selection preference for NAD^+-mediated initiation into perspective, half of the TSS selected by bacterial and eukaryotic RNAPs are +1A, whereas all of the TSS selected by yeast and mitochondrial RNAPs are +1A (Tsuchihara et al. 2009; Thomason et al. 2015; Bird et al. 2018). This preference for TSS further demonstrated that noncanonical capping is accomplished via transcription initiation, rather than post-transcriptional mechanisms.

Besides the TSS, the promoter sequence close to the TSS strongly affects the efficiency of capping. Bird et al. (2016) demonstrated that NAD^+ capping with the *E. coli* RNAI and gadY promoters exhibits higher efficiencies than with the PN25 and PT7A1 promoters. This is consistent with the relative extent of NAD^+-capped transcripts attributed to each after detection in vivo (Cahová et al. 2015). Further analysis revealed that the identity of the base −1 upstream of the TSS plays a particularly important role in NAD^+ capping efficiency. This may be due to the nicotinamide moiety of NAD^+ interacting with the −1 position, thus leading to different efficiencies depending on the identity of the −1-position base, with G facilitating NAD^+ capping and C repressing it (in the coding sequence) (Bird et al. 2016; Vvedenskaya et al. 2018). This trend was also observed in the *B. subtilis* veg promoter, where a T to C transition (in the coding strand) decreased the amount of NAD^+ capping by around 40% both in vitro and in vivo. Only 9% of the promoters of all NAD^+-capped RNAs contain a C at the −1 position (Frindert et al. 2018). Additionally, in *Staphylococcus aureus*, the efficiency of NAD^+ capping in RNAIII transcripts depends on the −1 position of the P3 promoter, further supporting this view (Morales-Filloy et al. 2020). All of these alterations in efficiency could be explained by the nicotinamide moiety experiencing severe steric hindrance with the template strand A or G at the −1 position (Vvedenskaya et al. 2018). However, later studies argued that the preference of NAD^+ at the −1 position is not specifically due to pairing of the nicotinamide moiety with the −1 base because the same trend was also observed for ATP (Julius and Yuzenkova 2017).

The efficiency of NAD^+ capping also depends on the identity of the nucleotides −3 and −2 upstream and + 2, +3, and + 4 downstream of the TSS. In particular, the +2 base has a large, 6–8-fold effect on the efficiency of noncanonical capping, which

makes it the second strongest determinant of capping with a noncanonical initiating nucleotide. Through the CapZyme-seq method, a consensus promoter sequence for the highest efficiency of NAD^+ capping was determined as HRRASWW (H, ATC; R, GA; S, GC; W, AT), where A is the +1 base in *E. coli*. Replacing the bases with their anti-consensus sequence, GYYAWSS (Y, TC), leads to a 40-fold decrease in NAD^+ capping efficiency (Vvedenskaya et al. 2018). Differing from *E. coli*, in yeast, the highly conserved promoter motif, YAAG, is associated with efficient NAD^+ incorporation and is more likely to be recognized by the yeast RNAP II (Zhang et al. 2020).

Comparable to NAD^+, capping by Np_4A in *E. coli* also depends on the identity of the base pair at position −1. The levels of capping are higher when the −1 base on the coding strand is a purine rather than a pyrimidine, whereas the −2 and − 3 positions only modestly affect Np_4A incorporation (Luciano and Belasco 2020). Taken together, it appears that the promoter sequence strongly affects the incorporation efficiency of noncanonical caps.

3.4 Cellular Metabolite Concentration

The intracellular concentration of NTPs and other noncanonical substrates utilized by RNAP for transcription plays a central role in regulating noncanonical cap initiation. Higher concentrations of NTPs lead to a greater chance of penetration into the active site of RNAP to initiate transcription (Haugen et al. 2008). RNAPs therefore seem to serve as both sensors to and actuators for the level of cellular metabolites, adjusting the transcriptional yield accordingly (Bird et al. 2018). For example, when high mitochondrial NAD(H) levels were changed to low levels, the levels of NAD(H)-capped mitochondrial RNAs changed from 15% to 0% (Bird et al. 2018).

Further support for the notion that cellular metabolite concentration influences the incorporation of noncanonical caps has been demonstrated in bacteria. In *E. coli*, the average cellular ATP concentration is about 1.54 mM, and the cellular NAD^+ concentration is about 0.6 mM, while the NADH concentration is up to 10 times lower than that of NAD^+. This predicts the probability of incorporation of each nucleotide, which would have an order of ATP > NAD^+ > NADH (Lin and Guarente 2003; Zhou et al. 2011; Yaginuma et al. 2014). The concentrations of other noncanonical substrates, such as dpCoA and FAD, are only around 10 μM to 600 μM, lower than that of NAD^+ (Takamura and Nomura 1988; Louie et al. 2003). These cellular concentrations are consistent with that of the respective, noncanonically capped RNA transcript levels detected in vivo (Chen et al. 2009; Kowtoniuk et al. 2009; Wang et al. 2019a). As for dinucleoside polyphosphates (Np_4N) in *E. coli*, concentrations are even lower than those of FAD, but this concentration elevates during oxidative stress. Thus, *E. coli* mRNA and sRNA can only acquire Np_4N caps under disulfide stress conditions that increase Np_4N cellular concentrations (Luciano et al. 2019; Luciano and Belasco 2020). Only one

substrate's concentration, UDP-GlcNAc, is comparable with that of NAD^+. UDP-GlcNAc is the most abundant noncanonical cap in vivo, consistent with the relative level of the cellular metabolite (>1 mM) in *E. coli* and human cells (Mao et al. 2006; Namboori and Graham 2008; Wang et al. 2019a). These studies indicate that the cellular concentration of noncanonical substrates is an important factor for transcript capping. Conversely, negative regulation by high NTP levels also leads to nascent transcription abortion (Turnbough and Switzer 2008). However, this negative regulation has not yet been reported during noncanonical capping.

3.5 Post-Transcription

Based on the limited research available, biosynthesis of noncanonically capped RNA by RNA polymerases during transcription initiation is the most common route. However, other mechanisms could also occur. For example, in mammalian cells, snoRNAs and small Cajal body RNAs (scaRNAs), some of which are produced from introns via splicing, also contain NAD^+ caps, particularly after the removal of the decapping enzyme, DXO, in cells. This observation led to the proposal that an alternate, post-transcriptional NAD^+ capping mechanism exists (Jiao et al. 2017).

There are capping mechanisms independent of those of RNAPs. In *E. coli*, some aminoacyl-tRNA synthetases, such as LysU, enable catalysis during the reaction of aminoacyl-adenylates with not only the 5′ triphosphate of mononucleotides but also with the triphosphorylated 5′ end of polynucleotides. This reaction produces Ap_4A capped *yeiP* RNA (Luciano et al. 2019). Additionally, in vitro, ribozymes that are able to incorporate NAD^+, FAD, and dpCoA into the 5′ terminal of RNA may also represent a potential method of capping in the in vivo synthesis of capped RNA (Huang 2003). Furthermore, m^7G-capped RNAs can undergo m^7G cap removal under specific conditions, and re-capping by NAD^+ may also be possible (Zhang et al. 2019a). Collectively, these studies suggest that alternative post-transcriptional noncanonical capping mechanisms may exist and need to be looked into.

In summary, RNA polymerases, initiation σ factors, Rif pockets, promoter sequences, and cellular metabolite concentration all influence the profile of NAD^+-capped RNA in organisms. However, the steady-state level of noncanonically capped RNAs may not only depend on such determinants. This level is also dynamically regulated by decapping mechanisms.

4 Decapping Enzymes of Noncanonical RNA Caps

Equally important to understanding the mechanisms involved in the modification of the 5′ end of RNA is understanding how noncanonical caps may be removed in a process referred to as decapping. While research in the decapping of noncanonical caps is only recently budding, decapping of the canonical eukaryotic m^7G structure

and conversion of the bacterial triphosphate 5′ end to a monophosphate have been studied extensively over the past several decades.

In eukaryotes, decapping of the m^7G structure is tied to the regulation of gene expression and is recognized to play a role in mRNA turnover. The degradation of mRNA can be accomplished through various mechanisms, where decapping is a critical step for 5′-to-3′ decay in particular. This decapping occurs in a deadenylation-dependent or deadenylation-independent manner. During deadenylation-dependent decay, deadenylases and associated proteins encourage decapping after poly(A) tail shortening, whereas during deadenylation-independent decay, decapping is triggered through mechanisms such as mRNA uridylation or endonucleolytic cleavage. With successful decapping, an exoribonuclease, such as the mammalian XRN1 or the plant XRN4, degrades RNA containing a 5′ end monophosphate (Łabno et al. 2016).

However, bacterial mRNA degradation occurs through different pathways, as bacterial RNAs do not contain the m^7G cap. Bacterial RNA largely contains a triphosphate at the 5′ end and a stabilizing hairpin structure at the 3′ end. Broadly, in bacteria, RNA degradation occurs through two pathways referred to as "direct access" and "5′-end-dependant" degradation (Hui et al. 2014; Kramer and McLennan 2019). Direct access degradation begins with cleavage by an endonuclease, such as RNase E in *E. coli*, and subsequently proceeds through 3′-to-5′ or 5′-to-3′ decay by exonucleases. On the other hand, 5′-end-dependant degradation initiates through the hydrolysis of the triphosphorylated 5′ end to a monophosphate by an enzyme, such as RppH, which makes the RNA susceptible to endonucleases and exonucleases, such as RNase E and RNase J, respectively.

As for noncanonical caps, the decapping process has been less extensively examined. Enzymes that are responsible for this decapping largely fall into two protein families: Nudix and DXO (Fig. 3). These two families are also involved in the hydrolysis of the canonical eukaryotic m^7G cap and the bacterial triphosphate 5′ end. Similar to canonical decappers, decapping proteins for noncanonical caps generally encourage the conversion of the 5′ end to a monophosphate, which subjects the RNA to further degradation.

4.1 Nudix Enzymes Involved in Decapping of Noncanonical Caps

The Nudix superfamily consists mainly of pyrophosphohydrolases that were initially classified for demonstrating activity on various nucleoside diphosphates linked to moiety X, although this family also includes proteins of other functionalities (Srouji et al. 2017). Nudix proteins are ancient, widespread, and evolutionarily conserved between all three branches of life, as well as viruses, with 13, 7, 22, and 28 Nudix genes found in *E. coli*, *Saccharomyces cerevisiae*, humans, and *Arabidopsis thaliana*, respectively (McLennan 2006; Yoshimura and Shigeoka 2015;

Carreras-Puigvert et al. 2017). Many proteins in this family can be defined by a conserved region termed the Nudix motif, $GX_5EX_7REUXEEXGU$, where U is a hydrophobic amino acid and X is any amino acid. This motif is critical for catalytic activity and the binding of divalent cations like Mg^{2+} and Mn^{2+}, which function as cofactors for the pyrophosphohydrolase activity (Mildvan et al. 2005; McLennan 2006). In particular, Nudix hydrolases have diverse functions and substrates, and were originally described as "housecleaning enzymes" that act to rid cells of toxic materials and reduce the accumulation of metabolites and intermediates (McLennan 2006). However, Nudix hydrolases have also recently been demonstrated to be proficient decappers of noncanonical RNA caps (Fig. 3).

4.1.1 Canonical Decappers: Dcp2, Nudt16, and Nudt3

The function of Nudix hydrolases in decapping has been recognized since the identification of the decapping abilities of Dcp2 (Wang et al. 2002). Dcp2 is conserved in eukaryotes and functions in the hydrolysis of the canonical m^7G cap. However, Dcp2 isn't the only enzyme responsible for the decapping of the m^7G structure. In vitro studies suggest that several other Nudix enzymes, such as human Nudt16 and Nudt3, could be involved in m^7G decapping (Song et al. 2013). Of the various Nudix enzymes that demonstrate m^7G decapping activity, only Nudt16 has recently shown potential in the hydrolysis of NAD^+-, FAD-, and dpCoA-capped RNA in vitro and (for NAD^+-capped RNA and FAD-capped RNA) in cells (Sharma et al. 2020).

4.1.2 NudC

NudC was the first member of the Nudix superfamily recognized to have activity on noncanonically capped RNA. In *E.coli*, NudC can hydrolyze NAD^+-capped RNA in the presence of Mg^{2+} to release nicotinamide mononucleotide (NMN) and monophosphorylated RNA, which is susceptible to further degradation by RNase E (Cahová et al. 2015; Bird et al. 2016; Kiledjian 2018). Following NudC deletion, NAD^+-capped RNA levels rise, supporting that NudC functions in cells as a regulator of NAD^+-capped RNA. Additionally, NudC can hydrolyze NAD^+ and NADH at lower efficiency compared to NAD^+-capped RNA but displays no significant activity against 5′ triphosphorylated RNA, indicating it may primarily serve to remove NAD^+ caps (Cahová et al. 2015; Höfer et al. 2016; Abele et al. 2020). In vitro, NudC also exhibits activity on RNA capped with NADH and dpCoA (Bird et al. 2016).

NudC prefers single-stranded substrates with three or more unpaired bases at the 5′ end and a purine as the first base of the RNA. In terms of RNA lengths, NudC can hydrolyze the NAD^+ cap of both longer, complex RNA and short RNA (Höfer et al. 2016). Structurally, NudC functions as a symmetric homodimer, where both monomers bind to an individual NAD^+. This dimerization is essential for substrate

recognition and binding, as the catalytic pocket containing the Nudix motif is comprised of residues from each monomer (Höfer et al. 2016; Zhang et al. 2016).

Recently, several close NudC homologs have been characterized. In mammals, Nudt12 was demonstrated to hydrolyze cytosolic NAD^+-capped RNA (Grudzien-Nogalska et al. 2019; Wu et al. 2019). Like NudC, loss of Nudt12 increases the levels of NAD^+-capped RNA, indicating that Nudt12 regulates the stability of a subset of NAD^+-capped RNA in cells. Specifically, Nudt12 may regulate transcripts involved in metabolism, as NAD^+-capped transcripts that increased after nutrient stress were responsive to Nudt12 decapping and included nuclear-encoded mitochondrial protein mRNAs. Nudt12 may also have a role in the regulation of circadian clock transcripts. Structurally similar to NudC, Nudt12 functions as a homodimer, with most of the structural differences occurring in the N-terminal domain instead of the C-terminal domain, which contains the conserved Nudix motif (Grudzien-Nogalska et al. 2019; Wu et al. 2019). Nudt12 interacts with bleomycin hydrolase (BLMH), forming a dodecamer that likely contains a BLMH hexamer and three Nudt12 dimers. The interaction between Nudt12 and BLMH is necessary to localize Nudt12 to cytoplasmic granules that are distinct from P-bodies. This sequestration of Nudt12 to cytoplasmic granules may be beneficial to regulate Nudt12 activity on m^7G and unmethylated caps, since Nudt12 can hydrolyze these structures to release m^7GMP and GMP/GDP, respectively. Finally, Nudt12 shows activity on NAD^+ and NADH, but prefers NADH (Grudzien-Nogalska et al. 2019).

Other close NudC homologs have been identified. Recently, in yeast, Npy1 was demonstrated to hydrolyze NAD^+-capped RNA in the cytosol (Zhang et al. 2020). Additionally, in vitro, Nudt19 in *Oryza sativa* contained NAD^+-decapping capabilities (Zhang et al. 2016).

4.1.3 RppH

A second bacterial protein involved in decapping is RppH (Deana et al. 2008). RppH is an RNA pyrophosphohydrolase with two differing prototypes, one from *E. coli* (EcRppH) and the other from *B. subtilis* (BsRppH). Orthologs of the EcRppH prototype are found within many classes of proteobacteria and flowering plants, while those of BsRppH are mainly restricted to the order Bacillales (Foley et al. 2015; Bischler et al. 2016). Both prototypes are involved in the hydrolysis of the 5′ triphosphate present in bacterial RNA, but key sequence and structural differences result in unique substrate specificity and function. EcRppH and BsRppH share only 23% identity, with much of the sequences outside of the Nudix motif differing significantly (Richards et al. 2011; Foley et al. 2015), leading to crucial differences between the two.

Recently, both BsRppH and EcRppH have been implicated in the removal of noncanonical caps. It has been demonstrated in vitro that BsRppH can decap NAD^+-capped RNA, resulting in monophosphorylated RNA and NMN (Frindert et al. 2018). This removal of the NAD^+ cap is enhanced by Mn^{2+} ions and the presence of guanosine at the second base position, but is inhibited by double-stranded

structures present at the 5′ end. However, loss of BsRppH did not significantly affect NAD^+-capped RNA levels, suggesting that NAD^+ cap removal may not be the primary function of this enzyme in vivo (Frindert et al. 2018). Similarly, EcRppH has been shown to decap NAD^+-capped RNA in vitro in some studies (Frindert et al. 2018; Grudzien-Nogalska et al. 2019), although this finding remains debatable due to contrasting studies that demonstrate that EcRppH has little efficiency on NAD^+-capped RNA (Cahová et al. 2015; Bird et al. 2016; Abele et al. 2020). Widely, it is instead theorized that the NAD^+ cap may serve to protect transcripts from EcRppH-dependent degradation (Cahová et al. 2015; Bird et al. 2016; Abele et al. 2020). Other than NAD^+-capped transcripts, EcRppH has been demonstrated to hydrolyze Np_nN caps to a 5′ monophosphate, although methylation of the Np_nN cap structure can inhibit this activity (Luciano et al. 2019; Hudeček et al. 2020).

4.2 DXO Enzymes Involved in Noncanonical Decapping

A second family of proteins that is recognized for having activity on a variety of RNA caps is the DXO family of proteins (Fig. 3). This protein family shares an active site with six conserved motifs, which function in cleavage, RNA binding, and the coordination of divalent cations. Outside of this active site, there is little conservation between proteins in this family (Xiang et al. 2009; Chang et al. 2012; Wang et al. 2015). An important difference between the DXO family of proteins and the Nudix superfamily of proteins is that the two cleave noncanonical caps at different locations, with the DXO family removing the entire cap structure (Fig. 3). In this section, three prototypes in the DXO family will be discussed: Rai1, Dxo1, and DXO.

4.2.1 Rai1

The fungal Rai1 is present in the nucleus and was initially found to be a pyrophosphohydrolase with activity on 5′ triphosphorylated RNA, releasing diphosphate and RNA with a monophosphorylated 5′ end (Xiang et al. 2009). Association of Rai1 with the 5′-to-3′ exoribonuclease Rat1 affords degradation of the remaining 5′ monophosphorylated RNA product and stimulates both cleavage by Rai1 and 5′ to 3′ exonuclease activity by Rat1 (Xiang et al. 2009; Jiao et al. 2010). Subsequent to the revelation that Rai1 functions on 5′ triphosphorylated RNA, it was demonstrated that Rai1 could remove the canonical m^7G cap, but was most efficient in removing unmethylated caps, releasing the entire cap structure, GpppN (Jiao et al. 2010). In addition, Rai1 homologs can have triphosphonucleotide hydrolase activity, releasing pppN (Wang et al. 2015). These functions indicated that the primary role of Rai1 was surveillance against aberrantly capped RNA. Recently, this role has expanded to include decapping of noncanonical caps. Rai1 cleaves NAD^+-capped RNA to release NAD^+ and also has activity on RNA capped with dpCoA and FAD in vitro (Jiao

et al. 2017; Vvedenskaya et al. 2018; Doamekpor et al. 2020a). Finally, the complex formed by Rai1 and Rat1 can also degrade 5' OH RNA (Doamekpor et al. 2020b).

4.2.2 Dxo1

Dxo1 works together with Rai1 in some yeast species to monitor aberrantly capped RNA, but is present in both the cytoplasm and the nucleus, indicating that there may be a hierarchical order to this surveillance (Zhang et al. 2020). Unlike Rai1, Dxo1 displays no pyrophosphohydrolase activity on 5′ triphosphorylated RNA; however, Dxo1 is highly efficient at removal of unmethylated GpppN cap structures (Chang et al. 2012). Additionally, this protein can remove the canonical m^7G cap more efficiently than Rai1, although it prefers unmethylated caps (Chang et al. 2012). Unlike Rai1, which generally depends on Rat1 for exonuclease activity, Dxo1 contains 5′ to 3′ exonuclease activities of its own, though it is prone to stalling at secondary structures (Chang et al. 2012). Finally, Dxo1 has activity on NAD^+-, dpCoA-, and FAD-capped RNA in vitro (Jiao et al. 2017; Doamekpor et al. 2020a).

4.2.3 Mammalian DXO

The predominately nuclear mammalian homolog, DXO, is a pyrophosphohydrolase, RNA-specific 5′ to 3′ exonuclease, and decapper of canonical and noncanonical RNA caps. DXO may have a preference for activity on pre-mRNA (Jiao et al. 2013) and can release diphosphate from 5′ triphosphorylated RNA, GpppG from RNA with unmethylated caps, and NAD^+, dpCoA, and FAD from RNA with noncanonical caps (Jiao et al. 2013, 2017; Doamekpor et al. 2020a). Additionally, DXO is efficient against methylated caps (Jiao et al. 2013) and can remove a 5′-OH dinucleotide before degrading 5′ OH RNA, making it a hydroxyl dinucleotide hydrolase (Doamekpor et al. 2020b). In cells, NAD^+-capped RNA and FAD-capped RNA levels rise when DXO activity is absent (Jiao et al. 2017; Doamekpor et al. 2020a). DXO likely functions on distinct subsets of these RNAs with a potential tie to RNA involved in environmental stress, such as heat shock (Grudzien-Nogalska et al. 2019). Due to the high activity of DXO on a variety of cap structures, this protein must be highly regulated. For example, cap binding proteins such as CBP20 and eIF4E can inhibit DXO activity, effectively protecting properly capped RNA (Jiao et al. 2013). The 2′-O-methylated cap structure also protects RNA from degradation by DXO (Picard-Jean et al. 2018).

4.2.4 Plant DXO1

In plants, the only DXO homolog present is the nuclear and cytoplasmic DXO1, which was also demonstrated to have deNADding, exoribonuclease, and hydroxyl dinucleotide hydrolase activity (Kwasnik et al. 2019; Doamekpor et al. 2020b; Pan

et al. 2020). However, this protein does contain a plant-specific modification of the active site that hampers its 5′ to 3′ exonuclease activity and its activity on other 5′ RNA modifications. Despite this modification and independent of its role as a potent deNADding enzyme, *Arabidopsis* DXO1 has likely evolved to have a role in chloroplast-, development-, and immunity-related processes. For example, the N-terminal extension of the protein may promote chloroplast functions, potentially serving as a connection between nuclear and plastid signaling (Kwasnik et al. 2019; Pan et al. 2020).

4.3 Other Enzymes Involved in the Decapping of Noncanonical RNA Caps

Other enzymes outside of these two protein families may also be capable of decapping noncanonical RNA caps. For example, a bis (5′-nucleosyl)-tetraphosphatase (ApaH) was demonstrated to be able to efficiently remove Np_nN caps (Hudeček et al. 2020). Additionally, CD38, a human glycohydrolase, can process NAD^+-capped RNA in vitro (Abele et al. 2020). Diverse other enzymes could be involved in the decapping of RNA with noncanonical caps, and further research is required to delve into these possibilities.

5 Potential Molecular and Biological Functions of Noncanonical Capping

5.1 Does Noncanonical Capping Promote RNA Stability or Decay?

The 5′ terminal structure can affect the stability of an RNA. In *E. coli*, 5′ triphosphate RNA generally has a longer half-life than 5′ monophosphate RNA, while in eukaryotes, the m^7G cap plays a central role in mRNA stability. However, whether noncanonical RNA caps also regulate mRNA stability remains somewhat controversial (Fig. 4). In *E. coli*, the 5′ end of triphosphorylated RNA can be hydrolyzed by the Nudix protein, RppH, to yield a 5′ monophosphorylated RNA, thereby triggering RNase-E-mediated decay (Deana et al. 2008). In vitro experiments showed that 5′ end modification with NAD^+ strongly decelerates processing by RppH, thus heightening stability against RNase E (Cahová et al. 2015). NAD^+ capping also resulted in a three to fourfold increase in RNA stability in vivo (Bird et al. 2016). Similarly, NAD^+ capping in *B. subtilis* stabilized mRNA against exonucleolytic decay by RNase J1, which prefers degrading 5′ monophosphorylated RNA (Frindert et al. 2018).

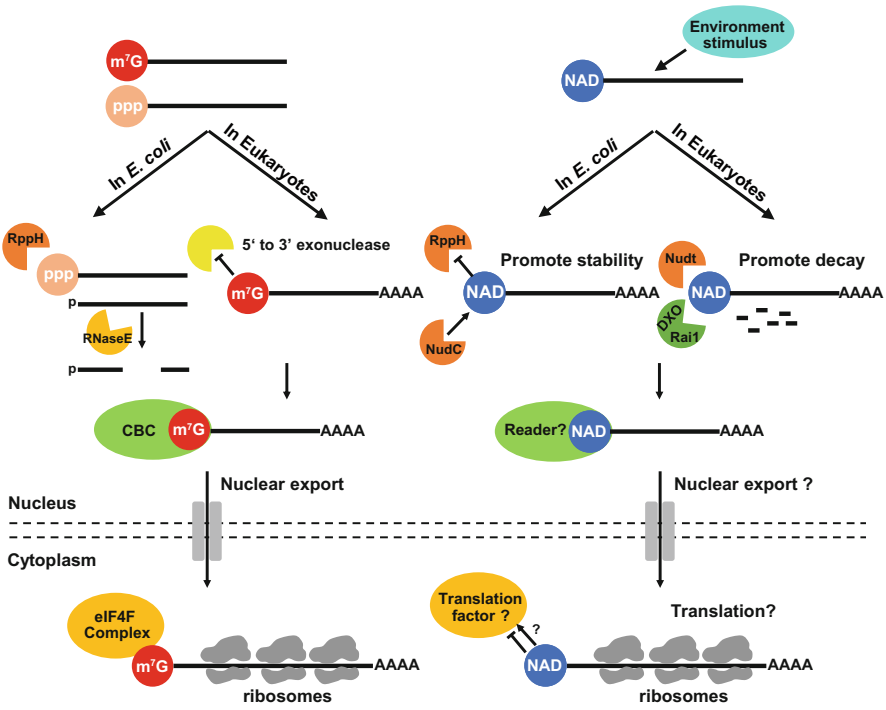


Fig. 4 Model of the potential molecular functions of the NAD^+ cap in RNA in *E. coli* and eukaryotes. NAD^+-capped RNA is altered dynamically in vivo through regulation by capping and decapping enzymes. In *E. coli*, pppRNA undergoes 5′ to 3′ decay enabled by RppH pyrophosphohydrolase and RNase E endonuclease activity, while the NAD^+ cap promotes RNA stability against RppH and RNase E. In eukaryotes, the m^7G cap protects mRNA against decay, while the NAD^+ cap promotes 5′ to 3′ decay through recruitment of deNADding enzymes. CBC (cap binding complex) binds to the m^7G cap to mediate splicing, polyadenylation, and nuclear export, though these steps remain unclear for the NAD^+ cap. m^7G-capped RNA recruits the eIF4F complex to initiate translation, while NAD^+-capped RNA does exist on plant ribosomes but does not support translation in vitro or in transfected human cells. NAD^+-capped RNA can be regulated by environmental stimuli and growth conditions, but the exact molecular and biological functions need to be further investigated

Contrary to observations in bacteria, the 5′ NAD^+ cap promoted decay of RNAs in eukaryotes. In human cells, transfected NAD^+-capped and polyadenylated luciferase mRNA was less stable and decayed via deNADding followed by 5′-3′ decay by DXO. The observed opposite response to NAD^+ capping of prokaryotic and eukaryotic cells is perhaps due to the different features between these organisms, as well as differences in experimental methods. In *E. coli*, the main machinery for RNA degradation is a complex of the endoribonuclease, RNase E, and an exoribonuclease. Therefore, the inhibition of this complex by NAD^+ capping could stabilize the RNA transcripts. However, in eukaryotes, most of the RNA transcripts are under the protection of the m^7G cap at the 5′ end, while less than 10% of RNA transcripts

are NAD^+-capped. Thus, the NAD^+ cap is more likely to be a 5′ end mark to recruit DXO and mediate decay of RNA that is unneeded or non-functional, save for special conditions (Jiao et al. 2017). Knockdown of mDXO or AtDXO in human or plant cells, respectively, causes the enrichment of NAD^+-capped RNA (Jiao et al. 2017; Pan et al. 2020). In particular, most enriched NAD^+ RNA in human cells are sno/scaRNAs, which are highly resistant to exonucleolytic degradation. This suggests that NAD^+ capping for sno/scaRNAs probably triggers DXO-mediated decay (Filipowicz and Pogacić 2002; Jiao et al. 2017).

In addition, noncanonical capping may mediate RNA stability indirectly via a 5′-independent mechanism. For example, most of the NAD^+-capped RNA revealed by NAD^+ captureSeq in *E. coli* are short fragments, which might imply that RNA degradation also occurs without removing the 5′ cap (Cahová et al. 2015). A direct entry and attack mechanism by RNase E might not need a 5′ monophosphate end and perhaps could be induced by noncanonical capping (Bouvier and Carpousis 2011). In plants, sRNA biogenesis is an alternative way to degrade NAD^+-capped RNA when there is a loss of the decapping enzyme DXO (Pan et al. 2020). Conversely, noncanonical capping might promote RNA stability by blocking the polyadenylation process that initiates degradation in *E. coli*, as its poly(A) polymerase prefers monophosphorylated substrates (Kushner 2004). However, how RNA stability is altered by noncanonical capping remains largely undefined and still requires further experimental support.

5.2 *Is Noncanonical Capping of RNA Involved in Translation Regulation?*

The initiation step of translation is critical to protein production. It requires the delivery of the ribosomal subunit to an mRNA, usually at the 5′ end. In eukaryotes, translation initiation is primarily achieved by the 5′ mRNA m^7G cap through binding with the eIF4F complex, which recruits the ribosomal subunit pre-bound to a complex of initiation factors (Mitchell and Parker 2015). Caps other than m^7G may not be recognized by this translation complex (Issur et al. 2013). Therefore, whether noncanonically capped RNA possesses the ability to be translated remains uncertain (Fig. 4).

In vitro translation experiments for yeast nuclear NAD^+-capped transcripts suggest that NAD^+-capped RNAs are unable to be translated, producing even less protein than triphosphorylated RNA and monophosphorylated RNA (Zhang et al. 2020). NAD^+-capped and polyadenylated luciferase mRNA transfected into human cells displayed a translation signal no greater than that for uncapped RNA, similarly suggesting that NAD^+-capped RNA is unable to initiate translation (Jiao et al. 2017). However, this study was performed using artificial, exogenous NAD^+-capped RNA, which may not reflect the natural conditions in vivo. An alternate study in plants demonstrated that NAD^+-capped mRNAs are enriched in the polysome fraction with

translating ribosomes and therefore can probably be translated (Wang et al. 2019b). So far, there are no studies that report NAD^+ capping mediating translation initiation or observations of the translation initiation complex binding with NAD^+-capped RNA.

In eukaryotes, there exist other translation mechanisms that are independent of the 5′ end cap. Some mRNAs contain specific internal ribosome entry sites (IRES) to recruit ribosomal subunits, and m^6A modification in the 5′ UTR can promote the translation of a transcript (Mitchell and Parker 2015). It is possible that NAD^+-capped RNAs enriched in the polysome fraction might undergo translation through a cap-independent mechanism involving internal ribosome entry. Alternatively, additional modifications could promote the translation of NAD^+-capped RNA. For example, the presence of m^6Am modification on the second nucleotide next to the m^7G cap increases translation initiation (Meyer et al. 2015). It is unclear if NAD^+-capped 5′ ends contain these m^6Am modified nucleotides. A recent report shows that the m^6Am next to the m^7G cap can be specifically demethylated by fat mass and obesity-associated protein (FTO), whose activity is enhanced by binding with NADP (Mauer et al. 2017; Wang et al. 2020). Thus, the possibility exists that NAD^+ capping can recruit a protein factor to promote or inhibit translation initiation. In pathogens, NAD^+ capping in RNAIII impairs the translation of its target gene, hla. This is perhaps due to the pseudo-base pairing between the nicotinamide of NAD^+ and the target RNA (Morales-Filloy et al. 2020).

5.3 The Relationship Between Noncanonical Capping and Cellular Metabolism

NAD^+, which is one of the most common organic cofactors, plays a critical role in cellular metabolism. Genes involved in the NAD^+-NADP synthesis pathway, or encoding NAD^+-NADP utilizing enzymes, were observed to produce NAD^+-capped RNAs in different organisms (Morales-Filloy et al. 2020). For example, L-threonine 3 dehydrogenase (tdh) catalyzes an NAD^+-dependent oxidation reaction in *B. subtilis*. NAD^+-capped tdh mRNA may directly provide a regulatory feedback mechanism for the synthesis of this protein (Frindert et al. 2018). Another gene involved in NAD^+ synthesis, nadA, is usually regulated by the nadA motif in the 5′ UTR that binds ligands, and might also be modulated by NAD^+ RNA capping (Malkowski et al. 2019). These findings imply that NAD^+ RNA capping may substitute for direct feedback regulation by the cofactor NAD^+ to regulate NAD^+ synthesis.

5.4 *Regulation of Noncanonical Capping by Developmental and Environmental Stimuli*

Cellular NAD^+ plays a vital role in metabolism and acts as a factor linking cellular metabolism, transcript level, and environmental stimulus (Gakière et al. 2018). Perhaps due to the roles of NAD^+ in the cell, NAD^+-capped RNA is affected by developmental stage and environmental condition. For instance, NAD^+-capped transcripts in the stationary phase of *E. coli* are twofold higher than in the exponential phase (Bird et al. 2016), and yeast cultures in synthetic media result in more NAD^+-capped transcripts compared to those in rich media (Frindert et al. 2018). These results demonstrate that NAD^+ capping could be modulated in response to environmental changes.

Additionally, it was found that NAD^+-capped RNA significantly increased when human cells were exposed to either heat shock or glucose deprivation, while cellular NAD^+ levels did not consistently demonstrate the same response. This suggests that NAD^+ capping can be directly modulated under stress and isn't only altered though sensing the cellular NAD^+ level by RNAP. Moreover, the target NAD^+-capped transcripts of DXO or Nudt12 were altered, further indicating that distinct regulation of NAD^+-capped RNA is undertaken following different stresses (Grudzien-Nogalska et al. 2019). Likewise, Np_4N in bacteria is thought to act as alarmones through receptor mediated signaling in environmental stress response. However, the generation of Np_4A-capped RNAs under disulfide stress implies that the physiological responses previously attributed to Np_4A signaling might be due to an Np_4A RNA capping mechanism (Luciano et al. 2019).

6 Conclusion and Outlook

For a long time, the hallmark for mRNA capping in eukaryotes was the traditional m^7G cap. After about 50 years of research, the molecular and biological function of m^7G RNA capping in different organisms has been uncovered. In recent years, the new discovery of the NAD^+ cap on mRNA opened a novel and exciting research field for RNA biology. With the present detection strategies, NAD^+-capped RNAs appear widespread in various prokaryotes and eukaryotes. NAD^+ capping occurs mainly on mRNA but also on noncoding RNAs. Additionally, NAD^+-capped transcripts encode proteins involved in a range of biological processes, particularly cellular metabolism and stress responses.

The mechanism of incorporation of noncanonical caps like NAD^+ continues to be elucidated. At present, NAD^+ is only known to be introduced into the 5′ end of RNA by RNA polymerases during transcription initiation. However, many questions remain. Do RNA polymerases deposit NAD^+ differently at unique genes? At one gene, are the transcription initiation sites different when NAD^+ vs. ATP is used as the initiating nucleotide? Besides this capping mechanism, are there alternative

mechanisms of NAD^+ capping? After capping by RNA polymerases, how are NAD^+-capped transcripts exported out of the nucleus? Are there any "readers" that recognize such transcripts? Finally, can these transcripts be translated by m^7G-cap-independent mechanisms? A plethora of questions remain unanswered concerning the mechanisms surrounding RNA with noncanonical caps.

Decapping enzymes are involved in maintaining the steady-state levels of noncanonically capped transcripts in vivo. The Nudix and DXO families of proteins, which have long been known as hydrolases for various cellular metabolites (Ogawa et al. 2008), have been demonstrated to possess potent decapping activities that target different noncanonically capped transcripts. How decapping enzymes specifically regulate noncanonical capping and perform uncharacterized biological functions highlights a major bottleneck to obtaining a full understanding of RNA capping by NAD^+ and other metabolites.

Besides capping by NAD^+, other noncanonical substrates (FAD, dpCoA, UDP-Glc, UDP-GlcNAc, Np_nN) have been identified in RNA in some organisms. Unfortunately, we still await robust sequencing technologies for such noncanonically capped RNA, which will pave the way to understanding their profiles in various transcriptomes. However, so far, no phenotypic changes were observed upon increasing or decreasing these noncanonically capped RNA in vivo. This brings up another critical inquiry: What are the functions of RNA noncanonical capping? How did noncanonical capping come to exist in evolution? Was it an accidental or specific event? We trust that further work on RNA with noncanonical caps will shed light on various questions in the epitranscriptomics field and will afford more practical applications.

References

Abele F, Höfer K, Bernhard P et al (2020) A novel NAD-RNA decapping pathway discovered by synthetic light-up NAD-RNAs. Biomol Ther 10:513

Barvík I, Rejman D, Panova N et al (2017) Non-canonical transcription initiation: the expanding universe of transcription initiating substrates. FEMS Microbiol Rev 2:131–138

Bird JG, Basu U, Kuster D et al (2018) Highly efficient 5′ capping of mitochondrial RNA with NAD^+ and NADH by yeast and human mitochondrial RNA polymerase. elife 7:e42179

Bird JG, Zhang Y, Tian Y et al (2016) The mechanism of RNA 5′ capping with NAD+, NADH and desphospho-CoA. Nature 535:444–447

Bischler T, Hsieh P, Resch M et al (2016) Identification of the RNA pyrophosphohydrolase RppH of helicobacter pylori and global analysis of its RNA targets. J Biol Chem 292:1934–1950

Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307

Boo SH, Kim YK (2020) The emerging role of RNA modifications in the regulation of mRNA stability. Exp Mol Med 52:400–408

Bouvier M, Carpousis AJ (2011) A tale of two mRNA degradation pathways mediated by RNase E. Mol Microbiol 82:1305–1310

Breaker RR (2012) Riboswitches and the RNA world. Cold Spring Harb Perspect Biol 4:a003566

Cahová H, Winz ML, Höfer K (2015) NAD capture Seq indicates NAD as a bacterial cap for a subset of regulatory RNAs. Nature 519:374–377

Campbell EA, Korzheva N, Mustaev A et al (2001) Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. Cell 104:901–912
Carreras-Puigvert J, Zitnik M, Jemth AS et al (2017) A comprehensive structural, biochemical and biological profiling of the human NUDIX hydrolase family. Nat Commun 8:1541
Chang J, Jiao X, Chiba K et al (2012) Dxo1 is a novel eukaryotic enzyme with both decapping and 5′-3′ exoribonuclease activity. Nat Struct Mol Biol 19:1011–1017
Chen YG, Kowtoniuk WE, Agarwal I et al (2009) LC/MS analysis of cellular RNA reveals NAD-linked RNA. Nat Chem Biol 5:879–881
Chen K, Zhao BS, He C (2016) Nucleic acid modifications in regulation of gene expression. Cell Chem Biol 23:74–85
Chen LQ, Zhao WS, Luo GZ (2020) Mapping and editing of nucleic acid modifications. Comput Struct Biotechnol J 18:661–667
Cohn WE (1960) Pseudouridine, a carbon-carbon linked ribonucleoside in ribonucleic acids: isolation, structure, and chemical characteristics. J Biol Chem 235:1488–1498
David R, Burgess A, Parker B et al (2017) Transcriptome-wide mapping of RNA 5-methylcytosine in Arabidopsis mRNAs and noncoding RNAs. Plant Cell 29:445–460
Deana A, Celesnik H, Belasco JG (2008) The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. Nature 451:355–358
Doamekpor SK, Gozdek A, Kwasnik A et al (2020b) A novel 5′-hydroxyl dinucleotide hydrolase activity for the DXO/Rai1 family of enzymes. Nucleic Acids Res 48:349–358
Doamekpor SK, Grudzien-Nogalska E, Mlynarska-Cieslak A et al (2020a) DXO/Rai1 enzymes remove 5′-end FAD and dephospho-CoA caps on RNAs. Nucleic Acids Res 48:6136–6148
Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485:201–206
Durniak KJ, Bailey S, Steitz TA (2008) The structure of a transcribing T7 RNA polymerase in transition from initiation to elongation. Science 322:553–557
Filipowicz W, Pogacić V (2002) Biogenesis of small nucleolar ribonucleoproteins. Curr Opin Cell Biol 14:319–327
Foley PL, Hsieh PK, Luciano DJ et al (2015) Specificity and evolutionary conservation of the Escherichia coli RNA pyrophosphohydrolase RppH. J Biol Chem 290:9478–9486
Frindert J, Zhang Y, Nübel G et al (2018) Identification, biosynthesis, and decapping of NAD-capped RNAs in B. subtilis. Cell Rep 24:1890–1901.e8
Fu Y, He C (2012) Nucleic acid modifications with epigenetic significance. Curr Opin Chem Biol 16:516–524
Gakière B, de Bont L, Fernie AR et al (2018) NAD^+ biosynthesis and signaling in plants. Crit Rev Plant Sci 37:259–307
Galloway A, Cowling VH (2019) mRNA cap regulation in mammalian cell function and fate. Biochim Biophys Acta Gene Regul Mech 1862:270–279
Göpel Y, Lüttmann D, Heroven AK et al (2011) Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in Enterobacteriaceae. Nucleic Acids Res 39:1294–1309
Gray MW, Lang BF (1998) Transcription in chloroplasts and mitochondria: a tale of two polymerases. Trends Microbiol 6:1–3
Grosjean H, Droogmans L, Roovers M et al (2007) Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. Methods Enzymol 425:55–101
Grudzien-Nogalska E, Bird JG, Nickels BE et al (2018) "NAD-capQ" detection and quantitation of NAD caps. RNA 24:1418–1425
Grudzien-Nogalska E, Wu Y, Jiao X et al (2019) Structural and mechanistic basis of mammalian Nudt12 RNA deNADding. Nat Chem Biol 15:575–582
Haugen SP, Ross W, Gourse RL (2008) Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Microbiol 6:507–519
Helm M, Motorin Y (2017) Detecting RNA modifications in the epitranscriptome: predict and validate. Nat Rev Genet 18:275–291

Hillen HS, Temiakov D, Cramer P (2018) Structural basis of mitochondrial transcription. Nat Struct Mol Biol 25:754–765
Höfer K, Li S, Abele F et al (2016) Structure and function of the bacterial decapping enzyme NudC. Nat Chem Biol 12:730–734
Huang F (2003) Efficient incorporation of CoA, NAD and FAD into RNA by in vitro transcription. Nucleic Acids Res 31:e8
Hudeček O, Benoni R, Reyes-Gutierrez PE et al (2020) Dinucleoside polyphosphates act as 5'-RNA caps in bacteria. Nat Commun 11:1052
Hui MP, Foley PL, Belasco JG (2014) Messenger RNA degradation in bacterial cells. Annu Rev Genet 48:537–559
Issur M, Bougie I, Despins S et al (2013) Enzymatic synthesis of RNAs capped with nucleotide analogues reveals the molecular basis for substrate selectivity of RNA capping enzyme: impacts on RNA metabolism. PLoS One 8:e75310
Jiao X, Chang JH, Kilic T et al (2013) A mammalian pre-mRNA 5′ end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. Mol Cell 50:104–115
Jiao X, Doamekpor SK, Bird JG et al (2017) 5′ end nicotinamide adenine dinucleotide cap in human cells promotes RNA decay through DXO-mediated deNADding. Cell 168:1015–1027.e10
Jiao X, Xiang S, Oh C et al (2010) Identification of a quality-control mechanism for mRNA 5′-end capping. Nature 467:608–611
Joyce GF (2001) RNA cleavage by the 10-23 DNA enzyme. Methods Enzymol 341:503–517
Julius C, Riaz-Bradley A, Yuzenkova Y (2018) RNA capping by mitochondrial and multi-subunit RNA polymerases. Transcription 9:292–297
Julius C, Yuzenkova Y (2017) Bacterial RNA polymerase caps RNA with various cofactors and cell wall precursors. Nucleic Acids Res 45:8282–8290
Julius C, Yuzenkova Y (2019) Noncanonical RNA-capping: discovery, mechanism, and physiological role debate. Wiley Interdiscip Rev RNA 10:e1512
Kellner S, Burhenne J, Helm M (2010) Detection of RNA modifications. RNA Biol 7:237–247
Kiledjian M (2018) Eukaryotic RNA 5′-end NAD. Trends Cell Biol 28:454–464
Kowtoniuk WE, Shen Y, Heemstra JM et al (2009) A chemical screen for biological small molecule-RNA conjugates reveals CoA-linked RNA. Proc Natl Acad Sci U S A 106:7768–7773
Kramer S, McLennan AG (2019) The complex enzymology of mRNA decapping: enzymes of four classes cleave pyrophosphate bonds. Wiley Interdiscip Rev RNA 10:e1511
Kulbachinskiy A, Mustaev A (2006) Region 3.2 of the sigma subunit contributes to the binding of the 3′-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. J Biol Chem 281:18273–18276
Kushner SR (2004) mRNA decay in prokaryotes and eukaryotes: different approaches to a similar problem. IUBMB Life 56:585–594
Kwasnik A, Wang VY, Krzyszton M et al (2019) Arabidopsis DXO1 links RNA turnover and chloroplast function independently of its enzymatic activity. Nucleic Acids Res 47:4751–4764
Łabno A, Tomecki R, Dziembowski A (2016) Cytoplasmic RNA decay pathways - enzymes and mechanisms. Biochim Biophys Acta 1863:3125–3147
Li X, Xiong X, Wang K et al (2016) Transcriptome-wide mapping reveals reversible and dynamic N1-methyladenosine methylome. Nat Chem Biol 12:311–316
Limbach PA, Paulines MJ (2017) Going global: the new era of mapping modifications in RNA. Wiley Interdiscip Rev RNA 8
Lin SJ, Guarente L (2003) Nicotinamide adenine dinucleotide, a metabolic regulator of transcription, longevity and disease. Curr Opin Cell Biol 15:241–246
Liu K, Lat PK, Yu HZ et al (2020) CLICK-17, a DNA enzyme that harnesses ultra-low concentrations of either cu^{+} or Cu^{2+} to catalyze the azide-alkyne 'click' reaction in water. Nucleic Acids Res 48:7356–7370
Louie TM, Xie XS, Xun LY (2003) Coordinated production and utilization of FADH(2) by NAD(P)H-flavin oxidoreductase and 4 hydroxyphenylacetate 3-monooxygenase. Biochemistry 42:7509–7517

Luciano DJ, Belasco JG (2020) Np4A alarmones function in bacteria as precursors to RNA caps. Proc Natl Acad Sci U S A 117:3560–3567
Luciano DJ, Levenson-Palmer R, Belasco JG (2019) Stresses that raise Np4A levels induce protective nucleoside tetraphosphate capping of bacterial RNA. Mol Cell 75:957–966.e8
Malkowski SN, Spencer TCJ, Breaker RR (2019) Evidence that the nadA motif is a bacterial riboswitch for the ubiquitous enzyme cofactor NAD^+. RNA 25:1616–1627
Malygin AG, Shemyakin MF (1979) Adenosine, NAD and FAD can initiate template-dependent RNA synthesis catalyzed by Escherichia coli RNA polymerase. FEBS Lett 102:51–54
Mao Z, Shin HD, Chen RR (2006) Engineering the E. coli UDP-glucose synthesis pathway for oligosaccharide synthesis. Biotechnol Prog 22:369–374
Mauer J, Luo X, Blanjoie A et al (2017) Reversible methylation of m6Am in the 5′ cap controls mRNA stability. Nature 541:371–375
McLennan AG (2006) The Nudix hydrolase superfamily. Cell Mol Life Sci 63:123–143
Meyer KD, Patil DP, Zhou J et al (2015) 5' UTR m(6)a promotes cap-independent translation. Cell 163:999–1010
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149:1635–1646
Mildvan AS, Xia Z, Azurmendi HF et al (2005) Structures and mechanisms of Nudix hydrolases. Arch Biochem Biophys 433:129–143
Mishima E, Jinno D, Akiyama Y et al (2015) Immuno-northern blotting: detection of RNA modifications by using antibodies against modified nucleosides. PLoS One 10:e0143756
Mitchell SF, Parker R (2015) Modifications on translation initiation. Cell 163:796–798
Morales-Filloy HG, Zhang Y, Nübel G et al (2020) The 5' NAD cap of RNAIII modulates toxin production in Staphylococcus aureus isolates. J Bacteriol 202:e00591–e00519
Nachtergaele S, He C (2018) Chemical modifications in the life of an mRNA transcript. Annu Rev Genet 52:349–372
Namboori SC, Graham DE (2008) Enzymatic analysis of uridine diphosphate N-acetyl-D-glucosamine. Anal Biochem 381:94–100
Nübel G, Sorgenfrei FA, Jäschke A (2017) Boronate affinity electrophoresis for the purification and analysis of cofactor-modified RNAs. Methods 117:14–20
Ogawa T, Yoshimura K, Miyake H et al (2008) Molecular characterization of organelle-type Nudix hydrolases in Arabidopsis. Plant Physiol 3:1412–1424
Paget MS (2015) Bacterial sigma factors and anti-sigma factors: structure, function and distribution. Biomol Ther 5:1245–1265
Pan S, Kai-en L, Huang W et al (2020) Arabidopsis DXO1 possesses deNADding and exonuclease activities and its mutation affects defense-related and photosynthetic gene expression. J Integr Plant Biol 62:967
Picard-Jean F, Brand C, Tremblay-Létourneau M et al (2018) 2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. PLoS One 13:e0193804
Ramanathan A, Robb GB, Chan SH (2016) mRNA capping: biological functions and applications. Nucleic Acids Res 44:7511–7526
Richards J, Liu Q, Pellegrini O et al (2011) An RNA pyrophosphohydrolase triggers 5-′-exonucleolytic degradation of mRNA in Bacillus subtilis. Mol Cell 43:940–949
Ringel R, Sologub M, Morozov YI et al (2011) Structure of human mitochondrial RNA polymerase. Nature 478:269–273
Rostovtsev VV, Green LG, Fokin VV et al (2002) A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. Angew Chem Int Ed Engl 41:2596–2599
Roundtree IA, Evans ME, Pan T et al (2017) Dynamic RNA modifications in gene expression regulation. Cell 169:1187–1200
Ryvkin P, Leung YY, Silverman IM et al (2013) HAMR: high-throughput annotation of modified ribonucleotides. RNA 19:1684–1692

Schwartz S, Motorin Y (2017) Next-generation sequencing technologies for detection of modified nucleotides in RNAs. RNA Biol 14:1124–1137
Sharma S, Grudzien-Nogalska E, Hamilton K et al (2020) Mammalian Nudix proteins cleave nucleotide metabolite caps on RNAs. Nucleic Acids Res 48:6788–6798
Shi H, Wei J, He C (2019) Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers. Mol Cell 74:640–650
Shuman S (2015) RNA capping: progress and prospects. RNA 21:735–737
Sikorski PJ, Warminski M, Kubacka D et al (2020) The identity and methylation status of the first transcribed nucleotide in eukaryotic mRNA 5′ cap modulates protein expression in living cells. Nucleic Acids Res 48:1607–1626
Song MG, Bail S, Kiledjian M (2013) Multiple Nudix family proteins possess mRNA decapping activity. RNA 19:390–399
Srouji JR, Xu A, Park A et al (2017) The evolution of function within the Nudix homology clan. Proteins 85:775–811
Takamura T, Nomura G (1988) Changes in the intracellular concentration of acetyl-CoA and Malonyl-CoA in relation to the carbon and energy-metabolism of Escherichia-Coli-K12. J Gen Microbiol 134:2249–2253
Thomason MK, Bischler T, Eisenbart SK et al (2015) Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in Escherichia coli. J Bacteriol 197:18–28
Thüring K, Schmid K, Keller P et al (2016) Analysis of RNA modifications by liquid chromatography-tandem mass spectrometry. Methods 107:48–56
Tsuchihara K, Suzuki Y, Wakaguri H et al (2009) Massive transcriptional start site analysis of human genes in hypoxia cells. Nucleic Acids Res 37:2249–2263
Turnbough CL, Switzer RL (2008) Regulation of pyrimidine biosynthetic gene expression in bacteria: repression without repressors. Microbiol Mol Biol Rev 72:266–300
Vvedenskaya IO, Bird JG, Zhang Y et al (2018) CapZyme-seq comprehensively defines promoter-sequence determinants for RNA 5′ capping with NAD. Mol Cell 70:553–564.e9
Walters RW, Matheny T, Mizoue LS et al (2017) Identification of NAD^+ capped mRNAs in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 114:480–485
Wang J, Alvin Chew BL, Lai Y et al (2019a) Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. Nucleic Acids Res 47:e130
Wang Z, Jiao X, Carr-Schmid A et al (2002) The hDcp2 protein is a mammalian mRNA decapping enzyme. Proc Natl Acad Sci U S A 99:12663–12668
Wang L, Song C, Wang N et al (2020) NADP modulates RNA m6A methylation and adipogenesis via enhancing FTO activity. Nat Chem Biol
Wang VY, Jiao X, Kiledjian M et al (2015) Structural and biochemical studies of the distinct activity profiles of Rai1 enzymes. Nucleic Acids Res 43:6596–6606
Wang Y, Li S, Zhao Y et al (2019b) NAD^+-capped RNAs are widespread in the Arabidopsis transcriptome and can probably be translated. Proc Natl Acad Sci U S A 116:12094–12102
Werner M, Purta E, Kaminska KH et al (2011) 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. Nucleic Acids Res 39:4756–4768
Wetzel C, Limbach PA (2016) Mass spectrometry of modified RNAs: recent developments. Analyst 141:16–23
Wu H, Li L, Chen KM et al (2019) Decapping enzyme NUDT12 partners with BLMH for cytoplasmic surveillance of NAD-capped RNAs. Cell Rep 29:4422–4434.e13
Xiang S, Cooper-Morgan A, Jiao X et al (2009) Structure and function of the 5′>3′ exoribonuclease Rat1 and its activating partner Rai1. Nature 458:784–788
Yaginuma H, Kawai S, Tabata KV et al (2014) Hiromi Imamura diversity in ATP concentrations in a single bacterial cell population revealed by quantitative single-cell imaging. Sci Rep 6:6522
Yang H, Yang T, Baur JA et al (2007) Nutrient-sensitive mitochondrial NAD+ levels dictate cell survival. Cell 130:1095–1107

Yoshimura K, Shigeoka S (2015) Versatile physiological functions of the Nudix hydrolase family in Arabidopsis. Biosci Biotechnol Biochem 79:354–366

Zhang D, Liu Y, Wang Q et al (2016) Structural basis of prokaryotic NAD-RNA decapping by NudC. Cell Res 26:1062–1066

Zhang H, Zhong H, Zhang S et al (2019a) NAD tagSeq reveals that NAD^+-capped RNAs are mostly produced from a large number of protein-coding genes in Arabidopsis. Proc Natl Acad Sci U S A 116:12072–12077

Zhang LS, Liu C, Ma H et al (2019b) Transcriptome-wide mapping of internal N7-methylguanosine methylome in mammalian mRNA. Mol Cell 74:1304–1316.e8

Zhang YQ, Kuster D, Schmidt T et al (2020) Extensive 5′-surveillance guards against non-canonical NAD-caps of nuclear mRNAs in yeast. Nat Commun 11:5508

Zhou Y, Wang L, Yang F et al (2011) Determining the extremes of the cellular NAD(H) level by using an Escherichia coli NAD(+)-auxotrophic mutant. Appl Environ Microbiol 77:6133–6140

ADAR-Mediated RNA Editing and Its Therapeutic Potentials

Beatrice Casati, Dimitra Stamkopoulou, Rafail Nikolaos Tasakis, and Riccardo Pecori

Contents

Abstract Adenosine-to-inosine (A-to-I) RNA editing mediated by the ADAR (adenosine deaminase acting on RNA) protein family is the primary type of epitranscriptomic modification known to occur in mammal cells. Recently, several technologies have been developed to re-target this RNA modification to desired locations within specific transcripts. This possibility opened a scenario in which targeted RNA-base editing tools can be used as therapeutic strategies to correct

Beatrice Casati and Dimitra Stamkopoulou contributed equally to this work.

B. Casati · R. N. Tasakis
Division of Immune Diversity, Program in Cancer Immunology, German Cancer Research Centre (DKFZ), Heidelberg, Germany

Faculty of Biosciences, University of Heidelberg, Heidelberg, Germany

D. Stamkopoulou
Division of Immune Diversity, Program in Cancer Immunology, German Cancer Research Centre (DKFZ), Heidelberg, Germany

Medical Faculty of Mannheim, University of Heidelberg, Mannheim, Germany

R. Pecori (✉)
Division of Immune Diversity, Program in Cancer Immunology, German Cancer Research Centre (DKFZ), Heidelberg, Germany
e-mail: r.pecori@dkfz-heidelberg.de

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_18

mutations at the RNA level. The chapter will go into detail about the therapeutic potentials of these different RNA base-editing technologies, after providing a brief overview of the roles and functions of ADAR family members. The chapter aims to review the recent advancements of targeted RNA-base editing methodologies and their translation to therapeutic settings. We will discuss strategies leveraging exogenous and endogenous ADAR to create a wholesome perspective on the potential of this molecular mechanism as a tool to correct disease-causing G-to-A point mutations. In this context, clinically relevant approaches and their potential future applications, as well as their currently challenging limitations, will be evaluated.

Keywords A-to-I · RNA editing · ADAR · Targeted RNA-base editing tools · RNA therapy · Cas13 · RESTORE · LEAPER

Abbreviations

A	Adenosine
A1AT	α1-antitrypsin
AAV	Adeno-associated virus
AAV8	Adeno-associated virus vector 8
ADAR	Adenosine deaminase acting on RNA
$ADAR2_{DD}$	Deamination domain of ADAR2
$ADAR2^{E488Q}$	Hyperactive ADAR2 mutant carrying the E488Q mutation
ADAT	Adenosine deaminase acting on tRNAs
adRNA or AD-gRNA	ADAR guiding RNA
adV	Adenovirus
ALS	Amyotrophic lateral sclerosis
arRNA	ADAR recruiting RNA
ASOs	Antisense oligonucleotides
A-to-C	Adenosine to cytosine
A-to-I	Adenosine to inosine
BG	O^6-benzylguanine
C	Cytosine
CDA	Cytidine deaminase acting on mononucleotides
CFTR	Cystic fibrosis transmembrane conductance regulator
CIRTS	CRISPR-Cas-inspired RNA targeting system
CNS	Central nervous system
dCas13b	Catalytically inactive mutant of PspCas13b protein
DMD	Duchenne muscular dystrophy
dsRBD	Double-stranded RNA-binding domain
dsRBM	Double-stranded RNA-binding motif
dsRNA	Double-stranded RNA
eCFP	Enhanced cyan fluorescent protein
eGFP	Enhanced green fluorescent protein

EONs	Editing oligonucleotides
G	Guanosine
GluR2	Glutamate ionotropic receptor
gRNA	Guide RNA
G-to-A	Guanosine to adenosine
HEK	Human embryonic kidney
HeLa cells	Henrietta Lacks cervical cancer cells
HEPN	Higher-eukaryotes and prokaryotes nucleotide binding domain
I	Inosine
IDUA	α-L-iduronidase
IFN	Interferon
IFN-I type	Interferon type I
irCLASH	Infrared-dye-conjugated and biotinylated ligation adapter-based CLASH protocol (Crosslinking, Ligation And Sequencing of Hybrids)
ISGs	Interferon stimulated genes
IV	Intravenous
IVT	Intravitreal
LEAPER	Leveraging endogenous ADAR for programmable editing of RNA
LINEs	Long interspersed elements
MAVS	Mitochondrial antiviral signalling protein
MCP	MS2 coat protein
MDA5	Melanoma differentiation-associated protein 5
MECP2	Methyl CpG binding protein 2
MEF	Mouse embryonic fibroblasts
MPS I	Mucopolysaccharidosis type I
NES	Nuclear export signal
NLS	Nuclear localisation signal
OTC	Ornithine transcarbamylase
PKR	Protein kinase R
PRRs	Pattern recognition receptors
Q/R site	Glutamine (Q) to arginine (R) site
R/G motif	Arginine (R) to glycine (G) motif
REPAIR	RNA Editing for Programmable A to I Replacement
RESCUE	RNA Editing for Specific C-to-U Exchange
RESTORE	Recruiting Endogenous ADAR to Specific Transcripts for Oligonucleotide-mediated RNA Editing
RIG-I	Retinoid acid-inducible gene I
RLR	RIG-I like receptor
RTT	Rett syndrome
ssRNA	Single-stranded RNA

TALENs	Transcription activator-like effector nucleases
tRNA	Transfer RNA
U	Uracil
UTRs	Untranslated regions
ZFNs	Zinc finger nucleases
λN	λ-phage N protein

1 Adenosine to Inosine (A-to-I) RNA Editing

RNA editing is a broad term that describes modifications at the RNA level introduced during or after transcription that may result in changes in the coding capacity of the edited RNA molecule (Keegan et al. 2001). First described in mitochondria of *Trypanosoma* (Benne et al. 1986), RNA editing can result in the insertion, deletion or modification of nucleotides. Among the best characterised RNA editing mechanisms are the conversions of cytosine (C) to uracil (U) and of adenosine (A) to inosine (I) (Keegan et al. 2001). Here we will focus on the latter modification, which is the most commonly found in mammals. The resulting inosine from an A-to-I editing event is interpreted by the cellular translation machinery as a guanosine (G) (Basilio et al. 1962). The enzymes that catalyse the deamination of adenosine in double-stranded RNA (dsRNA) substrates are members of the family of adenosine deaminases acting on RNA (ADAR) (Bass et al. 1997).

2 ADARs: Role and Function

In mammals, the ADAR family is composed of three members called ADAR, ADARB1 and ADARB2, or most commonly ADAR1, ADAR2 and ADAR3, respectively. ADAR1 and ADAR2 are RNA editors through their deaminase activity, while the function of ADAR3 is still incompletely understood although it is thought to be an RNA editor inhibitor (Oakes et al. 2017). These three enzymes are highly conserved in vertebrates (Savva et al. 2012). ADARs are thought to derive from adenosine deaminases acting on tRNAs (ADATs), which are responsible for A-to-I editing of transfer RNAs (tRNAs) at or near the anticodon position (Savva et al. 2012). Both families likely derive from the cytidine deaminase acting on mononucleotides (CDAs) (Gerber and Keller 2001).

ADAR structure is partially conserved among the different members of this family, with the catalytic domain located at the C-terminus of the protein (Wang et al. 2017). All ADARs have one to three repeats of the double-stranded RNA-binding domain (dsRBD), which binds the dsRNA.

2.1 ADAR1 and ADAR2

In order to perform editing activity, both ADAR1 and ADAR2 form homodimers, independently of RNA binding. The interaction between functional monomers is fundamental for successful editing and might also determine the site selectivity of each enzyme by placing specific adenosine residues in the correct position within the catalytic domain (Cho et al. 2003).

ADAR1 and ADAR2 can be set apart by the number of dsRBDs they contain and the distance between them (Nishikura 2016). These features play an important role in determining their differential specificity for certain editing sites (Wong et al. 2001). Furthermore, their specificity is also defined by secondary structures within ADAR substrates, as internal loops within the dsRNA helix have been observed to relate to the selection of the correct adenosine for deamination (Lehmann and Bass 1999). An exhaustive study about ADAR RNA substrate recognition has been recently conducted (Song et al. 2020). Here, a novel tool for the identification of ADAR substrates and binding preferences was developed. The new method, called irCLASH, combines an infrared-dye-conjugated and biotinylated ligation adapter with the CLASH protocol (Crosslinking, Ligation And Sequencing of Hybrids), a previously reported method for the detection of ribonucleoprotein complexes between RNA-binding proteins and specific RNA duplexes (Helwak and Tollervey 2014). Application of irCLASH at a transcriptome-wide level uncovered features and previously unknown details determining ADARs binding and editing activity. The unprecedented data revealed by irCLASH will be of great assistance in the development of ADAR-mediated targeted RNA base-editing strategies.

ADAR1 and ADAR2 are distributed widely in our body tissues. ADAR1 is ubiquitously expressed and exists in two isoforms: a full-length ADAR1 p150 and a shorter ADAR1 p110. The shorter isoform is the result of differential transcription start: exon 1, containing the first upstream start codon (AUG1), is skipped and transcription starts from a second downstream start codon in exon 2 (AUG296). ADAR1 p110 is constitutively expressed, while ADAR1 p150 is interferon-inducible (Patterson and Samuel 1995). The differential expression of these enzymes reflects their distinct functions and cellular localisations. ADAR1 p110 is mainly localised in the nucleus, while ADAR1 p150 is mainly detected in the cytoplasm (Strehblow et al. 2002).

As already mentioned, ADARs convert adenosine to inosine, which is then read as a guanosine by the translation machinery, in dsRNA substrates. If the editing event affects a protein-coding sequence, it could lead to non-synonymous codon change (recoding), and consequently to protein function alteration. Coding sites are predominantly edited by ADAR2 (Samuel 2019). Additionally, RNA editing by ADAR2 can regulate splicing (Rueter et al. 1999) and may sometimes influence splicing efficiency (Higuchi et al. 2000). Among the limited number of cellular genes that are site-selectively A-to-I edited by ADAR2 within their coding sequence, the vast majority includes neurotransmitter receptors and ion channels. The outcome of RNA editing of protein coding genes is most often the synthesis of different

variants of the same protein, which contributes to the diversification of protein functions. The first and best characterised example of ADAR2-mediated RNA editing is the modification of the pre-mRNA coding for the subunit 2 of the AMPA subtype of the glutamate ionotropic receptor (GluR2) (Fig. 1a) (Sommer et al. 1991). The mature glutamate receptor consists of four subunits that are assembled in different possible combinations to form an ion channel. Upon binding of its ligand, the receptor gets activated and the resulting calcium (Ca^{2+}) influx mediates fast excitatory synaptic transmission in the central nervous system (Twomey et al. 2017). GluR2 pre-mRNA is edited at various levels at different sites, resulting in amino acid changes that affect the Ca^{2+} permeability of the final receptor (Egebjerg and Heinemann 1993). The best studied examples are the substitution of a glutamine (Q) to an arginine (R) at the so-called Q/R site, which has been demonstrated to reduce GluR2 Ca^{2+} permeability, and the substitution of an arginine (R) to a glycine (G) at the so-called R/G site, which has been associated with different GluR2 recovery time from desensitization (Sommer et al. 1991; Seeburg et al. 1998). Thus, ADAR2-mediated editing of GluR2 mRNA provides an example of exon recoding and protein function regulation by RNA editing.

Most often, A-to-I RNA editing occurs in non-coding sequences, such as 5′ and 3′ untranslated regions (UTRs) and intronic retrotransposon elements, such as Alu repeats and long interspersed elements (LINEs) (Nishikura 2010). A-to-I editing of non-coding regions is mainly dependent on ADAR1 (Samuel 2019) and it has several potential functions and consequences. For example, editing of Alu dsRNAs may signal for their degradation, and so the expression of genes harbouring Alu dsRNAs can be controlled. This editing can also lead to heterochromatin formation and gene silencing in a region with an abundance of Alu sequences (Nishikura 2016). However, the main function of editing mediated by ADAR1 (p110) on endogenous dsRNAs is to prevent their recognition by dsRNA sensors, thus avoiding an interferon response against self dsRNAs (Liddicoat et al. 2015; Pestal et al. 2015). Instead, the interferon-inducible ADAR1 p150 is mainly responsible for A-to-I editing during viral infections upon interferon (IFN) signalling, whose aim is to mount an effective immune response against the pathogen (Chung et al. 2018).

While ADAR2 is mainly active in the central nervous system (CNS), ADAR1 role has a broader impact on the whole organism. It is established that ADAR1, more precisely its IFN-inducible isoform ADAR1 p150, is a master regulator in antiviral mechanisms and immunity (Lamers et al. 2019). In the event of viral infection, double-stranded RNA structures may derive from the replication of the virus inside the cell. Specialised receptors in the cytoplasm called pattern recognition receptors (PRRs) have evolved as sensory mechanisms to detect these molecular patterns and induce a response against them. The main system involved in the detection of dsRNAs is the RIG-I like receptor (RLR) signalling pathway. Its name derives from retinoid acid-inducible gene I (*RIG-I*), the first discovered member of the PRR family, soon followed by the melanoma differentiation-associated protein 5 (MDA5). Activation of the RLR pathway triggers a signalling cascade that leads to the production of antiviral type I interferons (IFN-I) (Borden 2019). The resulting IFNs induce an antiviral state by signalling the threat to neighbouring cells and by

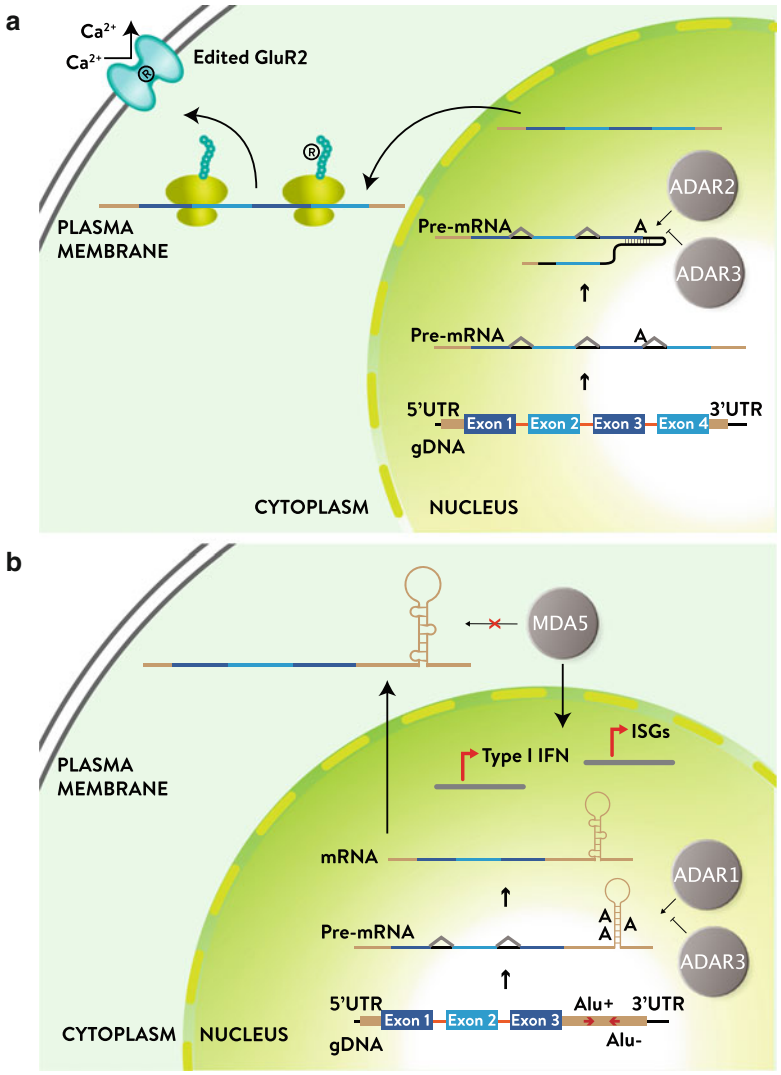


Fig. 1 Roles and functions of ADAR 1 and 2. (**a**) Illustration of the editing mechanism of the mRNA coding for the glutamate ionotropic receptor AMPA type subunit 2 (GluR2). GluR2 pre-mRNA contains a hairpin-shaped motif, the Q/R editing site, that is recognised by both ADAR2 and ADAR3. ADAR2 binding results in deamination of a specific adenosine (A) to inosine (I), while ADAR3 binding inhibits the editing mechanism by preventing ADAR2 from accessing GluR2 pre-mRNA. The A-to-I modification in the pre-mRNA leads to an amino acid change at the protein level: edited GluR2 presents a glutamine (Q) to arginine (R) substitution, which causes a reduction of Ca^{2+} permeability of the receptor. (**b**) Visual representation of ADAR1 role in innate immunity. By editing endogenous dsRNAs in the nucleus, ADAR1 prevents their recognition by dsRNA sensors, such as MDA5, which is essential to detect viral RNA molecules and induce an immune response against them. Activation of MDA5 triggers a signalling cascade that leads to the production of antiviral type I interferon (type I IFN) and expression of interferon stimulated genes (ISGs), whose products are proteins with different antiviral functions. ADAR1-mediated RNA editing of endogenous dsRNAs prevents self-mediated hyperactivation of IFN-stimulated

promoting the upregulation of interferon stimulated genes (ISGs), as summarised in Fig. 1b. Their products are proteins with different antiviral functions whose final aim is the elimination of the pathogen.

An important feature of this mechanism is the ability to differentiate between viral dsRNAs, which pose a real threat, and self and innocuous dsRNAs, which can naturally form from endogenous RNAs. It has been recently demonstrated that ADAR1-mediated A-to-I editing evolved in humans as a mechanism to distinguish self from non-self RNA (Liddicoat et al. 2015). By editing endogenous dsRNAs and introducing base pair mismatches in their sequence, ADAR1 destabilises their duplex structures and prevents them from being recognised by PRRs in the cytoplasm, thus avoiding hyperactivation of the interferon-induced dsRNA sensor protein kinase R (PKR), subsequent translational shutdown and cell death (Chung et al. 2018). This system ensures that self-mediated IFN production does not lead to autoinflammation and autoimmunity (Pestal et al. 2015).

The importance of accurate RNA editing has been further established thanks to several studies that focused on the impact of aberrant editing or malfunctioning RNA editing enzymes on our health. RNA editing deficiencies and alterations are associated both with physiological defects in different organisms and with human diseases and pathophysiologies, including cancer, highlighting its role in disease development and progression (Bajad et al. 2017).

Mouse models of ADAR deficiencies have been an invaluable tool for better characterising these enzymes. *ADAR1* knockout mice show embryonic lethality caused by hematopoiesis failure and a global interferon response that leads to systemic apoptosis (Wang et al. 2004). The effects of ADAR1 absence can be rescued by simultaneous knockout of either *MAVS* (mitochondrial antiviral signalling protein) or *MDA5*, thus further indicating that ADAR1 is fundamental in modulating the detection of self dsRNAs. More precisely, it was observed that mice embryos lacking ADAR1 p150 had the same phenotype as *ADAR1* knockout embryos, thus indicating that the p150 isoform plays a critical role in embryonic development (Ward et al. 2011). In humans, heterozygous mutations for functional null *ADAR1* gene causes dyschromatosis symmetrica hereditaria, an autosomal dominant pigmentary genodermatosis (Zhang et al. 2004). *ADAR1* is also mutated in Aicardi-Goutières syndrome, an autoinflammatory disorder characterised by spontaneous interferon production and neurologic sequelae (Rice et al. 2012).

Lack of ADAR2 is slightly less severe than ADAR1 depletion: mice survive until few weeks after birth, but they quickly develop seizures and die shortly afterwards. The introduction of a permanent mutation at the Q/R site of GluR2, mimicking ADAR2-mediated editing, is able to rescue the effect of lack of ADAR2, showing that lethality is dependent on GluR2 editing (Higuchi et al. 2000). In humans, several disorders of the central nervous system are caused by ADAR2-mediated

Fig. 1 (continued) pathways, autoinflammation and autoimmunity. The whole mechanism may be inhibited by ADAR3, which competes with ADAR1 for dsRNA binding, thus preventing its function

RNA-editing deficiencies, among which are sporadic amyotrophic lateral sclerosis (ALS), ischemia and certain neuropsychiatric disorders (Gallo et al. 2017).

Finally, ADARs editing plays also a relevant role in cancer. The first discovery of aberrant A-to-I editing in tumours was reported in 2001 when hypo-editing of the glutamate receptor *GluR2* mRNA was observed in malignant gliomas (Maas et al. 2001). Since then, altered A-to-I editing has been observed on several transcripts in many different types of cancer, thus demonstrating its involvement in carcinogenesis (Xu and Öhman 2019).

2.2 ADAR3

ADAR3 is vertebrate specific and it is mainly present in the brain. It is very similar to ADAR1 and ADAR2 in its structure, but it has not been proven to have deaminase activity in vivo so far. Nonetheless, it can bind dsRNA, as it contains a dsRNA binding domain, but also single-stranded RNA (ssRNA) via an arginine (R) rich domain that consists of arginine residues (Chen et al. 2000). This R-domain has been proposed to function as a nuclear localisation signal (NLS), as it can act as a mediator between ADAR3 and importin-α, the protein responsible for ADAR3 migration to the nucleus (Maas and Gommans 2009). It is highly expressed in the nervous system and it has been suggested to play an integral role in the correct cognitive functioning of the brain in mammals (Mladenova et al. 2018).

3 Targeted A-to-I RNA Base-Editing Technologies Towards the Development of Therapeutic Tools

While different tools have already been developed and employed to specifically modify DNA bases (Li et al. 2020), not as many are available for targeted RNA base-editing due to the newly arisen interest in this mechanism. So far, only A-to-I targeted RNA base-editing is possible and it employs ADAR proteins (Vogel and Stafforst 2019). By expressing the full ADAR protein or only its deaminase domain fused with a targeting domain, either a SNAP-tag, a λN peptide, an R/G motif, a Cas protein (dCas13b) or an MS2-tag, together with a guide RNA, targeted RNA editing can be induced at defined locations within specific transcripts. All these approaches rely on the delivery of both the editing enzyme and a guide RNA (gRNA), which is an RNA molecule complementary to the to-be-edited target sequence and able to recruit the editing enzyme. The fact that both these components need to be present makes their clinical applications more challenging. A great step forward for the development of RNA editing-based therapies would be achieved with the design of guides able to recruit endogenous ADARs. A few approaches towards this goal were recently published where endogenous ADAR was recruited by delivering

chemosynthetic antisense oligonucleotides (Merkle et al. 2019) or ADAR-recruiting RNAs (Qu et al. 2019).

The idea of targeted RNA editing as a new gene therapy approach was first presented in 1995 (Woolf et al. 1995). Since then, different studies have been performed with the aim of better understanding the biology of RNA editing and further exploring its therapeutic potential. The pioneering system for targeted RNA editing developed by Woolf and colleagues aimed at correcting a premature stop codon in the dystrophin mRNA fused to a luciferase reporter (Woolf et al. 1995). In order to induce targeted RNA editing, the mutated dystrophin mRNA was hybridised to a complementary 52-nucleotide RNA molecule and either treated with nuclear extract containing ADARs or delivered into single-cell *Xenopus* embryos, characterised by high levels of endogenous ADARs. In both cases, an increase in luciferase expression was observed in the presence of the hybrid. Based on these promising observations, the use of targeted RNA editing to therapeutically correct mutations in humans was suggested.

3.1 RNA vs DNA Editing: Advantages and Disadvantages

While RNA editing application in therapy is a relatively new field of research, DNA therapy has been extensively investigated over the past years. Both approaches have advantages and limitations that need to be considered when choosing one or the other technology for the treatment of a specific disease.

An undiscussed strength of RNA editing as a tool to correct mutations is the transient impact of introducing changes at the RNA level. Since the genetic information at the DNA level is not affected, potential errors or undesired unspecific modifications in the transcriptome, such as off-target editing, would not cause any permanent damage and could be reversed by stopping the treatment. Moreover, the duration of the effects of targeted RNA-base editing can be fine-tuned by shortening or extending the administration. On the other side, the transient nature of RNA base-editing approaches is also a limitation, because repeated administration is required to continuously correct a specific mutation. The opposite is true for DNA editing technologies: by introducing a modification at the genome level, a single administration is potentially enough to correct a mutation permanently, but the risk of unspecific DNA editing is also more concerning (Tang and Xu 2020).

The reports about off-target mutagenesis within DNA targeting gene therapies are still contradictive. Some studies show evidence that using a CRISPR-based system can cause significant off-target mutagenesis in various human cell types (Fu et al. 2013; Pattanayak et al. 2013), while others make considerable steps towards the reduction of such an effect and the optimisation of the specificity (Cho et al. 2014; Gaudelli et al. 2017; Wang et al. 2019; Zetche et al. 2015). Additional genome editing approaches that utilise other programmable nucleases like zinc finger nucleases (ZFNs) (Kim et al. 2011) and transcription activator-like effector nucleases (TALENs) (Miller et al. 2011) rely on protein–DNA interactions for their specificity,

which can often be unpredictable and can potentially have substantial off-target effects (Gabriel et al. 2011). Overall, potential off-target events pose a bigger safety issue in the context of DNA gene therapy than in that of RNA editing, as they can lead to genomic instability and destroy the functionality of otherwise normal genes (Tang and Xu 2020; Zhang et al. 2015).

Additionally, it is important to note that RNA-base editing approaches are being currently developed that can leverage endogenously expressed ADAR (Merkle et al. 2019; Qu et al. 2019). This could eliminate possible limitations of DNA editing systems, like CRISPR/Cas9, whose main component is of bacterial origin and upon delivery may elicit an undesired host immune response when trying to achieve in vivo gene therapy (Charlesworth et al. 2019). This could potentially be true also for systems using exogenous ADAR fused to non-human proteins, such as λ-phage N protein (Montiel-Gonzalez et al. 2013), illustrated in Fig. 2b, and bacteriophage MS2 coat protein (Katrekar et al. 2019), represented in Fig. 2f, thus further highlighting the importance of optimising strategies for the recruitment of endogenous ADARs.

However, by definition, ADAR-mediated RNA editing can only correct G-to-A mutations by reverting adenosine to inosine, then translated as guanosine. Instead, genome targeting approaches can correct several kinds of mutations, such as deletions and different point mutations, thus making DNA-base editing therapy more versatile and broadly applicable (Li et al. 2020). A crucial limitation for the translation of both approaches into the clinic is delivery. For genome editing approaches, nonviral delivery methods have been widely and successfully employed to deliver genetic editing elements in vitro, but when applied in vivo they resulted in limited editing efficiency (Mout et al. 2017; Li et al. 2020; Nóbrega et al. 2020). Instead, viral vectors such as retrovirus, lentivirus, adenovirus (AdV) and adeno-associated virus (AAV) have shown high efficiency (Li et al. 2020), but they also increase the possibility of introducing unintentional mutations (Hacein-Bey-Abina et al. 2008) and of inducing immune responses in some instances (Lundstrom 2018; Tang and Xu 2020). Viral particles have also been used for the delivery of RNA editing systems, which are usually easily packaged due to the small size of the components. Even when the construct consists of a ribonucleoprotein complex between enzymes and guide RNAs, it is still small enough to be delivered via a single AAV vector. Conversely, DNA targeting constructs are often too large to fit in a single AAV; therefore strategies that allow the simultaneous delivery of two different vectors are under investigation (Mout et al. 2017). For constructs that require chemical modifications to increase their stability, as in the case of some RNA oligonucleotides, AAV-mediated delivery is not possible, since these modifications can only be introduced via chemical synthesis. This poses an obstacle to the therapeutic translation of such approaches because an alternative delivery strategy needs to be found (Vogel et al. 2018; Merkle et al. 2019). On the other side, unmodified RNA constructs can be delivered via both viral and nonviral vectors and then expressed in vivo, but there is no insurance of the stability of these RNA molecules within the recipient cells (Qu et al. 2019).

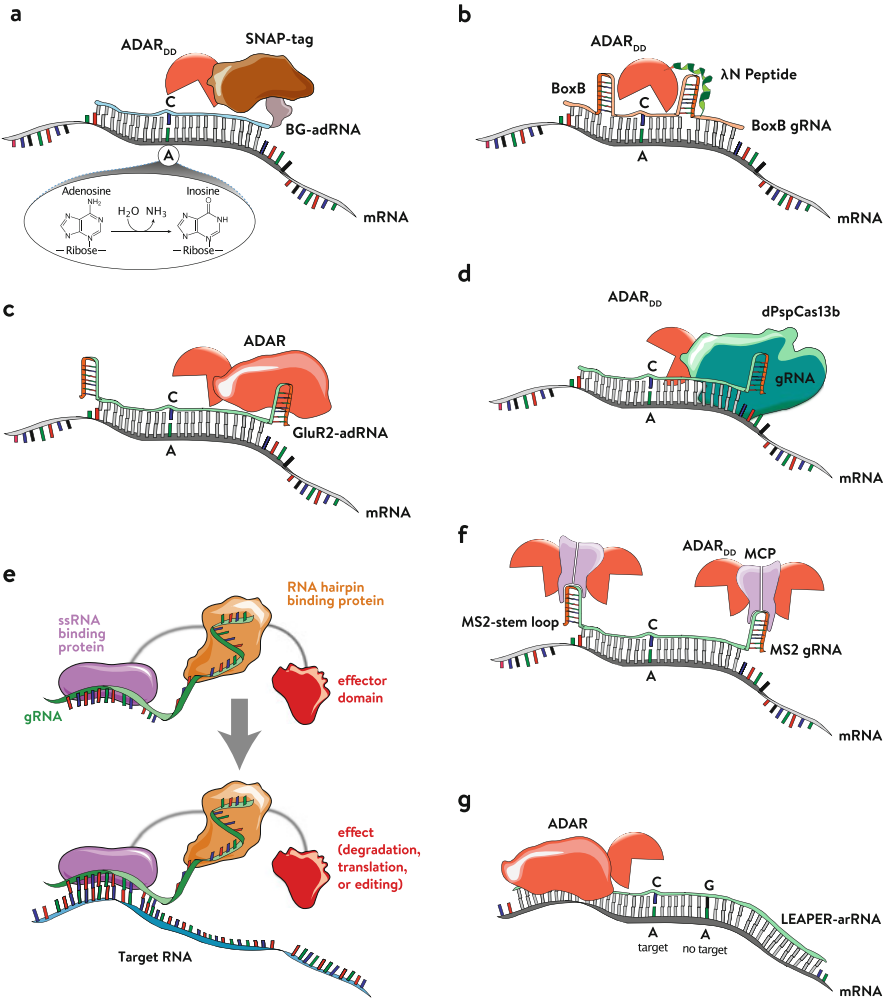


Fig. 2 Targeted A-to-I RNA base-editing technologies. (**a**) Illustration of the SNAP-ADAR system and the ADAR-mediated A-to-I editing mechanism. SNAP-ADARs are fusion proteins made of the deaminase domain of ADAR ($ADAR_{DD}$) and a SNAP-tag self-labeling protein derived from the human O^6-alkylguanine-DNA alkyltransferase. When combined with an ADAR recruiting guide RNA carrying the chemical tag O^6-benzylguanine (BG-adRNA), SNAP-ADARs covalently bind to the BG-adRNA, which is complementary to the target mRNA and thus determines SNAP-ADAR editing specificity. Opposite the targeted adenosine (A), a cytosine (C) mismatch in the BG-adRNA increases editing efficiency, leading to the deamination of adenosine into inosine. (**b**) Representation of the λN-ADAR system. The λN-ADAR system is based on the λ-phage N protein and its natural ability to recognise and bind boxB hairpin structures on RNA. Here, the deaminase domain of ADAR2 is fused to the λN peptide and the specific guide RNA complementary to the target sequence of interest carries the boxB structure. Upon non-covalent interaction between the λN peptide and the boxB structure, ADAR is recruited to the target mRNA sequence. Opposite the targeted adenosine (A), a cytosine (C) mismatch in the boxB-gRNA increases editing efficiency. (**c**) Illustration of the GluR2-ADAR system. The GluR2-ADAR system is based on the pre-mRNA encoding subunits of the GluR2 receptor containing a highly edited, specific, hairpin-shaped motif, the R/G editing site, that is recognised by the double-stranded RNA-binding motifs (dsRBMs) of

Overall, the relatively novel RNA editing approach holds promise for targeted therapy, despite the fact that some challenges still need to be addressed. Several different systems have been developed and described so far. In the following paragraphs we will elaborate on those who recruit ADAR enzymes to direct targeted editing to specific transcripts. The main features of each system are summarized in Table 1.

3.1.1 SNAP-ADAR

The SNAP-ADAR system (Stafforst and Schneider 2012) is based on fusion proteins made of the deaminase domain of ADAR and a SNAP-tag self-labeling protein derived from the human O^6-alkylguanine-DNA alkyltransferase (Keppler et al. 2003). When combined with guide RNAs carrying the chemical tag, O^6-benzylguanine (BG), SNAP-ADARs covalently bind to the guide RNAs, thus creating an efficient and precise ribonucleoprotein with editing activity. Since SNAP-ADARs lack their natural RNA binding domain, their specificity is determined by the gRNA, whose pairing with the complementary target sequence provides the double-stranded RNA substrate needed for efficient editing, as illustrated in Fig. 2a. The first demonstration of successful editing with this system was the

Fig. 2 (continued) ADAR2. ADAR-recruiting RNA carrying the R/G motif (GluR2-adRNA) can recruit ADAR2 to a specific mRNA in a way that mimics the natural editing process. Opposite the targeted adenosine (A), a cytosine (C) mismatch in the GluR2-adRNA increases editing efficiency. (**d**) Representation of the dPspCas13b-$ADAR_{DD}$ system. dPspCas13b-$ADAR_{DD}$ is a fusion protein between a catalytically inactive mutant of PspCas13b and the deamination domain of ADAR2 ($ADAR2_{DD}$). dPspCas13b-$ADAR_{DD}$ can be recruited to a specific site for targeted editing when administered together with a specific guide RNA (gRNA) containing a region complementary to the target mRNA, an A-to-C mismatch at the targeted site and a direct repeat sequence which forms the stem loop structure required for Cas13 recruitment. (**e**) Illustration of CIRTS (CRISPR-Cas-inspired RNA targeting system). CIRTS is based on four components: an RNA-hairpin-binding protein (orange) that can specifically bind to engineered guide RNA (gRNA) with high-affinity; a gRNA (green) with a complementary sequence to the target RNA (blue) and a specific structure that interacts with the RNA-hairpin-binding protein; a single-stranded RNA (ssRNA) binding protein (purple) that can stabilize and protect the gRNA before it interacts with the target; an effector domain (red) that acts on the targeted RNA once all the components are in place. (**f**) Representation of the MS2-MCP-ADAR system. The MS2-MCP-ADAR system is based on the interaction between MS2 stem loops naturally occurring in the phage genome and the MS2 coat binding protein (MCP). Here, adRNAs containing MS2 stem loops and a region complementary to the target mRNA can recruit the deaminase domain of ADAR ($ADAR_{DD}$) fused to MCP and specifically direct the fusion protein to the targeted site. (**g**) Illustration of the LEAPER system (leveraging endogenous ADAR for programmable editing of RNA). LEAPER is based on long, linear RNAs (LEAPER-arRNA) that are engineered so they can recruit endogenous ADAR proteins to target mRNAs for precise A-to-I editing. The arRNA is designed with an almost complementary sequence to its target transcript except for a specific mismatch that introduces a cytosine (C) opposite the targeted adenosine (A) and guanosine (G) opposite off-target adenosines along the LEAPER-arRNA-mRNA duplex

Table 1 Comparison between different RNA base-editing systems

	SNAP-ADAR	λN-ADAR	GluR2-ADAR	CAS-ADAR (REPAIR)	CIRTS	MS2-MCP-ADAR	RESTORE	LEAPER	AXIOMER
ADAR	Exogenous ADARs	Exogenous ADAR2	Exogenous ADAR2	Exogenous ADAR2	Exogenous ADAR2	Exogenous ADARs	Endogenous ADARs	Endogenous ADARs	Endogenous ADARs
Guide RNA (gRNA) length	17 nt	65–76–8692 nt	63 nt	66–86 - 106 - 120 nt	20–40 nt	20 nt	63–95 nt	111–151 nt	35 nt
gRNA modifications	BG tag, 2-′-methoxy groups on all rNTPs but the 3 centred around the targeted A, 2 phosphothioates at the 5′-term and 4 at the 3′-term	BoxB structures	GluR2 R/G motif	DR to recruit CAS	Hairpin structures	MS2 loops from phage genome	Synthetic ASOs with 2′-methoxy and phosphothioate modifications	None	Backbone modifications to allow RNA binding and avoid off-target
Mechanism	BG tagged gRNA recruits SNAP-ADAR by covalent interaction between BG tag and SNAP	boxB structures on the gRNA are recognised by the λN peptide fused to ADAR and they interact (non-covalent interaction)	ADAR2 recognises and interacts with the R/G motif on the gRNA and thus it gets recruited to the complementary target mRNA	DR sequence forms a loop structure on gRNA which recruits CAS which is fused to ADAR	Hairpin structure on gRNA interacts with dsRBD of a protein complex containing also an effector protein which is thus recruited to the target	MCP fused to ADAR recognises and binds MS2 loops on gRNA, thus driving ADAR to the target RNA complementary to the gRNA	gRNA creates dsRNA hybrid with the target, thus recruiting endogenous ADAR that will edit the A-to-C mismatch site	gRNA creates dsRNA hybrid with the target, thus recruiting endogenous ADAR that will edit the A-to-C mismatch site	EON binds a specific target thus creating a dsRNA duplex which recruits ADAR and allows editing at the targeted site
Delivery	Transfection	In cells: lipofection; in vivo: AAV	In cells: lipofection; in vivo: AAV	AAV	AAV	AAV in vivo	ASOs transfection	Plasmid or lentiviral vector, or arRNA transfection	In cells: transfection; in vivo: IVT naked EON injection or tail IV liposomal vectors injection

Clinically relevant corrected mutations	F5 Leiden G1746 > A mutation, KRAS, STAT1	CFTR, R106Q in Mecp2 mRNA causing Rett syndrome	W437Stop mutation in PINK1 associated with Parkinson's disease, dystrophin gene, OTC gene	Mutation in AVPR (associated with nephrogenic diabetes insipidus), mutation in FANCC (associated with Fanconi anemia) and others	KRAS4b transcript fused to a luciferase reporter	Dystrophin gene	Phosphotyrosine 701 in STAT1; PiZZ mutation causing α1-antitrypsin deficiency (E342K in SERPINA1)	αL-iduronidase catalytic activity (Hurler syndrome), TP53, COL3A1, MBPR2, AHI1, FANCC, MYBPC3, IL2RG	Murine versions of human IDUA W402X causing Hurler syndrome (W392X)
Cellular model	HEK293T	*Xenopus* oocytes, HEK293T, murine neurons carrying human R106Q Mecp2	HeLa, HEK293T	HEK293FT	HEK293T	HEK293T	Human cell lines and primary cells	Different human and mouse cell types, including primary fibroblasts from Hurler patient	Hepa1–6 cells and MEF from Hurler syndrome mouse model
In vivo model	None	Human R106Q Mecp2 mice	mdx mouse model for DMD, spf^{ash} mouse model for OTC deficiency	None	None	mdx mouse model	None	None	Hurler syndrome mouse model
Multiplexing	Same efficiency as singleplexing both when targeting two sites on the same transcript and two sites on two different transcripts	NA	NA	NA	Target multiple effectors to multiple targets	NA	Co-transfection of two ASOs targeting two different transcripts	Co-transfection of two arRNAs	NA

(continued)

Table 1 (continued)

	SNAP-ADAR	λN-ADAR	GluR2-ADAR	CAS-ADAR (REPAIR)	CIRTS	MS2-MCP-ADAR	RESTORE	LEAPER	AXIOMER
Targeted disease relevant endogenous transcripts	Kras, Stat1	Mecp2	Dystrophin gene, OTC gene	Ppib	Ppib, Nfkb, Nras, B4galnt1, Smarca4	Dystrophin gene	STAT1, SERPINA1	Idua, Tp53, Col3a1, Mbpr2, Ahi1, Fancc, Mybpc3, Il2rg	Idua
Concerns for immunogenicity	No	Yes	No	Yes	No	Yes	No	No	No

in vitro repair of a nonsense mutation within the open reading frame of the gene encoding the enhanced cyan fluorescent protein (eCFP). A 17-nucleotide long gRNA was designed with an A-to-C mismatch at the editing site to improve editing efficiency, as previously reported (Herbert and Rich 2001). Successful editing was determined by both fluorescence readout and sequencing of the target mRNA, which showed an efficiency of 60–90% (Stafforst and Schneider 2012). Based on these observations, a potential therapeutic application of this system to correct point mutations or to introduce specific modifications at the RNA level was suggested.

The guides used in this system were further optimised in order to improve specificity and efficiency in view of future applications in an in vivo setting (Vogel et al. 2014). Here, the SNAP-ADAR fusion protein can be genetically encoded to edit endogenous targets, while the gRNA carrying the chemical tag BG needs to be administered exogenously. Therefore, modifications of the gRNA are necessary in order to prevent its degradation by cellular RNases. For this purpose, the efficiency of chemically modified guides to induce editing was tested in vitro and in cells. An editing rate of 80% was obtained in vitro by using a gRNA carrying 2′-methoxy groups on all its ribonucleotides but the three centred around the targeted adenosine, two phosphothioate modifications at the 5′-terminus and four at the 3′-terminus. The modified gRNA was then proved to be highly efficient in a cellular environment, where instead unmodified gRNAs gave low editing yields, thus confirming the beneficial effect of the chemical modifications. Among the advantages of employing modified gRNAs, there is also an increase in the editing selectivity, especially on difficult targets such as adenosine-rich transcripts. The clinical relevance of this finding was demonstrated by the 70%-efficient in vitro editing of a disease-causing missense point mutation in the blood coagulation Factor 5 mRNA, the most common genetic risk factor for heritable multifactorial thrombophilia in the Caucasian population (Vogel et al. 2014).

A thorough analysis of how gRNAs can be modified to optimise the efficiency and selectivity of targeted RNA editing was performed (Schneider et al. 2014). Here, editing by ADAR1 and ADAR2 deaminase domains was systematically studied at four codons by using gRNAs carrying either a perfect match with uracil or a mismatch with cytosine opposite the targeted adenosine. In addition, the 5-′-neighbouring base next to the targeted site was matched or mismatched in all the possible combinations with the opposite base in the gRNA. Different codons were shown to have specific preferences in terms of both deaminase domain and gRNA structure; therefore optimisation is required for each individual desired target. Furthermore, it was demonstrated that introducing a guanosine mismatch opposite sites where off-target editing is observed provides a reliable strategy to decrease or even abolish unspecific editing (Schneider et al. 2014).

The SNAP-ADAR system was further optimised and proved to be highly efficient in editing endogenous transcripts, both by targeting them individually and simultaneously (Vogel et al. 2018). The possibility to apply targeted RNA editing to regulate activation of signalling pathways was demonstrated by editing two mRNAs coding for the signalling proteins KRAS and STAT1. Editing yields of 46–76% were achieved with the hyperactive version of SNAP-ADAR1 and no

off-target editing was detected along the gRNA-mRNA duplex thanks to the previously optimised gRNA carrying chemical modifications. In addition, global off-target editing was reduced by genomic integration of SNAP-ADARs under control of the doxycycline-inducible CMV promoter. Variation of induction times showed that off-target editing decreased with SNAP-ADAR expression level (Vogel et al. 2018).

3.1.2 λN-ADAR

The λN-ADAR system is based on the λ-phage N protein and its natural ability to recognise and bind boxB hairpin structures on RNA (Montiel-Gonzalez et al. 2013). Here, the deaminase domain of ADAR2 is fused to the λ-phage N peptide, and the specific guide RNA complementary to the target sequence of interest carries the boxB structure. Upon co-expression of λN-ADAR and the boxB-gRNA within a cell, the λN protein can interact with the boxB structure through non-covalent interactions, thus allowing the recruitment of the editing enzyme to the target RNA sequence, as shown in Fig. 2b.

It was demonstrated that the editing activity of λN-ADAR is maintained in vitro and that targeted editing can be achieved by expressing the enzyme together with an antisense oligonucleotide complementary to the sequence of interest containing boxB structures. Extensive editing can be induced by coupling λN-ADAR with a long gRNA carrying the λN-ADAR-recruiting boxB structures and complementary to the target mRNA coding for a potassium channel. Analysis of the different editing rates of the adenosine residues along the target led to the hypothesis that a more specific editing could be achieved by shortening the gRNA.

In order to show that targeted RNA editing can be employed to correct disease-relevant genetic mutations in vitro, a point mutation in the cystic fibrosis transmembrane conductance regulator (CFTR) associated with terminal cystic fibrosis was targeted. Upon optimisation of length and sequence of the gRNA and of the ratio between the three components, almost complete editing was achieved at the targeted site without off-target events. As last proof of the potential use of targeted RNA editing to correct mutations in cells, the λN-ADAR system was applied first in *Xenopus* oocytes and then in human cells. In *Xenopus* oocytes, correction of the above-mentioned mutation in CFTR mRNA was observed at both RNA and protein level with an efficiency of 20%. In HEK293T cells, instead, the λN-ADAR system was used to rescue a mutated eGFP (enhanced green fluorescent protein) reporter carrying a premature stop codon. Successful editing was detected as fluorescence signal with an efficiency of 20%. This study paved the way for further applications of the λN-ADAR system to correct G-to-A genetic mutations without directly modifying the genetic information at the DNA level.

Improvements in the λN-ADAR system were achieved by examining different parameters involved in the targeted editing mechanism (Montiel-Gonźalez et al. 2016). The structure of the gRNA was optimised by changing the position of the base facing the targeted adenosine and by modifying its distance from the boxB

structure. As previously reported by others (Herbert and Rich 2001; Wong et al. 2001; Stafforst and Schneider 2012; Schneider et al. 2014), an A-to-C mismatch was introduced at the target site in order to increase the editing efficiency. Editing efficiency in cells was evaluated with an efficient system based on two fluorescent reporters (mCherry and eGFP) expressed under the same promoter, with a nonsense mutation in the downstream gene. Cells transfected with this construct expressed only the first reporter in absence of editing and both reporters when editing occurred. In this way, accurate assessment of successful editing was possible; however, the overall efficiency of the system was low. The introduction of multiple λN peptides fused to ADAR deaminase domain was sufficient to improve the editing efficiency. Further improvement was obtained by adding more boxB hairpins in the gRNA structure. The use of the previously described hyperactive mutant ($ADAR2^{E488Q}$) of the editing enzyme also helped increase the editing efficiency (Kuttan and Bass 2012), but it also enhanced off-target editing. This effect was reduced, but not completely abolished, by decreasing the amount of gRNA transfected.

Since off-target events remained an unsolved problem for targeted RNA editing, further ways to optimise the λN-ADAR system were tested in order to limit the unknown and potentially detrimental effects of unspecific editing (Vallecillo-Viejo et al. 2018). It was demonstrated that nuclear localization of the editing enzyme reduced off-target editing, while maintaining on-target editing efficiency. Although this work was an important first step towards the solution of this problem, additional optimisation strategies are needed to further reduce off-target editing.

Site directed RNA editing based on the λN-ADAR system has been successfully employed to correct point mutations causing the debilitating neurological disorder called Rett syndrome (RTT) (Sinnamon et al. 2017). RTT is associated with sporadic mutations in the gene encoding the transcription factor methyl CpG binding protein 2 (MECP2), located on the X chromosome. As expected for an X-linked disorder, the consequences of carrying such mutations differ among female and male patients, with males being much more severely affected than females.

The most common and severe G-to-A mutation within the *MeCP2* gene (R106Q) was targeted. The R106Q mutation is located in the DNA binding domain of MeCP2 protein and affects its stability and ability to bind to chromatin. The λN-ADAR system was optimised in order to induce targeted RNA editing on an endogenous target. Neurons isolated from mice carrying the R106Q mutation in the *Mecp2* gene were transduced with adeno-associated viruses (AAV). The AAV vector expressed the recombinant $hADAR2^{E488Q}$-λN protein under the human synapsin I promoter and fused to three copies of the Simian virus 40 large T antigen nuclear localisation signal (NLS). The same AAV vector expressed also six copies of the *Mecp2* targeting guide under control of the human U6 small nuclear RNA polymerase III promoter. The *Mecp2* gRNA carried two boxB structures for recognition by the λN peptide and it was previously optimised by the addition of an A-to-C mismatch at the site of editing to increase editing efficiency (Herbert and Rich 2001). Moreover, an A-to-G mismatch was introduced 3 nucleotides upstream of the target adenosine in order to reduce the previously observed off-target editing at this site (Schneider et al. 2014). The *Mecp2* mRNA was edited with a 72 ± 5% efficiency and both the

MeCP2 protein level and its function were restored to a level similar to wild-type MeCP2. Off-target editing events were observed within the *Mecp2* cDNA, mainly within the sequence complementary to the gRNA, but none of them was of pathological concern (Sinnamon et al. 2017).

This approach was then applied in vivo in order to prove its efficiency not only in cell cultures but also in complex tissues (Sinnamon et al. 2020). The hippocampus of young mice carrying the human mutation R106Q in the mouse *Mecp2* gene was injected with AAV expressing the hyperactive catalytic domain of ADAR2 and the specific guide for *Mecp2*. First, expression of the editing enzyme was observed in the major cell populations of the hippocampus. Then, RNA was isolated both from the intact hippocampus and from three different subpopulations of hippocampal neurons and the editing level was determined. 35 ± 7% of on-target editing in the *Mecp2* mRNA from the intact hippocampus was observed, and around 50% of editing in the three individual neuronal populations was detected. Off-target editing events were also identified and whole-transcriptome RNA analysis showed that the majority of them occurred at rates ≤30%. As observed in vitro, the presence of an A-to-G mismatch decreased the rate of off-target editing at the corresponding site. On-target editing was guide specific and independent of the level of the editing enzyme, while off-target editing increased with the expression of ADAR. At the protein level, the localisation of MeCP2 to heterochromatin was restored at 50% of the wild-type value, as seen by confocal microscopy (Sinnamon et al. 2020).

The main strength of this approach is the successful editing of an endogenous transcript in vivo, which makes this system appealing for future therapeutic applications where the correction of a point mutation to rescue a cellular defect is desired. The administration of an exogenous editing enzyme might raise some concern, since the off-target editing rate increases with ADAR expression. However, the peripheral virus delivery used in this study to infect the whole brain results in a low viral titer, and therefore ADAR level, per cell, thus limiting the off-target editing events. An advantage of recruiting a recombinant, hyperactive deaminase domain of ADAR instead of the endogenous protein is the possibility to target sites located in sequence contexts that would not be naturally targeted. However, further studies to increase editing efficiency and specificity are needed. Moreover, in the case of Rett syndrome, understanding the level of editing that is necessary to fully reverse the pathological phenotype is necessary.

3.1.3 GluR2-ADAR

The GluR2-ADAR system is based on the pre-mRNA encoding subunits of the GluR2 receptor containing a highly edited, specific, hairpin-shaped motif, the R/G editing site, that is recognised by the double-stranded RNA-binding motifs (dsRBMs) of ADAR2 (Stefl et al. 2010). Due to the strong nature of this recognition by the dsRBDs of the full-length protein, the R/G motif can be used to recruit ADAR2 to a specific target in a way that mimics the natural editing process (Fig. 2c).

Guide RNAs containing an R/G motif and an mRNA binding sequence complementary to the target transcript except for a central A-to-C mismatch at the targeted site have been engineered to induce ADAR2 recruitment and A-to-I RNA-editing redirection (Wettengel et al. 2017). This approach was able to harness human ADAR2 and to successfully induce up to 10% editing of the Parkinson's associated gene PINK1 with a W437Stop mutation in HeLa cells. This was possible by co-transfecting an ADAR2 expressing construct and a specific gRNA to rescue the mutation. This system did not manage to recruit endogenously expressed ADAR2.

Recently, this approach has been optimised by overexpressing the hyperactive ADAR2 mutant (E488Q) together with a targeting ADAR guide RNA (adRNA), which contained the R/G motif as ADAR recruiting signal (Katrekar et al. 2019). This optimisation resulted in an increase of up to 40% of on-target editing rates of endogenous transcripts in HEK293T cells. Furthermore, the whole ADAR2 protein was packaged together with the GluR2-adRNA construct in a single adeno-associated virus vector 8 (AAV8). These AAV8-ADAR2-GluR2adRNA particles were administered to two experimental mouse models to demonstrate their potential in therapeutics. First, the *mdx* mouse model for Duchenne muscular dystrophy (DMD) was used, which carries a premature stop codon in exon 23 of the dystrophin gene. In order for it to be corrected, two consecutive adenosines must be edited (TAA to TGG), so a dual adRNA delivery approach was followed. The constructs were designed to express either wild-type or E488Q mutated ADAR2 (Kuttan and Bass 2012). Upon intramuscular administration of the AAV vector genomes, 1–2.5% protein restoration was achieved.

The second model was the spf^{ash} mouse model of ornithine transcarbamylase (OTC) deficiency. In this model a G-to-A point mutation in one of the OTC gene exons results in mRNA deficiency through incorrect splicing and in mutant protein production. After systemic delivery of AAV8-ADAR2-GluR2adRNA they observed 0.8–4.7% edited mRNA among the correctly spliced OTC mRNA. A further interesting observation was that the delivery of adRNA alone also resulted in low but significant RNA editing yields, suggesting the possibility of recruiting endogenous ADAR. The administration of the construct containing the hyperactive form of ADAR2 led to a higher editing efficiency (4.6–33.8%), as well as a reduction in the incorrectly spliced mRNA fraction. However, high off-target editing and toxicity in the animals were also reported, suggesting that the use of such mutants should be further studied and optimised, despite being already proven to yield higher editing efficiency (Katrekar et al. 2019).

This toolset as a whole requires further optimisation, but it shows potential in the field of RNA therapeutics mainly because it utilizes a naturally occurring ADAR recruiting signal for exogenous ADAR mediated RNA targeted editing.

By following a similar strategy, a first step towards leveraging endogenous ADAR2 in an in vitro approach has been taken (Fukuda et al. 2017) by using HEK293 cells in which native human ADAR2 could be overexpressed upon activation of a doxycycline-inducible promoter (Fukuda et al. 2012). Targeted A-to-I editing was achieved upon co-expression of ADAR guiding RNAs (AD-gRNAs) carrying the GluR2 hairpin motif for ADAR2 recruitment. Editing of a premature

stop codon in a reporter mRNA coding for GFP rescued the expression of the fluorescent protein with up to 3% efficiency, as assessed by fluorescence microscopy. These results indicate the possibility of the system to be further optimised towards recruitment of endogenous ADAR2, although additional studies are needed to achieve more efficient editing (Fukuda et al. 2017).

3.1.4 CRISPR/Cas

REPAIR

The CRISPR/Cas-based system depends on the protein dCas13b fused with the deaminase domain of hyperactive ADAR (Fig. 2d). dCas13b is a catalytically inactive mutant of PspCas13b, one of the so far known Cas13 enzymes. These proteins contain two higher-eukaryotes and prokaryotes nucleotide binding (HEPN) endoRNase domains that can cleave RNA molecules, preferably at sites flanked by protospacer sequences. It was demonstrated that the fusion protein of dCas13b with the deamination domain of ADAR2 ($ADAR2_{DD}$) can get recruited to a specific site for targeted editing when administered together with a specific gRNA (Cox et al. 2017). This specific gRNA contains a region complementary to the target RNA which defines the specificity of the system, an A-to-C mismatch at the targeted site to increase editing efficiency (Herbert and Rich 2001) and a direct repeat sequence which leads to the formation of a stem loop structure required for Cas13 recruitment (Abudayyeh et al. 2016).

First, successful editing was observed on a luciferase reporter, where the correction of a nonsense mutation through A-to-I editing could be detected as restoration of luminescence. Optimisation of the system led to the choice of dCas13b-$ADAR2_{DD}^{E488Q}$ as editing component and the development of REPAIRv1 (RNA Editing for Programmable A to I Replacement version 1). When applied to the endogenous transcript PPIB, REPAIRv1 showed an editing efficiency of 28%. After demonstrating that REPAIRv1 does not have any sequence constraints and that editing is observed within all target motifs, the ability of this system to correct two pathogenic G-to-A nonsense mutations among the 5769 reported in the ClinVar database (as of October 2020) was tested. The two selected mutations are located in two genes associated with X-linked nephrogenic diabetes insipidus (AVPR) and Fanconi anemia (FANCC). REPAIRv1 led to successful correction of the two mutations, with an editing efficiency of 35% for the AVPR gene and 23% for FANCC, respectively. When 34 other clinically relevant G-to-A mutations were tested, editing was achieved with up to 28% efficiency in 33 cases. In order to translate this system into the clinics, a therapeutically relevant administration strategy is needed, with AAV mediated delivery being one of the most suitable. In order to package dCas13b-$ADAR2_{DD}^{E488Q}$ into an AAV vector, its size, otherwise too large, was minimised by using a C-terminal truncated version of ADAR2, whose editing activity was maintained. Additionally, by rationally introducing mutations at different residues within the RNA duplex binding domain of $ADAR2_{DD}^{E488Q}$, a

more specific version of the system was developed, called REPAIRv2, which is characterised by a lower off-target editing rate combined with a still high on-target efficiency. Importantly, all the few off-target events induced by REPAIRv2 occurred in cancer-unrelated genes, thus dissolving any safety concern (Cox et al. 2017).

Rational mutagenesis at other residues might further improve efficiency and specificity of this system, as already demonstrated in a more recent study (Abudayyeh et al. 2019). Here, ADAR2 editing activity was expanded to cytosine deamination by rationally mutating key residues in the deaminase domain interacting with the RNA substrate. Further optimisation of the system led to the development of RESCUE (RNA Editing for Specific C-to-U Exchange), a new base editing tool to precisely introduce cytidine-to-uridine modifications at the RNA level.

Further engineering and optimisation of REPAIRv2 may offer additional advantages. By exploiting Cas13b pre-CRISPR-RNA processing activity, one could imagine designing a multiple guide RNA that, once cleaved in single guides, might allow multiplex editing at different sites. However, it has been shown that an R1079A mutation in Cas13 protein severely reduces its pre-crRNA processing activity (East-Seletsky et al. 2016). This mutation is located within the C-terminal Δ984–1090 truncation of dCas13b-ADAR$_{DD}$ that was introduced in order to fit the fusion protein within the packaging limit of AAV vectors (Cox et al. 2017). Therefore, multiplex targeted RNA editing might be achieved by administering a multiple gRNA together with the full-length dCas13b-ADAR$_{DD}$ fusion protein, but an alternative delivery strategy is needed to replace the AAV mediated delivery because of its size limitation. Another option to simultaneously allow multiplex targeted RNA editing and AAV delivery would be the identification of another truncated version of dCas13b-ADAR$_{DD}$ with a suitable size for AAV packaging and unaltered editing and pre-CRISPR-RNA processing activities.

On the other side, the main disadvantage of CRISPR-Cas-based approaches is the bacterial origin of CRISPR-Cas proteins and the fact that majority of the human population have circulating antibodies against them (Crudele and Chamberlain 2018). In this context, also the first administration of such proteins for therapies might be problematic and cause immune reactions.

CIRTS

In order to overcome this problem, a fully human-derived ribonucleoprotein complex called CIRTS (CRISPR-Cas-inspired RNA targeting system) was developed (Rauch et al. 2019). As illustrated in Fig. 2e, CIRTS is based on four components: an RNA-hairpin-binding protein that can specifically bind to engineered gRNA with high-affinity; a guide RNA with a complementary sequence to the target RNA and a specific structure that interacts with the RNA-hairpin-binding protein; a protein that can bind to the gRNA to stabilize and protect it before it interacts with the target; and an epitranscriptomic regulator that acts on the targeted RNA. The gRNA directs the RNA effector protein to a specific site in the transcriptome, where it can exert its function. This system offers an easily programmable, versatile platform for

modifying the epitranscriptome because it can deliver different effector proteins acting on RNA, among which RNA editing proteins such as ADARs. The possibility to direct the catalytic domain of ADAR2, both in the wild-type and hyperactive form, to edit an exogenous luciferase reporter was demonstrated. It has been shown that CIRTS can also deliver effector proteins to endogenous mRNAs, which would be ideal for a potential therapeutic application of targeted RNA base-editing. Another advantage of this system is the small size of CIRTS (432 aa), which allows AAV-packaging and delivery.

3.1.5 MS2-MCP-ADAR

The bacteriophage MS2-MCP tagging system has also been used for recruitment of ADAR. MS2 loops naturally occur in the phage genome and the MS2 coat binding protein (MCP) can recognise these structures and bind to them. adRNAs that contain MS2 loops can be paired with mRNA binding sequences and if MCP in turn is fused with ADAR deaminase domain, it can be utilised to be attracted to those loops and thus to the target mRNA, as shown in Fig. 2f. The potential of this system was examined by engineering adRNAs with a 20-nucleotide guide sequence with a C mismatch on the sixth nucleotide and with a pair of MS2 loops on either side of the guide (Katrekar et al. 2019). The adRNA was optimised to recruit synthetic fusion proteins made of MCP and the deaminase domain of ADAR1 or ADAR2, in the wild-type or hyperactive form. To evaluate the editing efficiency of the engineered constructs on different endogenous transcripts of HEK293T cells, Sanger and RNA sequencing analyses were performed. Both the constructs expressing $ADAR1_{DD}$ and $ADAR2_{DD}$ resulted in good editing rates ranging from 10 to 80%. Versions of the construct that contained nuclear export signal (NES) or expressed the hyperactive deaminase domains induced higher RNA editing at the target adenosine, but also led to a considerably higher off-target editing (Katrekar et al. 2019). Instead, lower off-target editing was achieved with constructs carrying a nuclear localisation signal (NLS), as previously observed with another system (Vallecillo-Viejo et al. 2018).

This system was applied to the already mentioned *mdx* mouse model, where the highest editing efficiency of almost 3% was reported 8 weeks after injection of AAV8 vectors expressing the MS2 gRNA with MCP-$ADAR1_{DD}$(hyperactive mutant)-NLS (Katrekar et al. 2019).

More recently, the challenge of the ratio of each component when it comes to delivery to actual patients was addressed, since in most existing protocols the amount of each component is more likely not compatible with potential clinical applications. The proposal involves the development of a system that introduces all the components (genes encoding gRNA, fusion MCP-$ADAR_{DD}$ and target genes) into a single plasmid vector construct (Tohama et al. 2020). An additional difference compared to the previous approach is the replacement of the conventional linker for the fusion of MCP and $ADAR1_{DD}$ with a flexible linker XTEN that contributes to protein stability (Chhabra et al. 2015). At the same time, the size of the MCP was reduced so as to limit the size of the fusion protein making it more applicable to gene

therapy. These single construct plasmids were transfected into HEK293T cells and the editing efficiency was measured both on a reporter gene and, more importantly, on endogenous genes and was found to be up to 40% (Tohama et al. 2020).

This is a promising finding in regard to a new approach to this system with possible applications in actual gene therapy. Further investigation into possible immunogenicity and off-target editing effects should be done.

3.2 Recruitment of Endogenous ADARs

All the above-mentioned approaches rely on the delivery of both the editing enzyme and the gRNA, which makes their application in the clinic more challenging. A step forward for the development of therapies based on RNA base-editing technologies has been achieved with the design of guides able to recruit endogenous ADARs.

3.2.1 RESTORE

A method for RNA editing called RESTORE (Recruiting Endogenous ADAR to Specific Transcripts for oligonucleotide-mediated RNA Editing) has been recently presented, where endogenous ADARs are recruited by delivering chemosynthetic antisense oligonucleotides (Merkle et al. 2019). The chemically modified antisense oligonucleotides (ASOs) comprise a programmable specificity domain complementary to the target mRNA and an invariant ADAR recruiting domain carrying the GluR2 motif to attract endogenous human ADAR to the ASO-mRNA hybrid for editing (Fig. 2c). The system was optimised by introducing chemical modifications along the ASO sequence, specifically phosphorothioate on 4 terminal residues at the 3′ end and 2′-O-methylations at all but three residues opposite the nucleotide triplet being targeted, as previously reported (Vogel et al. 2014), leading to higher editing efficiency. In addition, it was demonstrated that multiple transcripts can be targeted simultaneously by cotransfection of more ASOs, as already achieved with other systems (Vogel et al. 2018). Targeting the 3′ UTR of endogenous GAPDH yielded 5–35% editing, which was boosted by 1.5–2-fold upon addition of IFNα thanks to an increase in ADAR1 p150 expression. When switching the target within the ORF of GAPDH, editing was initially not observed. Optimisation of the system by lengthening the specificity domain of the ASO to 40 nucleotides and including locked nucleic acid modifications in the antisense domain finally led to successful editing with a 42.7 ± 1.5% efficiency.

In order to demonstrate the potential of a therapeutical application, RESTORE was applied to different primary cells. Editing of the phosphotyrosine 701 site of STAT1 was induced with an efficiency of 3–30% in primary fibroblasts, thus demonstrating the possibility to trigger the activation of a signalling factor via targeted RNA editing. Moreover, the PiZZ mutation (E342K) in SERPINA1, the most clinically relevant mutation causing α1-antitrypsin (A1AT) deficiency, was

successfully repaired with a 10–20% efficiency. By overcoming the need of overexpressing exogenous ADAR, this approach offers the advantage of minimizing the risk posed by off-target editing. In addition, since only a strategy to administer oligonucleotides is necessary for the application of this system, more delivery options are available, some of which are already successfully employed in clinical trials.

3.2.2 LEAPER

An alternative system that was developed for targeted RNA base-editing by endogenous ADAR is the LEAPER system (leveraging endogenous ADAR for programmable editing of RNA). Long, linear RNAs are engineered so they can recruit endogenous ADAR proteins to target transcripts and perform A-to-I editing precisely (Fig. 2g), giving rise to new potential for use of RNA editing in therapy. By using the REPAIR system (Cox et al. 2017), it has been reported that editing can be observed even in absence of dCas13a-$ADAR_{DD}$ fusion proteins and the Cas13a-recruiting scaffolds in the crRNA. Hence, ADAR recruiting RNAs (arRNAs) were designed and optimised for high editing efficiency (Qu et al. 2019). The mechanism is described as follows and is similar to others described previously. The arRNA is a long (71–151 nt) guide RNA that is designed with an almost complementary sequence to its target transcript except for a specific mismatch that introduces a C on the arRNA mispairing with the targeted A of the transcript (Schneider et al. 2014).

The LEAPER system was shown to be effective in recruiting ADAR1, while having minimal impact on the natural cell A-to-I editing and without showing RNA interference effects on endogenous transcripts. Furthermore, no immunogenicity or type I interferon pathways activation were observed. These findings suggest that LEAPER is safe for use in mammalian cells and therefore can be a promising therapeutic tool. A current disadvantage here, compared to RESTORE, is that the arRNA, having no further modifications, is not chemically protected from the environment in the cytoplasm and therefore it might be degraded upon delivery.

To assess the overall specificity of the construct, transcriptome-wide RNA sequencing was performed to identify off-target editing in predicted sites, upon transfection with arRNAs targeting the *PPIB* gene. This showed that LEAPER could retain transcriptome-wide specificity without off target edits.

The system was then tested in human primary cells in vitro. After co-transfection of an eGFP reporter gene and 151-nt arRNA for targeted editing in both human primary pulmonary fibroblasts and human primary bronchial epithelial cells, up to 45% of eGFP positive cells could be obtained. Additionally, editing of the endogenous gene *PPIB* was assessed in these two cell lines, as well as in a human primary T cell line, using a 151-nt arRNA targeting the specific transcript. The results showed 30–80% editing rates depending on the cell type and thus were deemed encouraging for further optimisation of the construct towards therapeutics.

Experimental approaches on clinically relevant targets showed that the system is capable of restoring protein function. Editing was induced to repair the tumour suppressor gene TP53—the most frequently mutated gene in human cancers—with a c.158G-to-A non-sense mutation (W53Stop), which results in a non-functional p53 protein due to the introduction of a premature stop codon. The resulting precise editing of up to 35% of the transcript could rescue the production of the full-length protein. Additionally, LEAPER's potential in treating a monogenic disease such as Hurler syndrome, the most severe subtype of Mucopolysaccharidosis type I (MPS I), was demonstrated through the restoration of IDUA enzyme (α-L-iduronidase) function. IDUA, whose deficiency is the causative factor for Hurler syndrome, is a lysosomal metabolic enzyme that is responsible for the degradation of mucopolysaccharides. In primary fibroblasts isolated from a Hurler syndrome patient carrying a homozygous TGG-to-TAG mutation in exon 9 of the *IDUA* gene (GM06214 cell line), the resulting W402Stop change in the protein causes the complete loss of its function. Two different arRNAs, one targeting the mature mRNA and one targeting the pre-mRNA of IDUA, both managed to successfully induce editing with an efficiency of 10% and 30%, respectively, as assessed by RNA-sequencing analysis. Restoration of IDUA catalytic activity was also evaluated and compared with the activity of IDUA in GM01323 cells, another primary fibroblast cell line from a patient with Scheie syndrome, a milder subtype of MPS-I than Hurler syndrome, which retains some activity of IDUA enzyme. Results showed that the catalytic activity of IDUA in GM06214 cells upon editing with the arRNA targeting IDUA pre-mRNA was higher than GM01323 cells. This gain of catalytic activity was consistent with the 30% efficiency of RNA editing on the transcript. A further useful observation was that editing was successfully induced without promoting expression of genes responsible for type-I interferon and pro-inflammatory responses (Qu et al. 2019).

It is evident that the LEAPER system is a very promising and versatile therapeutic tool that manages to repair both cancer-relevant premature codons and monogenic disease-related mutations. The advantages and possibilities are many, but there is still much to be investigated regarding the optimisation of editing efficiency and, most importantly, the implementation of an appropriate delivery system in the scope of successful therapy development. The delivery through more clinically relevant methods, such as the lentivirus-based expression, is already being tested. Other promising approaches for delivery may include plasmid, other viral vectors or direct delivery of synthetic oligonucleotides. Overall, the development of systems like LEAPER that rely on the exploitation of endogenous ADAR with the goal of targeted editing is a needed step towards the implementation of RNA-editing therapies.

4 Future Applications, Potentials and Limitations

The potential of RNA-based therapies has been broadly investigated in the past few decades, leading also to the rise of the more recent interest in the clinical application of RNA base-editing. Despite the development and availability of different targeted

RNA base-editing systems in research laboratories, none of them has made it into the clinic so far. The most advanced technology on the road to clinical trials was developed by the biotechnology company ProQR and is called Axiomer. This platform exploits the natural mechanism of RNA editing to correct pathogenic G-to-A point mutations at the RNA level: upon delivery of synthetic editing oligonucleotides (EONs), specifically designed to hybridize to the mutated mRNA, the resulting dsRNA duplex can recruit the endogenous enzyme ADAR and redirect its natural editing activity to repair the targeted mutation. The EONs are short single-stranded RNA molecules complementary to the mutated sequence of interest. Consistently with the literature and what we have presented so far, they also carry a C mismatch opposite the mutated site. The EONs are chemically modified to improve their efficacy, uptake and stability in the cells. Potentially, this approach could allow the correction of over 10,000 known G-to-A pathogenic mutations and provide a cure for the associated diseases, which are mostly untreatable.

To start with, ProQR is focusing on the development of an Axiomer-based therapy to correct the already mentioned most common mutation causing Hurler syndrome (W402X) (Qu et al. 2019). The preclinical studies are based on a mouse model for Hurler syndrome carrying the murine version of the human mutation (W392X). First, the efficacy of the system was proven by inducing targeted editing in mouse embryonic fibroblasts (MEF) transfected with an IDUA W392X reporter construct and treated with the specific EON. Up to 30% of editing was observed by Sanger sequencing and up to 40% restoration of iduronidase enzymatic activity was measured. Successful editing of endogenous IDUA was demonstrated also in mouse liver and lung fibroblasts and in hepatocytes isolated from W392X mice and transfected with the specific EON. Finally, the Axiomer platform was tested in vivo with two different delivery approaches. Intravenous (IV) injection of liposomal vectors carrying the specific EONs allowed to reach the liver of the W392X mouse model, where successful repair of the mutation was demonstrated by up to 40% increase of IDUA enzymatic activity (Aalto 2018). In the second approach, intravitreal (IVT) injection of naked EONs was used to correct the W392X mutation in the retina of the Hurler mouse model with an editing efficiency close to 10% (ProQR Therapeutics 2020). Treatment administration was different in the two approaches. IV injection was performed in 4 doses over 8 days, followed by assessment of the therapeutic outcome by IDUA enzymatic activity, while IVT injection was only performed once and the effect was determined after 7 days. Clearly, each delivery approach requires a specific administration protocol because of different efficiency, invasiveness and potential complications. Repetitive administration is not an issue for non-invasive strategies, such as IV injection, while it is a matter of concern for invasive delivery routes, such as IVT injection. The observation that a therapeutic effect was detectable one week after IVT delivery is promising, because it shows that naked EONs are stable for at least 7 days, excluding the need of a more frequent administration.

Further EON optimisation is currently ongoing. The establishment of a clinical proof of concept is the next step towards the launch of a clinical trial to translate this promising therapeutic approach into the clinic. The potential of the Axiomer platform is incomparable because of its versatility for numerous disparate genetic

diseases. Potentially, all 13,328 so far known G-to-A pathogenic mutations reported in the ClinVar database could be corrected. If successful in humans, it could change the paradigm of targeted therapy.

The therapeutic potential of RNA-base editing is currently held back by the challenging issue of delivery. The possibility to recruit endogenous ADARs limits the problem to one component of the editing system and highlights the need for RNA oligonucleotides delivery strategies. Major hurdles to the administration of RNA molecules are the different physiological mechanisms that take place inside the cells and in the extracellular space, which might lead to RNA degradation or sequestration, thus blocking its way to the targeted site (Roberts et al. 2020).

Several strategies to facilitate oligonucleotides delivery are available, with some being already approved and some others still being investigated. Improvements of RNA molecules pharmaceutical properties, such as pharmaco-kinetics, pharmaco-dynamics and biodistribution, have been achieved by addition of chemical modifications. By changing the chemistry of the RNA oligonucleotides, at either backbone, ribose sugar moiety or nucleobase level, delivery to different tissues can be enabled, even bypassing the need for an additional vector. Chemical modifications can increase stability, confer resistance to nucleases, allow internalization into the tissue of interest, facilitate interactions with intracellular proteins and thus accumulation at the site of action, modulate activity and efficiency and prevent immune responses that could rise upon formation of dsRNA structures in the cytoplasm (Roberts et al. 2020). However, it is always important to check that binding affinity for the target is not impaired by extensive modifications of the oligonucleotide.

Another strategy to facilitate the delivery of RNA oligonucleotides is the formation of bio-conjugates with other molecules such as lipids, peptides, antibodies, sugar moieties and aptamers. The different properties of such conjugates can allow to target specific cell types or tissues by promoting intracellular uptake upon interaction with specific cell surface receptors. The latest advancements in this field are provided by nanoparticles-mediated delivery systems, which offer the advantages, among others, of being highly optimizable for crossing biological barriers and ensuring target specificity (Roberts et al. 2020). For example, recently, RNA molecules have been successfully complexed on the surface of lipid nanoparticles and effectively delivered in vivo (Blakney et al. 2019). This is proposed as an alternative to encapsulating RNA within lipid nanoparticles, which is until now the delivery method of the only FDA-approved RNA-based therapy (Morrison 2018; Blakney et al. 2019).

As the interest in RNA editing-based therapy grows, finding delivery strategies to ensure safe, specific and effective administration of such potential drugs is of utmost importance. As optimisation of the editing efficiency is fundamental for the clinical application of the above-described systems, establishment of appropriate delivery technologies is crucial for their translation into the clinic. Significant advancements in both fields will allow new therapeutic opportunities to move forwards and give hope for the treatment of many currently incurable diseases.

References

Aalto A (2018) Axiomer technology–Therapeutic oligonucleotides for directing site-specific A-to-I editing by endogenous ADAR enzymes. https://www.proqr.com/files/2020-08/Axiomer-technology-TIDES-2018.pdf

Abudayyeh OO, Gootenberg JS, Franklin B et al (2019) A cytosine deaminase for programmable single-base RNA editing. Science 365:382–386

Abudayyeh OO, Gootenberg JS, Konermann S et al (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. In: Science 353(6299):aaf5573

Bajad P, Jantsch MF, Keegan L et al (2017) A to I editing in disease is not fake news. RNA Biol 14:1223–1231

Basilio C, Wahba AJ, Lengyel P et al (1962) Synthetic polynucleotides and the amino acid code. V. Proc Natl Acad Sci U S A 48:613–616

Bass BL, Nishikura K, Keller W et al (1997) A standardized nomenclature for adenosine deaminases that act on RNA. RNA 3:947–949

Benne R, Van Den Burg J, Brakenhoff JPJ et al (1986) Major transcript of the frameshifted coxll gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell 46:819–826

Blakney AK, McKay PF, Yus BI et al (2019) Inside out: optimization of lipid nanoparticle formulations for exterior complexation and in vivo delivery of saRNA. Gene Ther 26:363–372

Borden EC (2019) Interferons α and β in cancer: therapeutic opportunities from new insights. Nat Rev Drug Discov 18:219–234

Charlesworth CT, Deshpande PS, Dever DP et al (2019) Identification of preexisting adaptive immunity to Cas9 proteins in humans. Nat Med 25:249–254

Chen CX, Cho DS, Wang Q et al (2000) A third member of the RNA-specific adenosine deaminase gene binding domains. RNA 6:755–767

Chhabra ES, Moore N, Furcht C et al (2015) Evaluation of enhanced in vitro plasma stability of a novel long acting recombinant FVIIIFc-VWF-XTEN fusion protein. Blood 126:2279–2279

Cho SW, Kim S, Kim Y et al (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res 24:132–141

Cho DSC, Yang W, Lee JT et al (2003) Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. J Biol Chem 278:17093–17102

Chung H, Calis JJA, Wu X et al (2018) Human ADAR1 prevents endogenous RNA from triggering translational shutdown article human ADAR1 prevents endogenous RNA from triggering translational shutdown. Cell 172:811–824.e14

Cox DBT, Gootenberg JS, Abudayyeh OO et al (2017) RNA editing with CRISPR-Cas13. Science (80-) 358:1019–1027

Crudele JM, Chamberlain JS (2018) Cas9 immunity creates challenges for CRISPR gene editing therapies. Nat Commun 9:9–11

East-Seletsky A, O'Connell MR, Knight SC et al (2016) Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature 538:270–273

Egebjerg J, Heinemann SF (1993) Ca2+ permeability of unedited and edited versions of the kainate selective glutamate receptor GluR6. Proc Natl Acad Sci U S A 90:755–759

Fu Y, Foden JA, Khayter C et al (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31:822–826

Fukuda M, Kurihara K, Tanaka Y et al (2012) A strategy for developing a hammerhead ribozyme for selective RNA cleavage depending on substitutional RNA editing. RNA 18:1735–1744

Fukuda M, Umeno H, Nose K et al (2017) Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. Sci Rep 7:8–19

Gabriel R, Lombardo A, Arens A et al (2011) An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol 29:816–823

Gallo A, Vukic D, Michalík D et al (2017) ADAR RNA editing in human disease; more to it than meets the I. Hum Genet 136:1265–1278

Gaudelli NM, Komor AC, Rees HA et al (2017) Programmable base editing of a•T to G•C in genomic DNA without DNA cleavage. Nature 551:464–471
Gerber AP, Keller W (2001) RNA editing by base deamination: more enzymes, more targets, new mysteries. Trends Biochem Sci 26:376–384
Hacein-Bey-Abina S, Garrigue A, Wang GP et al (2008) Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1 find the latest version : insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest 118:3132–3142
Helwak A, Tollervey D (2014) Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). Nat Protoc 9:711–728
Herbert A, Rich A (2001) The role of binding domains for dsRNA and Z-DNA in the in vivo editing of minimal substrates by ADAR1. Proc Natl Acad Sci U S A 98:12132–12137
Higuchi M, Maas S, Single FN et al (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. Nature 406:78–81
Katrekar D, Chen G, Meluzzi D et al (2019) In vivo RNA editing of point mutations via RNA-guided adenosine deaminases. Nat Methods 16:239–242
Keegan LP, Gallo A, O'Connell MA (2001) The many roles of an RNA editor. Nat Rev Genet 2:869–878
Keppler A, Gendreizig S, Gronemeyer T et al (2003) A general method for the covalent labeling of fusion proteins with small molecules in vivo. Nat Biotechnol 21:86–89
Kim S, Lee MJ, Kim H et al (2011) Preassembled zinc-finger arrays for rapid construction of ZFNs. Nat Methods 8:7
Kuttan A, Bass BL (2012) Mechanistic insights into editing-site specificity of ADARs. Proc Natl Acad Sci U S A 109:E3295–E3304
Lamers MM, van den Hoogen BG, Haagmans BL (2019) ADAR1: "editor-in-chief" of cytoplasmic innate immunity. Front Immunol 10:1763
Lehmann KA, Bass BL (1999) The importance of internal loops within RNA substrates of ADAR1. J Mol Biol 291:1–13
Li H, Yang Y, Hong W et al (2020) Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. Signal Transduct Target Ther 5:1
Liddicoat BJ, Piskol R, Chalk AM et al (2015) RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. Science 349:1115–1120
Lundstrom K (2018) Viral vectors in gene therapy. Diseases 6(2):42
Maas S, Gommans WM (2009) Identification of a selective nuclear import signal in adenosine deaminases acting on RNA. Nucleic Acids Res 37:5822–5829
Maas S, Patt S, Schrey M et al (2001) Underediting of glutamate receptor Glur-B mRNA in malignant gliomas. Proc Natl Acad Sci U S A 98:14687–14692
Merkle T, Merz S, Reautschnig P et al (2019) Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. Nat Biotechnol 37:133–138
Miller JC, Tan S, Qiao G et al (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–150
Mladenova D, Barry G, Konen LM et al (2018) Adar3 is involved in learning and memory in mice. Front Neurosci 12:1–17
Montiel-Gonźalez MF, Vallecillo-Viejo IC, Rosenthal JJC (2016) An efficient system for selectively altering genetic information within mRNAs. Nucleic Acids Res 44:1–12
Montiel-Gonzalez MF, Vallecillo-Viejo I, Yudowski GA et al (2013) Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. Proc Natl Acad Sci U S A 110:18285–18290
Morrison C (2018) Alnylam prepares to land first RNAi drug approval. Nat Rev Drug Discov 17:156–157
Mout R, Ray M, Lee YW et al (2017) In vivo delivery of CRISPR/Cas9 for therapeutic gene editing: Progress and challenges. Bioconjug Chem 28:880–884

Nishikura K (2010) Funtions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem 79:321–349
Nishikura K (2016) A-to-I editing of coding and non-coding RNAs by ADARs. Nat Rev Mol Cell Biol 17:83–96
Nóbrega C, Mendonça L, Matos CA (2020) Non-viral vectors for gene therapy. In: A handbook of gene and Cell therapy. Springer International Publishing, Cham, pp 23–37
Oakes E, Anderson A, Cohen-Gadol A et al (2017) Adenosine deaminase that acts on RNA 3 (adar3) binding to glutamate receptor subunit B pre-mRNA inhibits RNA editing in glioblastoma. J Biol Chem 292:4326–4335
Pattanayak V, Lin S, Guilinger JP et al (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol 31:839–843
Patterson JB, Samuel CE (1995) Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. Mol Cell Biol 15:5376–5388
Pestal K, Funk CC, Snyder JM et al (2015) Isoforms of the RNA editing enzyme ADAR1 independently control nucleic acid sensor MDA5-driven autoimmunity and multi-organ development. Immunity 43:933–944
ProQR Therapeutics (2020) Axiomer technology–Therapeutic oligonucleotides for directing site-specific A-to-I editing by endogenous ADAR enzymes. https://www.proqr.com/files/2020-08/Axiomer public overview 2020.pdf
Qu L, Yi Z, Zhu S et al (2019) Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs. Nat Biotechnol 37:1059–1069
Rauch S, He E, Srienc M et al (2019) Programmable RNA-guided RNA effector proteins built from human parts. Cell 178:122–134.e12
Rice GI, Kasher PR, Forte GMA et al (2012) Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. Nat Genet 44:1243–1248
Roberts TC, Langer R, Wood MJA (2020) Advances in oligonucleotide drug delivery. Nat Rev Drug Discov 19(10):673–694
Rueter SM, Dawson TR, Emeson RB (1999) Regulation of alternative splicing by RNA editing. Nature 399:75–80
Samuel CE (2019) Adenosine deaminase acting on RNA (ADAR1), a suppressor of double-stranded RNA–triggered innate immune responses. J Biol Chem 294:1710–1720
Savva YA, Rieder LE, Reenan RA (2012) The ADAR protein family. Genome Biol 13:252
Schneider MF, Wettengel J, Hoffmann PC et al (2014) Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. Nucleic Acids Res 42:1–9
Seeburg PH, Higuchi M, Sprengel R (1998) RNA editing of brain glutamate receptor channels: mechanism and physiology. Brain Res Rev 26:217–229
Sinnamon JR, Kim SY, Corson GM et al (2017) Site-directed RNA repair of endogenous Mecp2 RNA in neurons. Proc Natl Acad Sci U S A 114:E9395–E9402
Sinnamon JR, Kim SY, Fisk JR et al (2020) In vivo repair of a protein underlying a neurological disorder by programmable RNA editing. Cell Rep 32:107878
Sommer B, Köhler M, Sprengel R et al (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. Cell 67:11–19
Song Y, Yang W, Fu Q et al (2020) irCLASH reveals RNA substrates recognized by human ADARs. Nat Struct Mol Biol 27:351–362
Stafforst T, Schneider MF (2012) An RNA-deaminase conjugate selectively repairs point mutations. Angew Chemie–Int Ed 51:11166–11169
Stefl R, Oberstrass FC, Hood JL et al (2010) The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific read out of the RNA minor groove. Cell 143:225–237
Strehblow A, Hallegger M, Jantsch MF (2002) Nucleocytoplasmic distribution of human RNA-editing enzyme ADAR1 is modulated by double- stranded RNA-binding domains, a leucine-rich export signal, and a putative dimerization domain. Mol Biol Cell 13:3822–3835

Tang R, Xu Z (2020) Gene therapy: a double-edged sword with great powers. Mol Cell Biochem 474:73–81

Tohama T, Sakari M, Tsukahara T (2020) Development of a single construct system for site-directed rna editing using MS2-ADAR. Int J Mol Sci 21:1–11

Twomey EC, Yelshanskaya MV, Grassucci R et al (2017) Channel opening and gating mechanism in AMPA-subtype glutamate receptors. Nature 549:60–65

Vallecillo-Viejo IC, Liscovitch-Brauer N, Montiel-Gonzalez MF et al (2018) Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme. RNA Biol 15:104–114

Vogel P, Moschref M, Li Q et al (2018) Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. Nat Methods 15:535–538

Vogel P, Schneider MF, Wettengel J et al (2014) Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA. Angew Chemie–Int Ed 53:6267–6271

Vogel P, Stafforst T (2019) Critical review on engineering deaminases for site-directed RNA editing. Curr Opin Biotechnol 55:74–80

Wang Q, Miyakoda M, Yang W et al (2004) Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. J Biol Chem 279:4952–4961

Wang F, Wang L, Zou X et al (2019) Advances in CRISPR-Cas systems for RNA targeting, tracking and editing. Biotechnol Adv 37:708–729

Wang C, Zou J, Ma X et al (2017) Mechanisms and implications of ADAR-mediated RNA editing in cancer. Cancer Lett 411:27–34

Ward SV, George CX, Welch MJ et al (2011) RNA editing enzyme adenosine deaminase is a restriction factor for controlling measles virus replication that also is required for embryogenesis. Proc Natl Acad Sci U S A 108:331–336

Wettengel J, Reautschnig P, Geisler S et al (2017) Harnessing human ADAR2 for RNA repair–recoding a PINK1 mutation rescues mitophagy. Nucleic Acids Res 45:2797–2808

Wong SK, Sato S, Lazinski DW (2001) Substrate recognition by ADAR1 and ADAR2. RNA 7:846–858

Woolf TM, Chase JM, Stinchcomb DT (1995) Toward the therapeutic editing of mutated RNA sequences. Proc Natl Acad Sci U S A 92:8298–8302

Xu L-D, Öhman M (2019) ADAR1 editing and its role in cancer. Genes (Basel) 10(1):12

Zetche B, Volz SE, Zhang F (2015) A split-Cas9 architecture for inducible genome editing and transcription modulation. Nat Biotechnol 33:139–142

Zhang XJ, He PP, Li M et al (2004) Seven novel mutations of the ADAR gene in Chinese families and sporadic patients with dyschromatosis symmetrica hereditaria (DSH). Hum Mutat 23:629–630

Zhang XH, Tee LY, Wang XG et al (2015) Off-target effects in CRISPR/Cas9-mediated genome engineering. Mol Ther–Nucleic Acids 4:e264

Mechanisms and Clinical Applications of RNA Pseudouridylation

Pedro Morais, Hironori Adachi, Jonathan L. Chen, and Yi-Tao Yu

Contents

Abstract Pseudouridylation is perhaps the most common epitranscriptomic modification among over 170 known chemical RNA modifications. Pseudouridine (Ψ) is highly conserved in various stable RNAs of all organisms. RNA pseudouridylation can be catalyzed by an RNA-independent mechanism by which stand-alone

P. Morais
ProQR Therapeutics, Leiden, Zuid-Holland, The Netherlands

H. Adachi · J. L. Chen · Y.-T. Yu (✉)
Department of Biochemistry and Biophysics, Center for RNA Biology, University of Rochester Medical Center, Rochester, NY, USA
e-mail: YiTao_Yu@URMC.rochester.edu

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_19

enzymes, known as pseudouridine synthases, recognize the substrate and catalyze the U-to-Ψ conversion reaction. Alternatively, pseudouridylation can be catalyzed by an RNA-guided mechanism, where a guide RNA (box H/ACA RNA), which is complexed with four core proteins (Cbf5/NAP57, Nhp2, Gar1, and Nop10), site-specifically directs the conversion of target uridine into a Ψ. Here, we discuss the underlying mechanisms of pseudouridylation as well as the methods for the detection of this modification. We also discuss pseudouridylation-linked diseases and potential clinical applications of this RNA modification.

Keywords Pseudouridine · box H/ACA RNP · Pseudouridine synthase · RNA modification · Epitranscriptome · Disease · Targeted pseudouridylation

1 Introduction

Pseudouridine (Ψ) is the most abundant RNA modification, and it is found in many different types of RNA (mRNA, rRNA, tRNA, snRNA, and other types of noncoding RNA), often in conserved positions. Pseudouridylation is a uridine-specific posttranscriptional RNA modification where a uridine residue is isomerized into a Ψ. The isomerization process initiates with the breakage of the N1-C1' bond followed by a 180° base rotation around the N3-C6 axis and formation of a rotatable C-C bond (C5-C1') (Fig. 1). This modified nucleotide has distinct chemical properties that can ultimately have an impact on RNA function and gene expression. In addition to the hydrogen bonding pattern of uridines, Ψ has an extra hydrogen bond donor group (N1H) in the major groove, which affects local RNA structure and increases base stacking and stability by promoting a C3'-*endo* conformation of the ribose moiety (Westhof 2019). Like uridine, Ψ can base-pair with adenosine

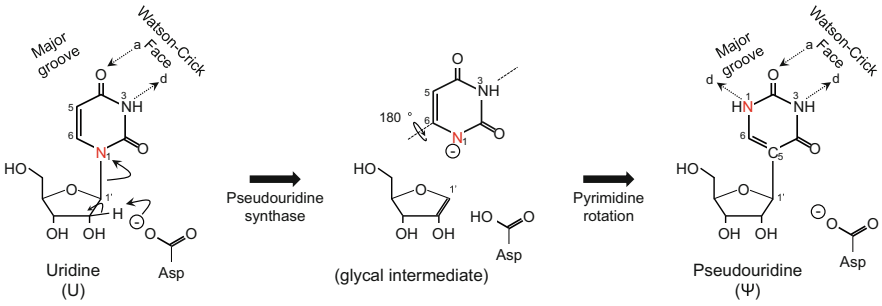


Fig. 1 Schematic representation of uridine (U)-to-pseudouridine (Ψ) isomerization—pseudouridylation. Ψ is a rotational isomer of uridine, in which the N-C glycosidic bond is broken to form the C-C bond (Veerareddygari et al. 2016). This reaction is catalyzed by pseudouridylase enzymes, containing an Asp at the active site (shown). When compared with uridine, Ψ has an extra hydrogen bond donor (d); however, the number of hydrogen bond acceptors (a) is unchanged

residues, but with greater thermodynamic stability in the context of a double helix (Kierzek et al. 2014). Thus, this modification can broaden the diversity of RNA structures and increase its potential for novel functions in gene expression processes such as pre-mRNA splicing and translation of proteins (Adachi et al. 2019a). Therefore, it is not surprising that Ψ is commonly found in RNA regions with functional importance. For instance, the decoding center of rRNA (Penzo and Montanaro 2018) and the branch site recognition regions of U2 snRNAs that base-pair with the branch site sequence of pre-mRNA (Yu et al. 2011) are all enriched with Ψ. Furthermore, pseudouridylation is not only constitutive, but it can also be induced in stress conditions. For instance, yeast U2 snRNA has, in addition to the three constitutive sites (Ψ35, Ψ42, Ψ44), at least two additional Ψ residues at positions 56 and 93 under stress conditions (heat shock or nutrient deprivation) (Wu et al. 2011). Given its unique chemical properties and its ability to influence RNA function, Ψ is believed to have potential clinical applications.

2 Mechanism of RNA Pseudouridylation

Targeted pseudouridylation can be catalyzed either by stand-alone PseudoUridine Synthases (Pus proteins) via an RNA-independent process or by box H/ACA RNPs (RNA-protein complexes) through an RNA-dependent process, where the RNA component of the RNPs directs site-specific pseudouridylation.

2.1 RNA-Dependent RNA Pseudouridylation

RNA-directed RNA pseudouridylation constitutes a major mechanism for targeted U-to-Ψ conversion, in which a small guide RNA known as box H/ACA RNA recognizes the substrate RNA through base-pairing and thus specifies the target uridine to be pseudouridylated (Fig. 2). Hence, this is an RNA-dependent mechanism (Yu and Meier 2014). Box H/ACA RNAs exist as ribonucleoprotein particles, each consisting of one unique box H/ACA RNA and four common core proteins called Nhp2, Gar1, Nop10, and dyskerin/NAP57 (in mammals) or Cbf5 (in yeast). Cbf5 is the catalytic component responsible for the U-to-Ψ chemical reaction. The box H/ACA RNAs each have a defined secondary structure, known as the hairpin-hinge (box H)-hairpin-tail (box ACA) structure. Thus, each box H/ACA RNA contains two independent hairpins. In each of the hairpins, there is an internal loop, known as the pseudouridylation pocket, which serves as a guide that base-pairs with the RNA substrate during modification. Notably, the base-pairing between the guide and the substrate RNA occurs in such a way that the target uridine is positioned at the base of the upper stem, left unpaired, and ready to be pseudouridylated by Cbf5 (Fig. 2).

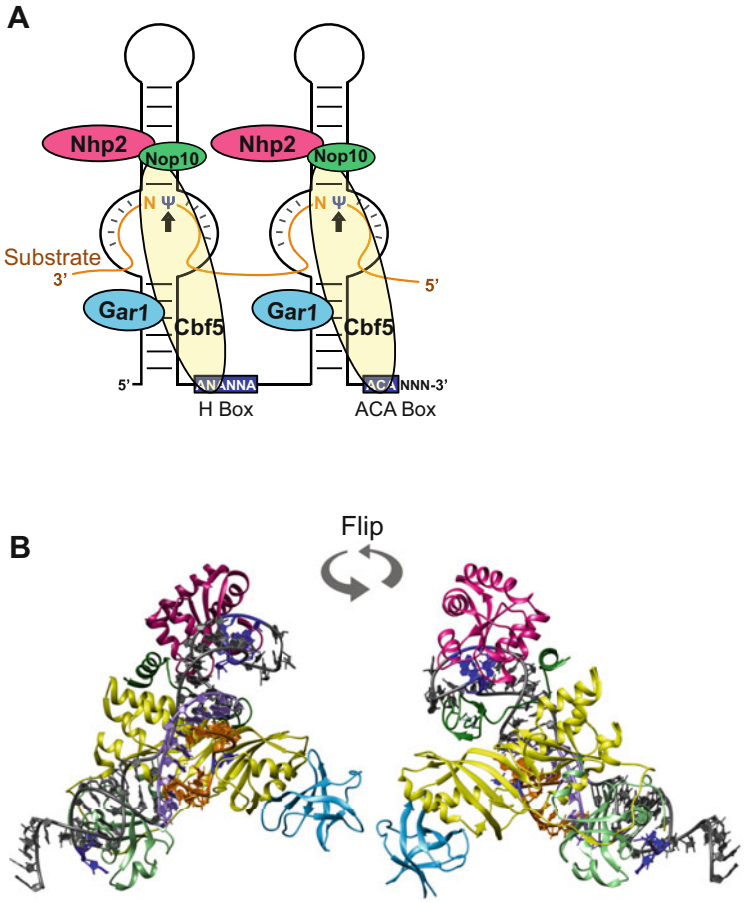


Fig. 2 (**a**) Schematic diagram of box H/ACA RNA-protein complex (RNP). The core components of a box H/ACA RNP, a box H/ACA RNA, and four proteins (Nhp2, Nop 10, Gar1, and Cbf5) are shown. An RNA substrate paired with the two internal loops of the box H/ACA RNA is also shown. The arrows indicate the target nucleotides for pseudouridylation. The H box (5'-ANANNA-3′) and ACA box (5'-ACA-3′) are indicated. (**b**) Crystal structure of box H/ACA snRNP with the substrate (PDB: 3HAY). Yellow: Cbf5. Light green: PUA domain of Cbf5. Purple: Ψ pocket. Orange: substrate. Blue: kink-turn, 3′ ACA tail, and Ψ (Duan et al. 2009)

The nature of the interaction between guide RNA and its substrate RNA has attracted a lot of attention over the years. For instance, two independent NMR studies analyzed this complex and showed that the guide-substrate duplex formed at the pseudouridylation pocket does not adopt a standard Watson-Crick geometry (Jin et al. 2007; Wu and Feigon 2007). Instead, the target RNA interacts with the guide on one side only, forming an unusual "helix." Besides, the guide-substrate RNA duplex can form even in the absence of the box H/ACA snoRNP proteins. More recently, two research groups interrogated the rules of guide-substrate

base-pairing and provided evidence that at least eight base-pairs in the pseudouridylation pocket are required for pseudouridylation to occur (De Zoysa et al. 2018; Kelly et al. 2019). There is a significant degree of flexibility in the pseudouridylation pocket to accommodate substrate sequences (De Zoysa et al. 2018; Majumder et al. 2020).

Advances have been made in the understanding of the box H/ACA snoRNPs structure, owing to the successful crystallization of partial or complete archaeal box H/ACA RNPs (from *Pyrococcus furiosus*), either alone or complexed with a single hairpin substrate (Li and Ye 2006; Rashid et al. 2006; Duan et al. 2009). Such efforts have provided a detailed insight into the structural organization of the protein components. Cbf5, Nop10, and L7Ae (the homolog of Nhp2, in archaea) proteins align the guide RNA and direct its pairing with the substrate by binding to the upper stem of the single hairpin. Gar1, on the other hand, regulates the loading and release of the substrate. Moreover, the structure of a yeast box H/ACA RNP has also been solved (Li et al. 2011), indicating that there is an overall structural resemblance with archaeal box H/ACA RNP, except for an archaeal-specific structural RNA motif located in the upper stem of the hairpin, which does not exist in the yeast structure.

2.2 RNA-Independent RNA Pseudouridylation

On the other hand, in RNA-independent pseudouridylation, stand-alone protein enzymes can each recognize their target uridine, presumably through recognizing the specific sequences surrounding the target uridine or the secondary structural features in the target RNA (Carlile et al. 2019). There are six families of Pus proteins: TruA (including Pus1, Pus2, and Pus3), TruB (Pus4; Cbf5, a catalytic component of box H/ACA RNP, also belongs to this family), RsuA (bacterial only and will not be discussed here), RluA (Pus5, Pus6, Pus8, and Pus9), TruD (Pus7), and Pus10. Many of these enzymes can act on mRNAs, whereas some can also modify several types of noncoding RNA (such as snRNA, tRNA, and rRNA). Here, we summarize the main features of each RNA-independent Pus enzyme (Rintala-Dempsey and Kothe 2017; Borchardt et al. 2020).

2.2.1 The TruA Family

Pus1 is localized in the nucleus in yeast and is known to target snRNAs (U44 of U2 snRNA and U28 of U6 snRNA), tRNAs (at positions 1, 26, 27, 28, 34, 36, 65, and 67), as well as mRNAs (Schwartz et al. 2014)—a total of approximately 80 Ψ modifications in various transcripts (Rintala-Dempsey and Kothe 2017). Interestingly, however, Pus1 deletion does not lead to cell death in yeast, except when an additional knockout of Pus4 is combined. Recently, it was further demonstrated that a combination of Pus1 deletion and mutations in $tRNA^{Gln}CUG$ (that decodes CAG codons) could also lead to lethality in yeast (Khonsari and Klassen 2020). Pus2 has

sequence homology with Pus1 and also possesses a TruA catalytic domain, although it is shorter than Pus1 (42 kDa vs 62 kDa). Pus2 is mainly localized in the mitochondrion. It is responsible for mitochondrial tRNA modification at two positions (U27 and U28) as well as mitochondrial mRNA modification at more than 20 sites. Pus3 also has a sequence similar to Pus1 (with the TruA catalytic domain) and is smaller in size (51 kDa). It catalyzes pseudouridylation of more than 70 sites in coding and noncoding RNAs, such as the highly conserved U38 and U39 sites of several mitochondrial and at least 19 cytoplasmic tRNAs. These Ψs might play an important functional role in supporting the ability of yeast to grow at higher temperatures (Han et al. 2015).

2.2.2 The TruB Family

Pus4 belongs to the TruB family of pseudouridine synthases and can be localized both in the nucleus and the mitochondrion of yeast. It catalyzes pseudouridylation of a single position in both cytoplasmic and mitochondrial tRNAs (U55) and several sites in mRNAs (a total of 50 Ψ in the transcriptome). Upon inspection of the target sites, a consensus substrate sequence (GUΨCNANNC, where N represents any nucleotides) was identified for Pus4 (Carlile et al. 2014). Cbf5, a catalytic component of box H/ACA RNP, also belongs to the TruB family. The role of Cbf5 in box H/ACA RNP-catalyzed pseudouridylation has been discussed above (see *2.1*).

2.2.3 The RluA Family

Pus5 is localized in the mitochondrion (Ansmant et al. 2000) and is responsible for the pseudouridylation of mitochondrial 21S rRNA at U2819 (Carlile et al. 2014). This pseudouridine synthase is the smallest of all: 29 kDa. Pus6, another member of the RluA family, is responsible for the pseudouridylation of U31 of cytoplasmic and mitochondrial tRNAs as well as some mRNA at various positions (a total of 12 Ψ sites). Pus8 is a cytoplasmic enzyme (Behm-Ansmant et al. 2004) and catalyzes the pseudouridylation of cytoplasmic tRNAs at U32. This enzyme has two additional domains flanking the catalytic core: an S4-like domain at the N-terminus and a deaminase domain at the C-terminal region. The deaminase domain, which plays no role in pseudouridylation, independently catalyzes the removal of an amino group from 2,5-diamino-6-(5-phospho-D-ribitylamino)pyrimidin-4(3H)-one (DRAP), an intermediate of the vitamin B2 (or riboflavin) biosynthesis pathway. Pus9 also has the S4-like domain but lacks the unusual deaminase domain. Besides, Pus9 harbors a mitochondrial targeting sequence in the N-terminal region, which probably could explain why it can be found in the mitochondrion, in addition to the nucleus and cytoplasm, and pseudouridylates mitochondrially localized tRNAs at U32.

2.2.4 The TruD Family

Pus7 is the only known member of TruD, and the Pus7 enzymes found in different organisms all have an insertion in their catalytical domain that is probably involved in RNA binding. Pus7 localizes in the cytoplasm and nucleus of cells and is responsible for pseudouridylation of various types of RNA at more than 45 sites, including the well-known U2 snRNA (U35 in yeast U2; also at U56 under heat shock conditions) (Wu et al. 2011), cytoplasmic tRNA (U13) (Behm-Ansmant et al. 2003), and various mRNAs (Carlile et al. 2014; Schwartz et al. 2014; Li et al. 2015; Nakamoto et al. 2017). It is also shown that Pus7 can catalyze pseudouridylation of $tRNA^{Tyr}$ at U35 (Behm-Ansmant et al. 2003), although this position can also be targeted by Pus1 (Behm-Ansmant et al. 2003).

2.2.5 The Pus10 Family

Pus10 was identified in 2006 (Roovers et al. 2006), and it alone constitutes the Pus10 family. Outside the context of archaeal organisms, the substrate(s) of Pus10 was unknown. Recently, however, several investigators suggested that Pus10 could be responsible for the pseudouridylation at position U54 (Deogharia et al. 2019) and U55 (Mukhopadhyay et al. 2020) of certain tRNAs in mammalian cells.

2.3 The Enzymatic Reaction Mechanism Converting Uridine to Ψ

Even though Pus proteins from different families have different primary sequences, they share a similar three-dimensional structure (McKenney et al. 2017), particularly in their active sites. Hence, all pseudouridine synthases, either RNA-dependent or RNA-independent, possess a conserved active site domain containing an aspartate residue with catalytic activity. This aspartate residue is responsible for the attack of the target uridine, breaking the covalent bond between the base and the sugar and forming an aspartate-intermediate. Subsequently, the base rotates and forms a new C-C bond, which rejoins the ribosyl moiety and the isomerized base, thus generating the Ψ. A final step, involving the deprotonation of C5 and protonation of N1, is also required to complete the reaction (Fig. 1). While for several years it was believed that the reaction was initiated by a nucleophilic attack either via a Michael addition (i.e., the aspartate carboxylate group would attack the C6 of the target uridine) or through an acylal mechanism (the aspartate attacks the C1' of the sugar instead), it was recently suggested that pseudouridine synthases operate via a glycal mechanism, where the aspartate deprotonates the sugar (C2'H), thus forming a glycal intermediate (Fig. 1) that results in the cleavage of the C1'-N covalent bond (Veerareddygari et al. 2016).

3 Methods for Detection of Pseudouridylation and Pseudouridine

Although being a nucleotide with distinct and unique chemical and physical properties, Ψ is not easily directly detected due to the fact that Ψ is an isomer of uridine and that Ψ is read as uridine in standard RT—and PCR-based sequencing methods (Potapov et al. 2018). To address this issue, great efforts have been put forth to develop novel techniques. As a result, a number of methods are currently available, and some of them can detect Ψ at single-nucleotide resolution.

3.1 Site-Specific Cleavage Followed by Labeling of Ψ and TLC

This simple TLC-based method can be used to detect a Ψ in a site-specific manner (Zhao and Yu 2004; Hengesbach et al. 2008; Wu et al. 2011; Liu et al. 2013). Briefly, the target Ψ site (either at its 5′ or 3′ side) is specifically cleaved by RNase H directed by a 2'-O-methyl-RNA/DNA chimeric oligonucleotide or by a DNAzyme in certain instances. If the cleavage is designed to occur at the site 5′ of the target Ψ, the cleaved 3′ half RNA fragment is phosphorylated with [γ-^{32}P]-ATP and PNK (polynucleotide kinase) and then splint-ligated, with T4 DNA ligase, to a known short RNA in the presence of a bridging DNA oligonucleotide. The ligated RNA is gel-purified and digested with P1 nuclease to completion. The digested mononucleotides are then resolved on thin-layer chromatography (TLC). Because only the target Ψ (or U if not completely modified) is ^{32}P-radiolabeled, TLC allows separation and visualization of uridylate from pseudouridylate, thus providing a quantitative measurement of pseudouridylation at the target site (Fig. 3). Alternatively, if the RNase H is directed to cleave at the site 3′ of the target Ψ, a 5′-end radioactively labeled known RNA oligo is splint-ligated to the cleaved 5′ half RNA with T4 DNA ligase in the presence of a bridging DNA oligonucleotide. The ligated RNA is then fully digested with RNase T2, and the free nucleotides are further separated via TLC, allowing quantification of pseudouridylation at the target site (Fig. 3). It appears that cleaving the RNA 3′ (as opposed to 5′) of the target Ψ generates cleaner and better results.

3.2 CMC-Modification Followed by Primer Extension

Alternatively, the detection of Ψ can be accomplished by applying a derivatization reaction where an electrophilic organic compound CMC (cyclohexyl-N-′-(2-morpholinoethyl)-carbodiimide metho-*p*-toluene-sulfonate) reacts with Ψ, generating a bulky modified nucleobase (Bakin and Ofengand 1993). Because this

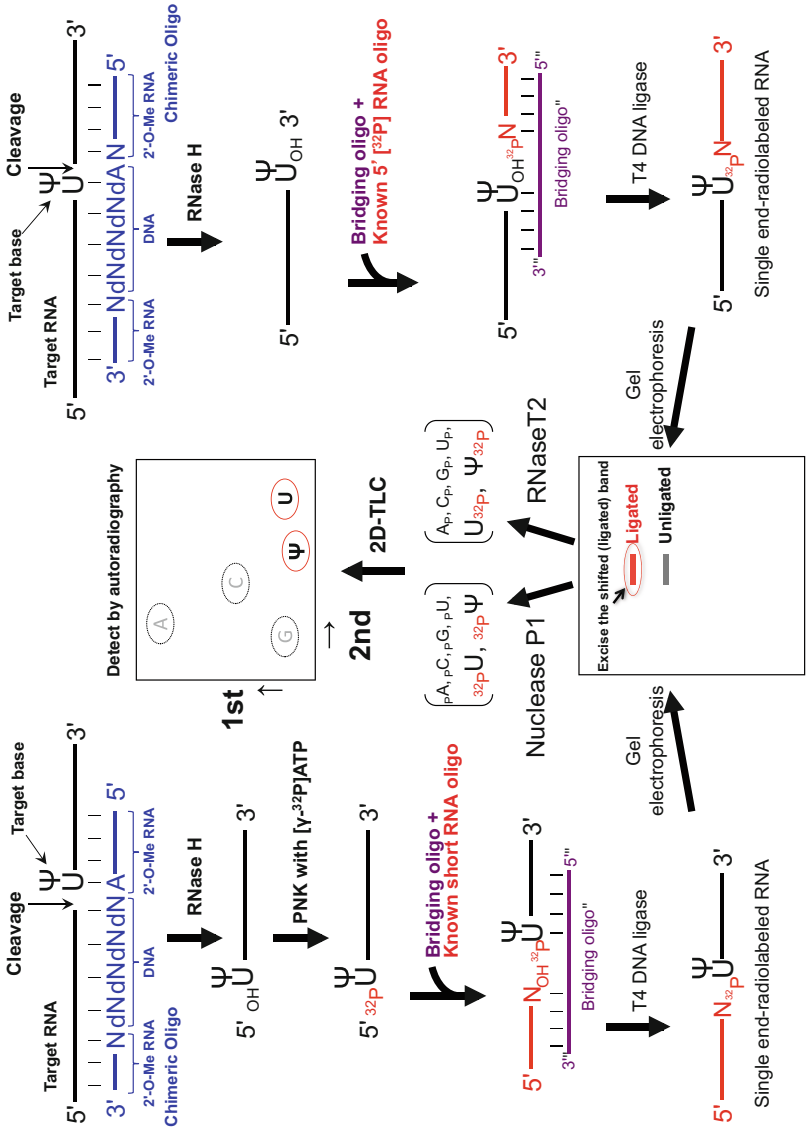


Fig. 3 Schematic description of the site-specific cleavage/labeling/TLC analysis. A stepwise illustration (2′-O-methyl RNA-DNA chimera-directed RNase H cleavage, labeling with phosphorylation/ligation, gel-purification, RNase digestion, and thin-layer chromatography using TLC cellulose) is provided. The relevant RNAs (including the RNA substrate and a known short RNA oligo), the bridging DNA oligo, the target nucleotide (U/Ψ), the enzymes/reagents, etc. are indicated

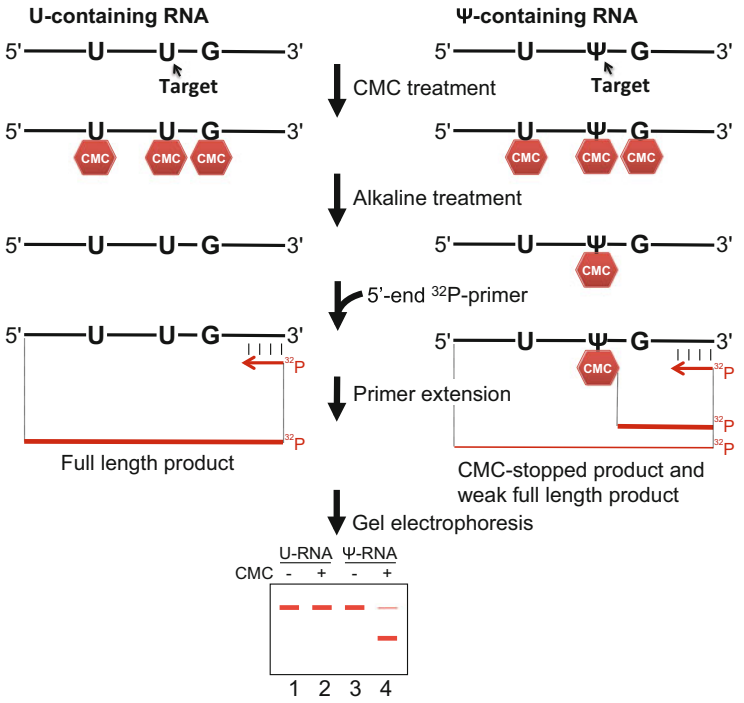


Fig. 4 Schematic description of CMC-modification/primer extension analysis. The unmodified (U-containing RNA) and modified (Ψ-containing RNA) RNAs (indicated) are subjected to pseudouridylation assay—CMC modification followed by primer extension. CMC molecules as well as the target nucleotides (a U and a Ψ) are indicated. The red lines indicate primer extension products, generated after CMC-modification/alkaline treatment. A short (additional) product/band, observed on a denaturing gel after primer extension, is generated only when the RNA contains a Ψ (lane 4)

chemical compound can also react with other nucleotides, i.e., guanosine and uridine, a subsequent alkaline hydrolysis treatment is required so that only the derivatized Ψ residues (Ψ-CMC) remain in the RNA (Bakin and Ofengand 1993; Adachi et al. 2019b). Upon primer extension, the modified Ψ (Ψ-CMC), in turn, can stop the reverse-transcriptase one nucleotide before it, enabling the mapping of Ψ in the RNA (Fig. 4). This approach is especially useful in the context of the identification of multiple Ψ residues in the same target RNA. While still being widely used, the CMC-derivatization method has some limitations, especially there is a need for the abovementioned alkaline treatment, which can generate false positives in the primer extension reaction, due to RNA degradation under those conditions. To address this limitation, a recently developed alternative Ψ derivatization method was proposed using hydrazine (HydraPsiSeq). This reagent creates abasic sites specifically at uridine positions by opening up the ring of the uridine bases. Subsequent treatment with aniline cleaves those abasic sites (Marchand et al. 2020). The Ψ

positions are insensitive to this hydrazine/aniline treatment and are thus mapped by sequencing.

3.3 Next-Generation Sequencing (Ψ—Seq)

Several research groups have adapted next-generation sequencing to carry out a transcriptome-wide mapping of Ψ modifications in yeast and human cells (Carlile et al. 2014; Lovejoy et al. 2014; Schwartz et al. 2014; Li et al. 2015; Nakamoto et al. 2017). Essentially, they couple the Ψ-CMC modification and next-generation sequencing to enable the identification of Ψ in novel RNA species, including low-abundance RNAs such as mRNA and many different types of noncoding RNAs. First, polyA-selected mRNAs are treated with CMC (Carlile et al. 2014, 2015; Lovejoy et al. 2014; Li et al. 2015; Nakamoto et al. 2017). RNAs are fragmented to uniformly cover the transcriptome. Alkaline treatment of CMC-treated RNAs in Pseudo-seq (Carlile et al. 2014, 2015), PSI-seq (Lovejoy et al. 2014), and Ψ-seq (Schwartz et al. 2014) removes CMC from guanidines and uridines, but not pseudouridines due to the resistance of N3-CMC-Ψ to alkaline hydrolysis. An adapter is ligated to the 3′ end of the RNAs, which are then hybridized to an RT primer and reverse transcribed. RT stops at CMC-Ψs to yield truncated cDNAs. cDNAs are circularized (Pseudo-seq and CeU-seq) or ligated to a 3′ adapter (PSI-seq and Ψ-seq), PCR amplified and sequenced. To accurately map Ψs, a separate library of CMC-untreated sequences is prepared to differentiate between CMC-induced and natural RT stops caused by RNA secondary structure (Carlile et al. 2015). A direct comparison between the two libraries leads to the identification and mapping of Ψs.

To detect lower abundance Ψ-sites in mRNAs, CeU-seq enriches Ψ-containing RNAs by conjugating biotin to CMC-treated mRNAs, which are then pulled down before reverse transcription (Li et al. 2015). CeU-seq enriched Ψ-containing RNAs ~15—to 20-fold compared to input RNA. PSI-seq was also used to identify sites of mRNA pseudouridylation by TgPus1, which is necessary for differentiation of the parasite *T. gondii* from active to chronic infection (Nakamoto et al. 2017). More recently, another method, termed RBS-seq, was developed to simultaneously detect Ψ, m^5C, and m^1A modifications. Treatment of RNA with bisulfite not only modifies m^5C and m^1A but also Ψ, yielding ring-opened Ψ-bisulfite isomers, which cause base-skipping during cDNA synthesis (Fleming et al. 2019; Khoddami et al. 2019). The deletion signatures are then amplified by PCR and quantified by sequencing. Read-through of Ψ-adducts in RBS-seq permits mapping of more than one Ψ per RNA strand.

3.4 Direct Nanopore RNA Sequencing

Due to the intrinsic limitations of conventional next-generation sequencing of RNA (through cDNA), especially for the mapping of posttranscriptional RNA modifications, the direct nanopore sequencing that avoids the cDNA preparation and amplification steps has attracted much attention in the epitranscriptomics field (Saletore et al. 2012; Garalde et al. 2018; Liu et al. 2019). The nanopore (protein nanopore) is set in a membrane, to which an electric field is applied. When the RNA sample is loaded, each RNA molecule passes through a nanopore in a stepwise manner, one nucleotide after another. Each passing nucleotide of the RNA can be detected by a small current variation that is specific for each of the nucleotides. The modified nucleotides also generate a unique current difference when compared to the original unmodified nucleotides, thus enabling a base-calling. For instance, in a recent study, N6-methyladenosine (m6A) modifications were detected in yeast RNA with an accuracy of 87% (Liu et al. 2019), increasing the hype surrounding this technique in epitranscriptomics. With this technology, it becomes possible to simultaneously detect and map different RNA modifications (Workman et al. 2019), showcasing how much the field has advanced. Because of the different electronic properties conferred by Ψ versus uridine, there was some expectation whether pseudouridine would also be detected by this new technology. Indeed, it has recently been shown that the nanopore sequencing technology can detect the single pseudouridine typically present in *E. coli* 16S rRNA (Smith et al. 2019). This direct sequencing technique can also be combined with Ψ-specific derivatization methods, i.e., acrylonitrile adducts (cyanoethyl Ψ), to successfully detect Ψ in RNA (Ramasamy et al. 2020).

3.5 Mass Spectrometry

An alternative Ψ detection method relies on the use of mass spectrometry (MS). Since Ψ has the same molecular mass as uridine, derivatization of Ψ with chemicals, such as CMC (Durairaj and Limbach 2008), acrylonitrile (Mengel-Jørgensen and Kirpekar 2002), or methyl vinyl sulfone (Emmerechts et al. 2005), has been used. A direct MS-based Ψ detection method has also been proposed. This method takes advantage of the unique glycosidic bond of Ψ (C-C instead of N-C) as well as the collision-induced dissociation mechanism to generate molecular mass fragments distinguishable from uridine-derived fragments (Pomerantz and McCloskey 2005). Tandem MS/MS analysis is performed after RNase digestion of the RNA sample, and subsequent fragmentation at the different nucleotide positions (at the phosphodiester backbone) allows the precise mapping of Ψ in the RNA.

4 Potential Therapeutic Applications of Pseudouridylation

Given the importance of Ψ in RNAs that perform critical functions in the cells, such as rRNAs in protein translation and snRNAs in pre-mRNA splicing, it is not surprising that any defects in pseudouridylation are potentially associated with human diseases.

4.1 Mutations in the Box H/ACA snoRNP Machinery

It has long been known that defective pseudouridylation due to lack or improper function of the guide RNA-dependent pseudouridylation machinery can be linked to diseases. For example, mutations in the gene that encodes the catalytic unit of box H/ACA RNPs, dyskerin, also known as NAP57, can lead to an X-linked form of dyskeratosis congenita (DC) (Grozdanov et al. 2009). These mutations are usually located in protein regions that interact with the box H/ACA snoRNA (Trahan et al. 2010) and could potentially lead to aberrant pseudouridylation of rRNA. Unsurprisingly, dyskerin (DKC1) knockdown results in a significant reduction in the levels of rRNA pseudouridylation (Schwartz et al. 2014). DC is a bone marrow failure syndrome that affects several tissues, mainly characterized by skin pigmentation, nail abnormalities, and oral leukoplakia, ultimately leading to anemia and cancer. Mutations in other core components of the pseudouridylation machinery, such as Nop10p and Nhp2p, can lead to autosomal forms of DC (Mason and Bessler 2011). Recently, additional phenotypes caused by new mutations in the DKC1 gene and NOP10 were described, particularly nephrotic syndrome with clouding of the eye lens, deafness, and inflammation of the digestive tract (Balogh et al. 2020). In this study, the structure of the box H/ACA snoRNPs was analyzed, and it was found that these new mutations were located in the interface between dyskerin and NOP10, thus preventing the catalytic activity of the snoRNPs and ultimately lowering the rRNA pseudouridylation levels in patients with these mutations.

Furthermore, a possible association between the overexpression of dyskerin and some cancers has been suggested by three different research teams. One of them analyzed patient material from 70 breast carcinomas and found a correlation between DCK1 expression (mRNA and protein) and the clinical outcome of patients. The tumors with lower mRNA DCK1 levels had a better prognosis, which is consistent with the assumption that cancer cells have higher needs of protein synthesis and hence robust rRNA biogenesis, including pseudouridylation (Montanaro et al. 2006). Another study showed that DCK1 is overexpressed in prostate cancers, and the overexpression might be even required for prostate cancer progression (Sieron et al. 2009). Finally, it was also found that DCK1 is overexpressed in glioma tissues, further supporting the hypothesis that DCK1 downregulation could be beneficial in certain types of cancer especially if such treatment would be performed in a localized manner, directly in the tumor (Miao et al. 2019).

4.2 Mutations in the Pus Enzymes

Just as box H/ACA snoRNPs, guide RNA-independent pseudouridine synthases also play an essential role in cellular homeostasis, and mutations that affect the function of these proteins have been correlated with human diseases. One missense mutation occurring in the active site of Pus1 has been identified as the responsible factor for mitochondrial myopathy and sideroblastic anemia (MLASA), a devastating disease caused by abnormalities in the oxidative phosphorylation process and iron metabolism (Bykhovskaya et al. 2004). This missense mutation (C656T) replaces an arginine residue with tryptophan (R116W) in a highly conserved domain (RTDKGV) that constitutes the active site where the catalytic aspartate residue responsible for the initiation of the U-to-Ψ isomerization is located. It has been hypothesized that this mutation could affect the pseudouridylation of mitochondrial tRNA and impact tRNA structure and function, thus causing the disease. Based on this initial assumption, cytoplasmic and mitochondrial tRNA fractions were purified from patient-derived cell lines and a control cell line, and tRNA pseudouridylation patterns were directly compared. It was found that several Pus1-dependent Ψ modifications at positions 27 and 28 were missing from several tRNAs in the MLASA patient-derived samples (Patton et al. 2005). Additional mutations have since been identified in the Pus1 gene in MLASA patients (Kasapkara et al. 2017), and hopefully, the availability of a Pus1-knock out mouse model will accelerate the search for new therapies for MLASA (Mangum et al. 2016).

Mutations in other Pus genes have also been correlated with disease. For example, a study identified a nonsense mutation Arg435X in the *Pus3* gene to be linked to an autosomal recessive form of intellectual disability (ID) (Shaheen et al. 2016). Pus3 catalyzes the guide RNA-independent pseudouridylation of tRNA at position 39 in the anticodon stem-loop. This study confirmed the link between the mutation and the phenotype by measuring decreased tRNA pseudouridylation levels in ID patient-derived cells. Mutations in Pus7, a gene that intervenes in pseudouridylation of mRNA and tRNA targets, are also related to cognitive impairment. It has been reported that several deleterious variants of Pus7 are responsible for a set of conditions such as speech delay and aggressive behavior, among others (de Brouwer et al. 2018). The authors hypothesized that the activity of Pus7 was hampered in affected patients, and they performed primer extension on tRNA extracted from patient-derived cell lines to inspect for abnormalities in the pseudouridylation pattern. Initially, the authors were able to show reduced pseudouridylation at position 13 of tRNA-Glu, in patient-derived samples. When performing a broader sequencing analysis using Ψ-seq, they found additional tRNAs with reduced pseudouridylation levels at the Pus7-targeted position (position 13) in patient-derived samples when compared to that from control subjects. More recently, an additional missense mutation was reported in the Pus7 gene, which is responsible for a milder cognitive impairment phenotype (Darvish et al. 2019).

Hopefully, future therapeutic strategies that make use of the concept of targeted pseudouridylation via artificial box H/ACA guide RNAs (discussed in Sect. 4.5)

could potentially correct the consequences of mutations in guide RNA-independent Pus proteins.

4.3 Ψ as a Potential Biomarker

Biomarkers are being increasingly used in the context of clinical development, as a faster way to predict the effect of a certain drug as well as an alternative to conventional clinical outcomes. Even before trials begin, biomarkers can be used to predict which patients are more likely to respond well to the effect of a drug and help the patient selection process. During and after the trials, biomarkers can quantify and assess the effects of drug treatment, from an efficacy and safety standpoint. The regulatory authorities are increasingly open to accepting biomarkers as surrogates for clinical endpoints. Many diseases have been associated with changes in RNA modifications such as Ψ. Since Ψ cannot be recycled back to uridine, it is usually excreted from the body. Ψ is thus a candidate for a potential biomarker. Several chromatographic (HPLC), spectrometric (NMR), and immunological-based methods (ELISA) have been developed to detect Ψ in biological matrices, such as urine, blood (plasma or serum), amniotic fluid, feces, and several human tissues (Stockert et al. 2020). If well established as a biomarker, the detection of Ψ in easily obtained matrices such as urine, by highly accurate techniques (mass spectrometry or NMR), could possibly constitute an excellent alternative to invasive methods, such as tissue biopsies. Hence the role of this RNA modification as a biomarker has drawn significant attention.

Ψ levels are increased in urinary samples of Alzheimer's disease (AD) patients when compared to healthy individuals (Lee et al. 2007), and this RNA modification could potentially be a useful biomarker to diagnose AD patients at early disease stages, especially if the mechanisms behind this correlation are well elucidated. Ψ also has a potential to be a biomarker for distinguishing malignant from benign tumors. For instance, patients with malignant prostate tumors have increased urinary Ψ levels when compared with individuals with benign prostatic hyperplasia (Pérez-Rambla et al. 2017). A link between Ψ and this disease was further established in a recent study (Stockert et al. 2019). Specifically, by measuring Ψ levels in various prostate cancer cell lines, representing different disease stages, the investigators established a correlation between levels of Ψ and disease progression. Ψ is also believed to play a role as a biomarker for kidney function. For instance, higher plasma Ψ concentrations were found in chronic kidney disease (CKD) patients, as compared to the Ψ levels in healthy individuals (Sekula et al. 2017). Furthermore, it has been reported that Ψ can be used as a potential biomarker for heart failure (Razavi et al. 2020). Here, the investigators profiled the metabolome of serum obtained from a population of individuals to whom echocardiography was performed. They found an association between Ψ plasma levels and a higher left ventricular mass index (LVMI), a parameter that, when increased, is associated with the early development of heart failure (Razavi et al. 2020). Recently, Ψ was also

identified as a potential biomarker for depression in type 2 diabetes mellitus (T2DM). This connection was based on the fact that significantly lower levels of Ψ were detected in the urine of post-stroke depression patients than in the urine of control individuals (non-depressed T2DM) (Liang et al. 2019).

With the standardization of advanced metabolomics techniques in clinical chemistry especially in the context of clinical trials, it is hopeful that additional biomarker roles for Ψ (and other RNA modifications) will be unveiled.

4.4 Alterations in Box H/ACA snoRNA Expression

Considering the crucial role box H/ACA snoRNAs play in guiding rRNA pseudouridylation, which is crucial for protein synthesis in the cell, it is not surprising that any alteration in the expression of these guide RNAs can have an impact on cell fate and potentially in carcinogenic processes. This potential association has been explored by several laboratories, especially in the context of tumorigenesis and the development of blood disorders.

In one study, SNORA42, a guide for 18S rRNA at U113 and U576 (Kiss et al. 2004), was identified as a highly expressed box H/ACA snoRNA in non-small cell lung cancer (NSCLC), and knockdown of this guide RNA in NSCLC cells can reduce tumorigenicity, both in vitro and in vivo (Mei et al. 2012). The increased expression of SNORA42 in tumor samples of NSCLC patients is correlated with poor survival rates. This snoRNA could also be a predictive biomarker for the monitorization of the progression of colorectal cancer (Okugawa et al. 2017). SNORA21, a guide for pseudouridylation of 28S rRNA at U4401 and U4480 positions, is also correlated with metastasis and tumor progression. A research group confirmed this link while looking for oncogenic markers with a role in colorectal cancer and measuring the expression levels of several snoRNAs in tissues from cancer patients (Yoshida et al. 2017). Likewise, three similar but independent snoRNA profiling studies identified multiple deregulated snoRNAs (including box H/ACA RNAs) in tissues from lung cancer patients (Liao et al. 2010; Gao et al. 2015; Gong et al. 2017). Specifically, two box H/ACA snoRNAs (SNORA21 and SNORA73B) were upregulated in all three studies. SNORA24, a guide responsible for pseudouridylation of 18S rRNA, was identified as a mediator of tumor development (McMahon et al. 2019). It is hypothesized that the lack of SNORA24-mediated rRNA pseudouridylation results in the production of deficient ribosomes in hepatocellular carcinoma (HCC) cells, thus increasing the error rate of translation of genes presumably important for the disease initiation and progression. In another study, an overall downregulation of snoRNAs (including box H/ACA snoRNAs) was observed in malignant plasma cells of patients suffering from multiple myeloma and SPCL (secondary plasma cell leukemia) when compared to healthy controls (Ronchetti et al. 2012). In a subsequent study, the same research group compared the expression levels of several box H/ACA snoRNA in cells from CLL (chronic lymphocytic leukemia) patients with that in cells from healthy individuals and

found a downregulation pattern of several guide RNAs (SNORA31, SNORA6, SNORA62, and SNORA71C) in CLL cells, particularly a lower SNORA70F level in patients with a poor disease prognosis (Ronchetti et al. 2013).

With the expanding use of next-generation sequencing methods and the increasing importance of epitranscriptomics in medicine, especially in diagnostics and monitorization of disease progression, it is expected that additional roles of box H/ACA snoRNA will be unveiled in the future, providing potentially new drug targets.

4.5 Nonsense Suppression by Targeted Pseudouridylation Directed by Artificial Box H/ACA RNAs

The therapeutic application of pseudouridylation with potentially more impact in human disease was presented as a completely new mechanism of action to suppress nonsense mutations (Karijolich and Yu 2011). Premature termination codons (PTC) are caused by (nonsense) mutations that introduce early stop codons in mRNA and ultimately lead to mRNA degradation by the nonsense-mediated mRNA decay (NMD) mechanism and premature translation termination (Kurosaki et al. 2019). Taking advantage that all PTCs (UAA, UAG, and UGA) have a uridine nucleotide at the first position, these investigators proposed box H/ACA RNA-guided pseudouridylation targeting that uridine as a means to convert those nonsense codons into sense codons (Karijolich and Yu 2011). They tested this idea using several yeast reporter systems. One of the reporter systems was the *ACT1-CUP1* reporter gene, the expression of which is required for cells to grow in the copper-containing media. The authors inserted a PTC in the *CUP1* gene and then used a designer yeast box H/ACA snoRNA (based on SNR81, a natural yeast snoRNA), in which only the pseudouridylation pocket sequence was changed, to target the first uridine of the PTC. Upon pseudouridylation, a full-length protein (read-through of the PTC) was generated. Moreover, the amino acids incorporated into those pseudouridylated stop codons were identified. Under the conditions they used (in yeast culture), ΨGA coded for tyrosine and phenylalanine, while ΨAA and ΨAG coded for serine and threonine (Karijolich and Yu 2011).

It was later suggested that the isomerization of the first uridine in the PTC could promote read-through by favoring unusual interactions between codon and anticodon, thus outcompeting the binding of release factors to the PTC (Fernández et al. 2013). This novel mechanism of action has several advantages. Firstly, it uses machinery (box H/ACA snoRNP) that is endogenously and globally expressed in human cells and tissues. Secondly, the artificial guide RNAs have a great potential to be changed in their pseudouridylation pockets to target virtually any nonsense mutation (De Zoysa et al. 2018). Thirdly, the suppression of nonsense mutations by targeted pseudouridylation can not only promote read-through of the PTC-containing mRNA but also inhibit the NMD (Morais et al. 2020). Since there

is a very high number of genetic disorders caused by nonsense mutations (Mort et al. 2008), the targeted pseudouridylation approach could potentially generate a meaningful impact on drug development.

4.6 Ψ in mRNA Therapeutics

Since the early in vivo proof-of-concept attempts showed that exogenously delivered mRNA had inherent immunogenicity, it was tested whether this innate immune response could be suppressed by replacing mRNA uridines with Ψs. Remarkably, this approach generated some success (Karikó et al. 2005) and boosted the development of an entirely new RNA-based therapeutic modality. The enormous potential for Ψ-modified mRNA was further demonstrated as a means to reprogram several human cell types into embryonic stem cells (Warren et al. 2010). Since these two landmark studies, mRNA has become a major player in drug discovery and development. Indeed, multiple therapeutic mRNA lead candidates, in which U is replaced with Ψ or its derivatives (such as N1-methyl-pseudouridine), have shown increased stability and reduced immunogenicity. This RNA modification has enabled the acceleration of a novel therapeutic technology field now well advanced in clinical trials, especially in the context of the current pandemic of COVID-19 and the current efforts for the development of mRNA vaccines (Haq et al. 2020).

5 Concluding Remarks

Since the discovery of Ψ decades ago, there has been an incredible accumulation of knowledge of the mechanisms of pseudouridylation. This critical mass of expertise expands the original biological scope of this modification and creates new opportunities to develop new technologies. The emerging direct sequencing methods will further accelerate this trend and unravel new cellular pathways where Ψ is expected to play a crucial role. At the same time, a direct clinical application for this modification is ongoing, where disorders caused by nonsense mutations are targeted. This RNA recoding technology has the potential to become an entirely new drug modality, which is a viable alternative to CRISPR-based drugs.

Acknowledgments We thank the members of the Yu lab for valuable discussions. The work performed in the Yu lab is supported by grants from the National Institute of Health (GM138387 and CA241111) and Cystic Fibrosis Foundation (YU20GO). Pedro Morais is a scientific director at ProQR Therapeutics.

References

Adachi H, De Zoysa MD, Yu Y-T (2019a) Post-transcriptional pseudouridylation in mRNA as well as in some major types of noncoding RNAs. Biochim Biophys Acta Gene Regul Mech 1862:230–239

Adachi H, DeZoysa MD, Yu Y-T (2019b) Detection and quantification of Pseudouridine in RNA. Meth Mol Biol 1870:219–235

Ansmant I, Massenet S, Grosjean H et al (2000) Identification of the Saccharomyces cerevisiae RNA:pseudouridine synthase responsible for formation of psi(2819) in 21S mitochondrial ribosomal RNA. Nucleic Acids Res 28:1941–1946

Bakin A, Ofengand J (1993) Four newly located pseudouridylate residues in Escherichia coli 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. Biochemistry 32:9754–9762

Balogh E, Chandler JC, Varga M et al (2020) Pseudouridylation defect due to DKC1 and NOP10 mutations causes nephrotic syndrome with cataracts, hearing impairment, and enterocolitis. Proc Natl Acad Sci U S A 117:15137–15147

Behm-Ansmant I, Grosjean H, Massenet S et al (2004) Pseudouridylation at position 32 of mitochondrial and cytoplasmic tRNAs requires two distinct enzymes in Saccharomyces cerevisiae. J Biol Chem 279:52998–53006

Behm-Ansmant I, Urban A, Ma X et al (2003) The Saccharomyces cerevisiae U2 snRNA pseudouridine-synthase Pus7p is a novel multisite-multisubstrate RNA:psi-synthase also acting on tRNAs. RNA 9:1371–1382

Borchardt EK, Martinez NM, Gilbert WV (2020) Regulation and function of RNA Pseudouridylation in human cells. Ann Rev Genet (Online ahead of print)

Bykhovskaya Y, Casas K, Mengesha E et al (2004) Missense mutation in pseudouridine synthase 1 (PUS1) causes mitochondrial myopathy and sideroblastic anemia (MLASA). Am J Hum Genet 74:1303–1308

Carlile TM, Martinez NM, Schaening C et al (2019) mRNA structure determines modification by pseudouridine synthase 1. Nat Chem Biol 15:966–974

Carlile TM, Rojas-Duran MF, Gilbert WV (2015) Transcriptome-wide identification of Pseudouridine modifications using pseudo-seq. Curr Protoc Mol Biol 112:4.25.1–4.25.24

Carlile TM, Rojas-Duran MF, Zinshteyn B et al (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515:143–146

Darvish H, Azcona LJ, Alehabib E et al (2019) A novel PUS7 mutation causes intellectual disability with autistic and aggressive behaviors. Neurol Genet 5:e356

de Brouwer APM, Abou Jamra R, Körtel N et al (2018) Variants in PUS7 cause intellectual disability with speech delay, microcephaly, short stature, and aggressive behavior. Am J Hum Genet 103:1045–1052

De Zoysa MD, Wu G, Katz R et al (2018) Guide-substrate base-pairing requirement for box H/ACA RNA-guided RNA pseudouridylation. RNA 24:1106–1117

Deogharia M, Mukhopadhyay S, Joardar A et al (2019) The human ortholog of archaeal Pus10 produces pseudouridine 54 in select tRNAs where its recognition sequence contains a modified residue. RNA 25:336–351

Duan J, Li L, Lu J et al (2009) Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. Mol Cell 34:427–439

Durairaj A, Limbach PA (2008) Improving CMC-derivatization of pseudouridine in RNA for mass spectrometric detection. Anal Chim Acta 612:173–181

Emmerechts G, Herdewijn P, Rozenski J (2005) Pseudouridine detection improvement by derivatization with methyl vinyl sulfone and capillary HPLC-mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 825:233–238

Fernández IS, Ng CL, Kelley AC et al (2013) Unusual base pairing during the decoding of a stop codon by the ribosome. Nature 500:107–110

Fleming AM, Alenko A, Kitt JP et al (2019) Structural elucidation of bisulfite adducts to Pseudouridine that result in deletion signatures during reverse transcription of RNA. J Am Chem Soc 141:16450–16460

Gao L, Ma J, Mannoor K et al (2015) Genome-wide small nucleolar RNA expression analysis of lung cancer by next-generation deep sequencing. Int J Cancer 136:E623–E629

Garalde DR, Snell EA, Jachimowicz D et al (2018) Highly parallel direct RNA sequencing on an array of nanopores. Nat Methods 15:201–206

Gong J, Li Y, Liu C-J et al (2017) A pan-cancer analysis of the expression and clinical relevance of small nucleolar RNAs in human cancer. Cell Rep 21:1968–1981

Grozdanov PN, Fernandez-Fuentes N, Fiser A et al (2009) Pathogenic NAP57 mutations decrease ribonucleoprotein assembly in dyskeratosis congenita. Hum Mol Genet 18:4546–4551

Han L, Kon Y, Phizicky EM (2015) Functional importance of Ψ38 and Ψ39 in distinct tRNAs, amplified for tRNAGln(UUG) by unexpected temperature sensitivity of the s2U modification in yeast. RNA 21:188–201

Haq EU, Yu J, Guo J (2020) Frontiers in the COVID-19 vaccines development. Exp Hematol Oncol 9:24

Hengesbach M, Meusburger M, Lyko F et al (2008) Use of DNAzymes for site-specific analysis of ribonucleotide modifications. RNA 14:180–187

Jin H, Loria JP, Moore PB (2007) Solution structure of an rRNA substrate bound to the Pseudouridylation pocket of a box H/ACA snoRNA. Mol Cell 26:205–215

Karijolich J, Yu Y-T (2011) Converting nonsense codons into sense codons by targeted pseudouridylation. Nature 474:395–398

Karikó K, Buckstein M, Ni H et al (2005) Suppression of RNA recognition by toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. Immunity 23:165–175

Kasapkara ÇS, Tümer L, Zanetti N et al (2017) A myopathy, lactic acidosis, Sideroblastic anemia (MLASA) case due to a novel PUS1 mutation. Turk J Haematol 34:376–377

Kelly EK, Czekay DP, Kothe U (2019) Base-pairing interactions between substrate RNA and H/ACA guide RNA modulate the kinetics of pseudouridylation, but not the affinity of substrate binding by H/ACA small nucleolar ribonucleoproteins. RNA 25:1393–1404

Khoddami V, Yerra A, Mosbruger TL et al (2019) Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. Proc Natl Acad Sci U S A 116:6784–6789

Khonsari B, Klassen R (2020) Impact of Pus1 Pseudouridine synthase on specific decoding events in Saccharomyces cerevisiae. Biomol Ther 10:729

Kierzek E, Malgowska M, Lisowiec J et al (2014) The contribution of pseudouridine to stabilities and structure of RNAs. Nucleic Acids Res 42:3492–3501

Kiss AM, Jády BE, Bertrand E et al (2004) Human box H/ACA pseudouridylation guide RNA machinery. Mol Cell Biol 24:5797–5807

Kurosaki T, Popp MW, Maquat LE (2019) Quality and quantity control of gene expression by nonsense-mediated mRNA decay. Nat Rev Mol Cell Biol 20:406–420

Lee SH, Kim I, Chung BC (2007) Increased urinary level of oxidized nucleosides in patients with mild-to-moderate Alzheimer's disease. Clin Biochem 40:936–938

Li S, Duan J, Li D et al (2011) Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase. Gen Dev 25:2409–2421

Li L, Ye K (2006) Crystal structure of an H/ACA box ribonucleoprotein particle. Nature 443:302–307

Li X, Zhu P, Ma S et al (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol 11:592–597

Liang Z-H, Jia Y-B, Li Z-R et al (2019) Urinary biomarkers for diagnosing poststroke depression in patients with type 2 diabetes mellitus. Diabetes Metab Syndr Obes 12:1379–1386

Liao J, Yu L, Mei Y et al (2010) Small nucleolar RNA signatures as biomarkers for non-small-cell lung cancer. Mol Cancer 9:198

Liu H, Begik O, Lucas MC et al (2019) Accurate detection of m6A RNA modifications in native RNA sequences. Nat Commun 10:4079
Liu N, Parisien M, Dai Q et al (2013) Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA 19:1848–1856
Lovejoy AF, Riordan DP, Brown PO (2014) Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in S. cerevisiae. PLoS One 9:e110799
Majumder M, Mukhopadhyay S, Kharel P et al (2020) The presence of the ACA box in archaeal H/ACA guide RNAs promotes atypical pseudouridylation. RNA 26:396–418
Mangum JE, Hardee JP, Fix DK et al (2016) Pseudouridine synthase 1 deficient mice, a model for mitochondrial myopathy with Sideroblastic anemia, exhibit muscle morphology and physiology alterations. Sci Rep 6:26202
Marchand V, Pichot F, Neybecker P et al (2020) HydraPsiSeq: a method for systematic and quantitative mapping of pseudouridines in RNA. Nucleic Acids Res (Online ahead of print)
Mason PJ, Bessler M (2011) The genetics of dyskeratosis congenita. Cancer Genet 204:635–645
McKenney KM, Rubio MAT, Alfonzo JD (2017) The evolution of substrate specificity by tRNA modification enzymes. Enzymes 41:51–88
McMahon M, Contreras A, Holm M et al (2019) A single H/ACA small nucleolar RNA mediates tumor suppression downstream of oncogenic RAS. elife 8:e48847
Mei Y-P, Liao J-P, Shen J et al (2012) Small nucleolar RNA 42 acts as an oncogene in lung tumorigenesis. Oncogene 31:2794–2804
Mengel-Jørgensen J, Kirpekar F (2002) Detection of pseudouridine and other modifications in tRNA by cyanoethylation and MALDI mass spectrometry. Nucleic Acids Res 30:e135
Miao F-A, Chu K, Chen H-R et al (2019) Increased DKC1 expression in glioma and its significance in tumor cell proliferation, migration and invasion. Investig New Drugs 37:1177–1186
Montanaro L, Brigotti M, Clohessy J et al (2006) Dyskerin expression influences the level of ribosomal RNA pseudo-uridylation and telomerase RNA component in human breast cancer. J Pathol 210:10–18
Morais P, Adachi H, Yu Y-T (2020) Suppression of nonsense mutations by new emerging technologies. Int J Mol Sci 21:4394
Mort M, Ivanov D, Cooper DN et al (2008) A meta-analysis of nonsense mutations causing human genetic disease. Hum Mutat 29:1037–1047
Mukhopadhyay S, Deogharia M, Gupta R (2020) Mammalian nuclear TRUB1, mitochondrial TRUB2 and cytoplasmic PUS10 produce conserved pseudouridine 55 in different sets of tRNA. RNA (Online ahead of print)
Nakamoto MA, Lovejoy AF, Cygan AM et al (2017) mRNA pseudouridylation affects RNA metabolism in the parasite toxoplasma gondii. RNA 23:1834–1849
Okugawa Y, Toiyama Y, Toden S et al (2017) Clinical significance of SNORA42 as an oncogene and a prognostic biomarker in colorectal cancer. Gut 66:107–117
Patton JR, Bykhovskaya Y, Mengesha E et al (2005) Mitochondrial myopathy and sideroblastic anemia (MLASA): missense mutation in the pseudouridine synthase 1 (PUS1) gene is associated with the loss of tRNA pseudouridylation. J Biol Chem 280:19823–19828
Penzo M, Montanaro L (2018) Turning uridines around: role of rRNA Pseudouridylation in ribosome biogenesis and ribosomal function. Biomol Ther 8:38
Pérez-Rambla C, Puchades-Carrasco L, García-Flores M et al (2017) Non-invasive urinary metabolomic profiling discriminates prostate cancer from benign prostatic hyperplasia. Metabolomics 13:52
Pomerantz SC, McCloskey JA (2005) Detection of the common RNA nucleoside pseudouridine in mixtures of oligonucleotides by mass spectrometry. Anal Chem 77:4687–4697
Potapov V, Fu X, Dai N et al (2018) Base modifications affecting RNA polymerase and reverse transcriptase fidelity. Nucleic Acids Re 46:5753–5763
Ramasamy S, Sahayasheela VJ, Yu Z et al (2020) Chemical probe-based Nanopore Sequencing to Selectively Assess the RNA modifications. bioRxiv:2020.05.19.105338

Rashid R, Liang B, Baker DL et al (2006) Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. Mol Cell 21:249–260
Razavi AC, Bazzano LA, He J et al (2020) Pseudouridine and N-formylmethionine associate with left ventricular mass index: metabolome-wide association analysis of cardiac remodeling. J Mol Cell Cardiol 140:22–29
Rintala-Dempsey AC, Kothe U (2017) Eukaryotic stand-alone pseudouridine synthases–RNA modifying enzymes and emerging regulators of gene expression? RNA Biol 14:1185–1196
Ronchetti D, Mosca L, Cutrona G et al (2013) Small nucleolar RNAs as new biomarkers in chronic lymphocytic leukemia. BMC Med Genet 6:27
Ronchetti D, Todoerti K, Tuana G et al (2012) The expression pattern of small nucleolar and small Cajal body-specific RNAs characterizes distinct molecular subtypes of multiple myeloma. Blood Cancer J 2:e96
Roovers M, Hale C, Tricot C et al (2006) Formation of the conserved pseudouridine at position 55 in archaeal tRNA. Nucleic Acids Res 34:4293–4301
Saletore Y, Meyer K, Korlach J et al (2012) The birth of the Epitranscriptome: deciphering the function of RNA modifications. Genome Biol 13:175
Schwartz S, Bernstein DA, Mumbach MR et al (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell 159:148–162
Sekula P, Dettmer K, Vogl FC et al (2017) From discovery to translation: characterization of C-Mannosyltryptophan and Pseudouridine as markers of kidney function. Sci Rep 7:17400
Shaheen R, Han L, Faqeih E et al (2016) A homozygous truncating mutation in PUS3 expands the role of tRNA modification in normal cognition. Hum Genet 135:707–713
Sieron P, Hader C, Hatina J et al (2009) DKC1 overexpression associated with prostate cancer progression. Br J Cancer 101:1410–1416
Smith AM, Jain M, Mulroney L et al (2019) Reading canonical and modified nucleobases in 16S ribosomal RNA using nanopore native RNA sequencing. PLoS One 14:e0216709
Stockert JA, Gupta A, Herzog B et al (2019) Predictive value of pseudouridine in prostate cancer. Am J Clin Exp Urol 7:262–272
Stockert JA, Weil R, Yadav KK et al (2020) Pseudouridine as a novel biomarker in prostate cancer. Urol Oncol (Online ahead of print)
Trahan C, Martel C, Dragon F (2010) Effects of dyskeratosis congenita mutations in dyskerin, NHP2 and NOP10 on assembly of H/ACA pre-RNPs. Hum Mol Genet 19:825–836
Veerareddygari GR, Singh SK, Mueller EG (2016) The Pseudouridine synthases proceed through a Glycal intermediate. J Am Chem Soc 138:7852–7855
Warren L, Manos PD, Ahfeldt T et al (2010) Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell 7:618–630
Westhof E (2019) Pseudouridines or how to draw on weak energy differences. Biochem Biophys Res Commun 520:702–704
Workman RE, Tang AD, Tang PS et al (2019) Nanopore native RNA sequencing of a human poly (a) transcriptome. Nature Meth 16:1297–1305
Wu H, Feigon J (2007) H/ACA small nucleolar RNA pseudouridylation pockets bind substrate RNA to form three-way junctions that position the target U for modification. Proc Natl Acad Sci US A 104:6655–6660
Wu G, Xiao M, Yang C et al (2011) U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. EMBO J 30:79–89
Yoshida K, Toden S, Weng W et al (2017) SNORA21–an oncogenic small nucleolar RNA, with a prognostic biomarker potential in human colorectal cancer. EBioMedicine 22:68–77
Yu AT, Ge J, Yu Y-T (2011) Pseudouridines in spliceosomal snRNAs. Protein Cell 2:712–725
Yu Y-T, Meier UT (2014) RNA-guided isomerization of uridine to pseudouridine--pseudouridylation. RNA Biol 11:1483–1494
Zhao X, Yu Y-T (2004) Detection and quantitation of RNA base modifications. RNA 10:996–1002

Functions of MicroRNA Methylations in Cancer: From Bench to Bedside

Manon Duforestel, Florestan Courant, Louis Paré, Aurélien A. Sérandour, and Pierre-François Cartron

Contents

M. Duforestel · F. Courant · L. Paré
CRCINA, INSERM, Université de Nantes, Nantes, France

LaBCT, Institut de Cancérologie de l'Ouest (ICO), Saint Herblain, France

Cancéropôle Grand-Ouest, réseau Epigénétique (RepiCGO), Nantes, France

EpiSAVMEN network, Nantes, France

A. A. Sérandour
CRCINA, INSERM, Université de Nantes, Nantes, France

LaBCT, Institut de Cancérologie de l'Ouest (ICO), Saint Herblain, France

Cancéropôle Grand-Ouest, réseau Epigénétique (RepiCGO), Nantes, France

EpiSAVMEN network, Nantes, France

Ecole Centrale de Nantes, Nantes, France

P.-F. Cartron (✉)
CRCINA, INSERM, Université de Nantes, Nantes, France

LaBCT, Institut de Cancérologie de l'Ouest (ICO), Saint Herblain, France

Cancéropôle Grand-Ouest, réseau Epigénétique (RepiCGO), Nantes, France

EpiSAVMEN network, Nantes, France

LabEx IGO "Immunotherapy, Graft, Oncology", Nantes, France
e-mail: pierre-francois.cartron@inserm.fr; pierre-francois.cartron@univ-nantes.fr

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_20

Abstract MicroRNAs (miRNAs) play a crucial role in regulation of gene expression. The functionality of miRNAs can be regulated by chemical modifications. Therefore, deciphering these molecular mechanisms appears as a crucial point to understand the miRNA functionality and their implication in cancer. In this chapter, we present the main modifications of microRNAs. We also focus our discussion on the three main methylations (adenosine, guanosine, and cytosine methylation) occurring in miRNA and on proteins regulating these three methylations. The therapeutic targeting of miRNA methylations actors will be also discussed since this field could constitute, in the next years, an extensive research axis comparable to the one targeting the epigenetic actors.

Keywords miRNA · Epigenetics · Epitranscriptomics · 5-Cytosine methylation (m^5C) · 6-Adenine methylation (m^6A) · 7-Guanine methylation (m^7G) · Glioblastoma multiforme · Biomarkers

1 Introduction

In the 1960s, fairly quickly after the discovery of mRNAs as pure intermediates of genetic information, the translational machinery was gradually dissected with the discovery of tRNAs and rRNAs (Crick 1958; Holley et al. 1965; Hoagland et al. 1958). Probably for technological reasons, the following research struggled to reconsider the dogma that RNA cannot be other than messenger, transfer, or ribosomic. For a long time, what was not a tRNA or rRNA was by default classified as mRNA. In the 1980s, the submerged part of the iceberg began to surface with the discovery of small RNAs with regulatory activities. The publication of the human genome in 2001 revealed that only 1.2% of the genome codes for proteins and achieved to shed the light on the noncoding area (Lander et al. 2001). Since then,

studies have highlighted a multitude of noncoding RNAs (ncRNAs), divided into two major classes, separated by length: the short regulatory noncoding RNAs (sncRNAs) less than 200 nts and the long regulatory noncoding RNAs (lncRNAs) over 200 nts. Within these two families, ncRNAs are distinguished mainly according to their biogenesis and their type of targets. To these linear ncRNAs families are added circular ncRNAs, more stable but less expressed than the linear ones. Their functions are still poorly described. What unites all these ncRNAs, including housekeeping ncRNAs (rRNA, tRNA, snRNA, snoRNA), are their regulatory functions, diverse enough to affect all levels of gene expression: nuclear architecture, transcription, transcript stability, translation, and even protein activity (Hombach and Kretz 2016).

MicroRNA (miRNA) is the prevalent class of noncoding RNAs and was firstly discovered in *C. elegans* by Lee RC and colleagues in 1993 (Lee et al. 1993). MiRNAs are endogenous RNAs between 21 and 24 nucleotides long. MiRNAs play a major role in RNA interference by partial or complete base-pairing with their target mRNAs and are key regulators of numerous physiological and pathological processes. They predominantly anneal to the 3'UTR regions of their targets which lead to translational repression or mRNA degradation. With few exceptions, miRNAs transcription is carried out by RNA pol II and is regulated by various transcription factors and epigenetics cofactors modifying the DNA and the histones at the miRNA promoters. MiRNAs are transcribed as hairpin-containing long primary miRNAs (pri-miRNAs), flanked by a cap and a poly-A tail. They then undergo a multiple-step posttranscriptional processing. They are cropped by the microprocessor complex (DROSHA and DGCR8) to isolate small hairpin-shaped fragments called precursor-miRNAs (pre-miRNAs). These are then exported to the cytoplasm by the exportin 5 through the nuclear pores (Yi et al. 2003). The pre-miRNAs meet the DICER protein and its cofactors in the cytoplasm that cut their loops and release the final dsRNA products of 21–24 nts (Bernstein et al. 2001). The pre-RNA-induced silencing complex (pre-RISC) containing the AGO protein is now able to load the mature miRNA into its mRNA targets. One of the RNA strands is discarded, while the other serves as a guide for the AGO protein in the mature RISC complex. Throughout this process, the miRNA processing proteins are subjected to posttranslational modifications (acetylation, phosphorylation, etc.) essential for the tight regulation of their localization, activity, and affinity (Krol et al. 2010).

MiRNAs are also subjected to various forms of regulations such as editing of adenosine in inosine, tailing, and single nucleotide polymorphism, able to interfere with processing, stability, and sequence identity. More recently, modifications occurring during miRNAs maturation have been described as an important regulators of miRNAs regulation. In the large family of noncoding RNAs, tRNAs are a class heavily subjected to chemical changes during their maturation. More widely studied, this dozen tRNA modification is directly implicated in numerous pathologies including cancer (Torres et al. 2014). Posttranscriptional modifications of miRNAs are still poorly described. Interestingly, they seem susceptible of alterations in pathological conditions. For now, the research field has focused on the epigenetic transcriptional regulation of miRNAs (DNA methylation of miRNA gene

promoters) but not on their epitranscriptional regulation (miRNA modifications) (Dakhlallah et al. 2013). A large number of studies reporting the methylation status of miRNAs in cancer actually discuss the methylation of their gene promoters, thus creating confusion. The study of aberrant expression of miRNAs in cancer has already given rise to the use of miRNAs as powerful cancer biomarkers. However, very few articles assess miRNA modifications as potential biomarkers despite the fact that miRNA cytosine-methylation appears to be a widespread chemical modification (Squires et al. 2012). Recently, the use of circulating DNA methylome as a biomarker has already shown its significance in the cancer field (Duforestel et al. 2020), and the associated detection tools, although perfectible, are easy to use. A very recent study has just highlighted the value of blood circulating DNA methylation signatures in multi-cancer detection and in specific identification of the tissue of origin. In comparison to WGS and targeted mutation approaches, their DNA methylation targeted assay is less tedious and more tissue specific, allows deeper sequencing, and ensures this way a low rate of false positives. Currently, this kind of large-scale studies is not applicable to the epitranscriptome, as it is hampered by the little information available on miRNA modifications. However, it is very likely that the miRNA methylome is as relevant and powerful as the DNA methylome. For this reason, we report here the state-of-the-art about miRNA methylation. We emphasize the value of miRNA methylation in cancer research and draw attention to their potential use as diagnostic and therapeutic tools.

2 Overview of "Non-base Methylation" MicroRNA Modifications in Cancer

A large number of RNA modifications exist, but their presence in miRNAs remains understudied. The main explanation of this lack of study is due to the fact that this type of investigations requires the adaptation of conventional methods to detect RNA modifications to small size RNAs. Despite this technical challenge, the literature reports the presence of several miRNA modifications that can be subdivided according to their chemical nature and their localization. In this subsection, we discuss the non-methylation modifications of miRNAs, i.e., RNA editing, ribonucleotide addition, and methylation of phosphate or ribose (Fig. 1).

2.1 Phospho-Dimethylation of MicroRNAs

The 5′ monophosphate end of miRs that is generated by the successive processing is important for subsequent miRNA functions. Xhemalce et al. report that BCDIN3D phospho-dimethylates pre-miR-145 (Xhemalce et al. 2012). Specifically, they highlighted that pre-miR-145 phospho-dimethylation blocks its Dicer-dependent

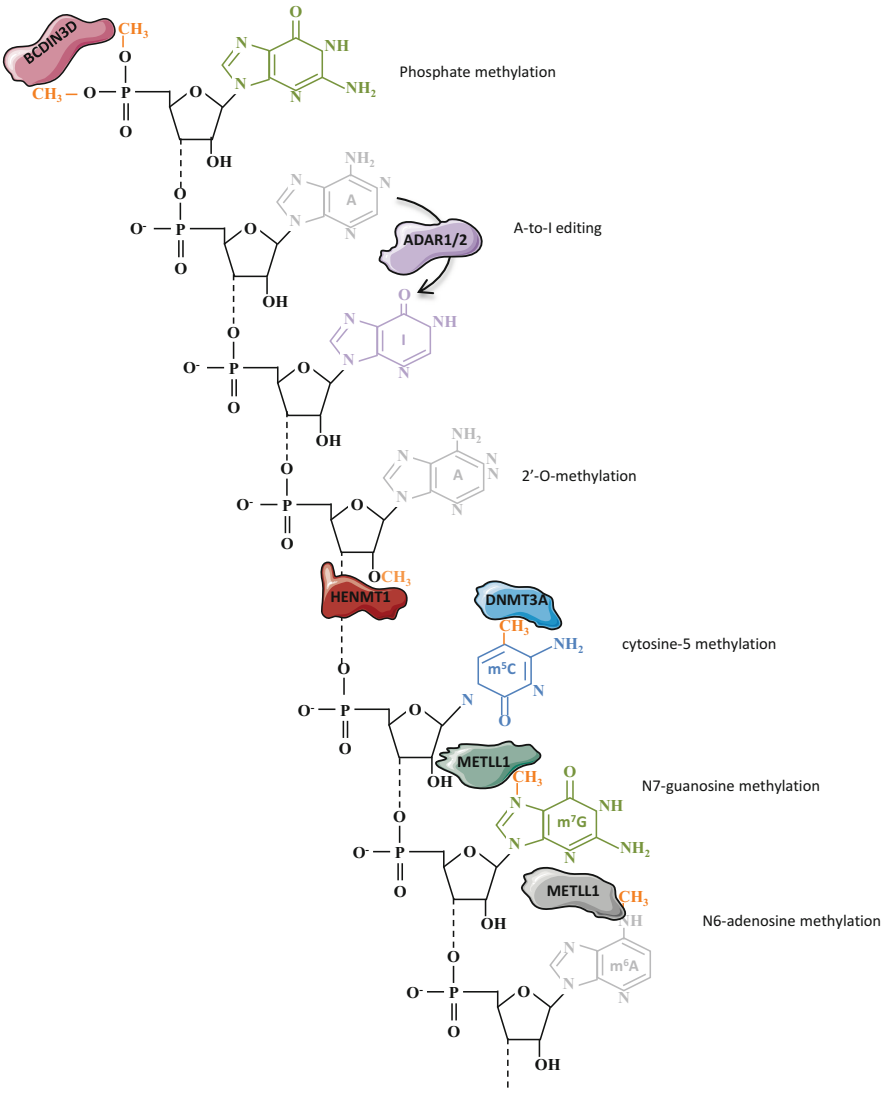


Fig. 1 Representation of the main modifications of microRNAs

processing since Dicer recognizes the 5'terminal negative charge of pre-miRNAs for efficient and accurate cleavage (Park et al. 2011). Therefore BCDIN3D inhibits the miR-145 biogenesis and inhibits its tumor suppressor activity in breast cancer cells. BCDIN3D seems to be a pro-tumoral actor. The work published by Yao et al. goes in the same direction by revealing that the high expression of BCDIN3D predicts worse clinical outcome in breast cancer patients, especially in triple negative breast cancer (Yao et al. 2016). This idea is also supported by the identification of BCDIN3D as a signature of tumorigenic breast cancer cells (Liu et al. 2007).

2.2 A-to-I Editing of MicroRNAs

A-to-I editing is the deamination of adenosine into inosine. This biochemical modification is catalyzed by ADAR1 and ADAR2, two enzymes for which the knock-outs in mice lead to premature lethality (Wang et al. 2000; Higuchi et al. 2000). The A-to-I editing of pri-miRNAs influences their maturation (Yang et al. 2006; Kawahara et al. 2008). At the functional level, the presence of A-to-I editing in the seed sequence of miRNA can alter and/or redirect the targetome of miRNA (Kawahara et al. 2008). In the cancer field, literature reports that the ADAR1-mediated A-to-I editing of the seed sequence of miR-455-5p altered the ability of this miRNA to bind the CPEB1 3'UTR and that this mechanism contributes to the melanoma tumor growth and metastasis (Shoshan et al. 2015). Wang et al. have analyzed the miRNA editing profiles of 8595 samples across 20 cancer types from miRNA sequencing data of The Cancer Genome Atlas and identified 19 adenosine-to-inosine (A-to-I) RNA editing hot spots (Wang et al. 2017). By integrating deep-sequencing, microarray array approaches, and molecular studies, Tomaselli et al. showed that about 90 miRNAs were "edited" in U118 glioblastoma cells (Tomaselli et al. 2015). Particularly, this study reported that ADAR2 can edit miR-222/221 and miR-21 precursors and decrease the expression of the corresponding mature onco-miRNAs in vivo and in vitro, with important effects on cell proliferation and migration.

2.3 2'-O-Methylation of MicroRNAs

While the 2'-O-methylation (2'OMe) can occur on all four ribonucleotides, its detection in miRNAs remains scanty. To date, the most reported example of presence of 2'-O-methylation in miRNA is about adenosine residues of miR-487b-3p (van der Kwast et al. 2018). Besides, it appears that the 2'-O-methylation may protect adenosine from A-to-I editing (Yi-Brunozzi et al. 1999).

Recently, Liang et al. (2020) report that 2'OMe occurs at the 2'hydroxyl group on the ribose at the 3′-end of miR-21-5p in lung cancer (Liang et al. 2020). HENMT1 is identified as the methyltransferase responsible for this 3′-terminal 2'OMe that enhances the stability of miR-21-5p and its association with AGO2. In plants, the 3'terminal 2'OMe is known to protect miRNAs against 3′ truncation and uridylation (Zhao et al. 2012). However this point is not yet reported in mammals.

2.4 Nontemplated Nucleotide Additions (NTA) of MiRNAs

Two major NTA commonly occur in pre- and mature miRNA: the addition of uridines and adenosines (Landgraf et al. 2007).

Several poly(A)polymerase are involved in the 3'adenylation of microRNAs. For example, the poly(A)polymerase GLD2 (also known as PAPD4 or TENT2) catalyzes the 3'adenylation of miR-122 to promote its stabilization (Hojo et al. 2020), and the poly(A)polymerase PAPD5 catalyzes the 3'adenylation of miR-21 to promote its degradation (Boele et al. 2014).

The 3'uridylation of miRNAs is catalyzed by the Terminal UridylTransferases (TUT). Thus, TUT1 is defined as a global regulator of miRNA abundance (Knouf et al. 2013). For instance, literature shows that the TUT1-mediated uridylation upregulate the miRNA-24 and miRNA-29a expression in osteosarcoma (Zhu et al. 2014). TUT2, TUT4, and TUT7 also catalyzed the urydylation of miRNA (Faehnle et al. 2017) (Yamashita et al. 2019). Heo and colleagues reported that TUT2, TUT3, and TUT4 catalyzed the pre-miRNA mono-uridylation of let-7 and regulated its biogenesis, while the polyuridylation of let-7 precursors inhibits the production of its mature form (Heo et al. 2012).

3 Base Methylation of MicroRNAs

Currently, three base methylations have been described in miRNAs: N6-adenosine methylation (m^6A), 7-guanosine methylation (m^7G), and cytosine methylation (m^5C). Echoing to these three distinct types of methylation, several actors of miRNA methylation/demethylation have been identified, and those known are common with mRNAs (Yang et al. 2019). As for any epigenetic process, the actors of miRNAs methylation can be divided into three groups: writers (enzymes adding epimarks to RNA bases), erasers (enzymes removing epimarks), and readers (proteins recognizing and binding epimarks to generate a molecular response) (Table 1). In this subsection, we present the adenosine, guanosine, and cytosine methylations of miRNA and their actors (Fig. 2).

3.1 Adenosine Methylation of MicroRNA

The most described methylation in miRNAs is methylation of the sixth N atom of adenosine (A). Therefore, we talk about N6-methylAdenosine (m^6A). This miRNA epitranscriptomic mark (epimark) is directed by a protein complex containing METTL3, METTL14, and WTAP whose the main play seems to be METLL3. This complex is already described as causing m^6A methylation on mRNAs together with other interaction partners such as HAKAI, KIAA1429, RBM15/RBM15B, and ZC3H13 (which have not yet been characterized in miRNA6-Adenosine methylation) (Lewis et al. 2017). In mRNAs, m^6A methylation occurs on a RRACH or DRACH consensus motif (R = G or A; D = A,G or U; and H = A, C, or U) located in highly conserved regions (Ping et al. 2014). Concerning the miRNAs, a couple of research studies have shown that the METTL3 writing complex recognizes and binds a (R/G)GAC sequence (Alarcón et al. 2015a, b; Zhang et al. 2019). This

Table 1 The main enzymes involved in the microRNA modifications

miRNA-modifying enzyme	Uniprot reference	Type of enzyme	Function
BCDIN3D BCDIN3 domain-containing protein	Q7Z5W3	Writer	Phospho-dimethylation
ADAR1 Adenosine deaminase acting on double-stranded RNA 1	P55265	Writer	A-to-I editing
ADAR2 Adenosine deaminase acting on double-stranded RNA 2	P78563	Writer	A-to-I editing
HENMT HEN1 methyltransferase homolog 11	Q5T8I9	Writer	2'-O-Methylation
GLD2	Q6PIY7	Writer	3'Adenylation
TUT1 Terminal uridylyltransferase 1	Q9H6E5	Writer	3'Urydilation
METTL3 Methyltransferase like 3 protein	Q86U44	Writer	Adenosine methylation
FTO Fat mass and obesity-associated protein	Q9C0B1	Eraser	Adenosine methylation
ALKBH5 Human AlkB homolog H5	Q6P6C2	Eraser	Adenosine methylation
HNRNPA2B1 Heterogeneous nuclear ribonucleoproteins A2/B1	P22626	Reader	Adenosine methylation
NKAP NF-κB-associated protein	8N5F7	Reader	Adenosine methylation
METTL1 Methyltransferase like 1 protein	Q9UBP6	Writer	Guanosine methylation
DNMT3A DNA methyltransferase 3A	Q9Y6K1	Writer	Cytosine methylation

sequence appears to be a consensus fixation motif, but this needs to be confirmed. The existence of multiple or tinier motifs is conceivable given the small size of the miRNAs and therefore the reduced probability of finding a consensus sequence. Using the RNA immunoprecipitation (RIP) with an anti-m^6A-antibody followed by RNA-seq, the work of Berulava and colleagues revealed that a significant fraction of miRNAs contains m^6A (Berulava et al. 2015). Thus, this princeps study underlined the existence of 239 adenosine methylated miRNAs in HEK293 cells.

Surprisingly, a study from Yuan et al. highlights that the tRNA methyltransferase NSUN2 (NOP2/SUN methyltransferase 2), commonly associated with m^5C in miRNA, can also deposit m^6A on RRACH and AAC sequences on mir-125b (Yuan et al. 2014). A mutation in one of this motif abolishes the 6-adenosine methylation. These results support the hypothesis of the existence of multiple consensus motifs and suggest the possibility for NSUN2 to methylate both adenosine and cytosine in RNAs. However this needs to be confirmed, as to our knowledge, no other publication has ever demonstrated bifunctional ability of methylation enzymes.

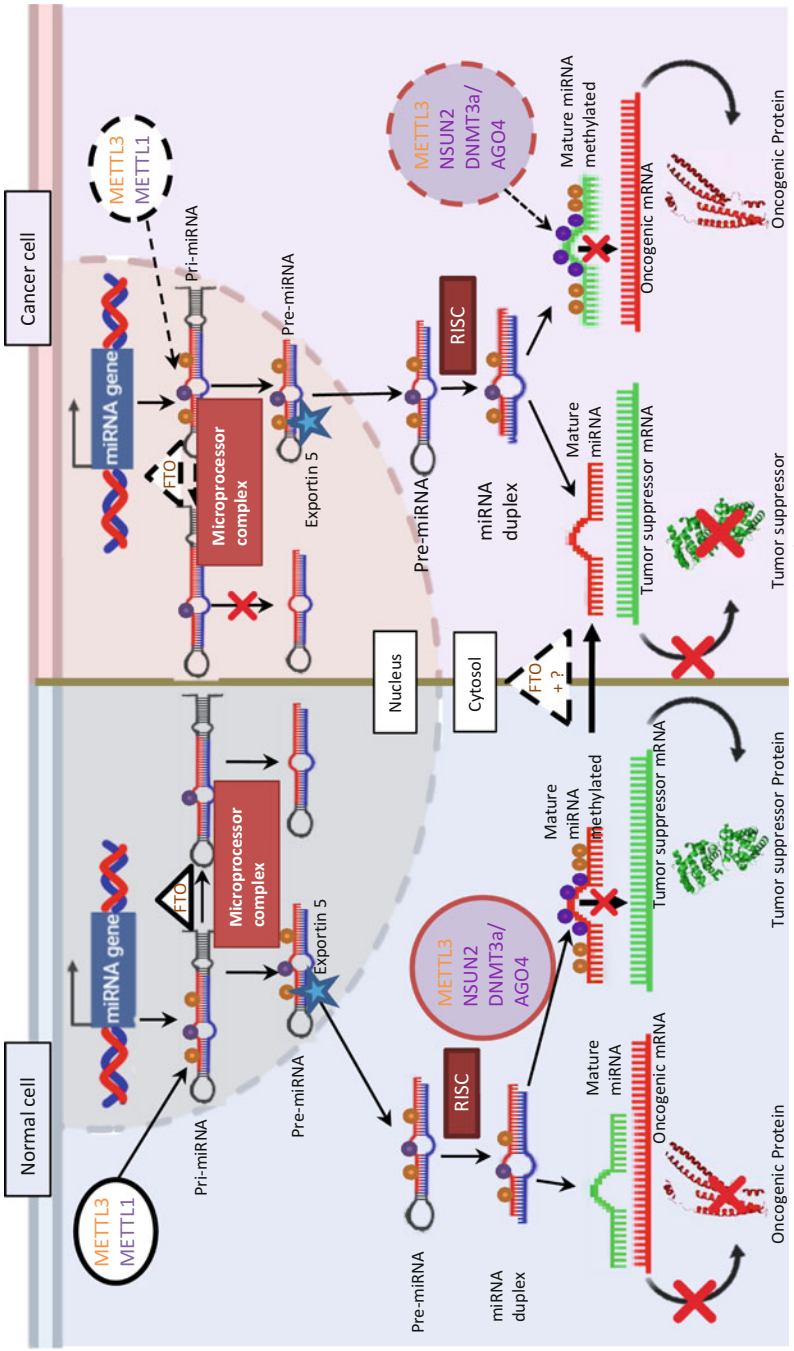


Fig. 2 Adenosine, guanosine, and cytosine methylation of microRNAs. Pri-miRNA methylation (m6A and m7G) by METTL3 and METTL1 on pri-miRNA are necessary for the first step of miRNA maturation. m6A demethylation by FTO inhibits the miRNA maturation. The microprocessor complex recognizes m6A on pri-miRNA thanks to DGCR8 subunit and cleaves the pri-miRNA and releases a hairpin-shaped pre-miRNA. The pre-miRNA leaves the nucleus thanks to the exportin 5, and the RISC complex performs the end of miRNA maturation. METTL3 and/or NSUN2 and DNMT3a/AGO4 are able to methylate (m6A/m5C)

Two enzymes able to erase m^6A methylation in miRNAs are ALKBH5 and FTO (Fat mass and obesity-associated protein) (Xu et al. 2014; Gerken et al. 2007). The study performed by Berulava and colleagues also incriminated FTO as an eraser of adenosine methylation of miRNAs (Berulava et al. 2015). Interestingly, these both enzymes are alpha-ketoglutarate-dependent dioxygenase, thus suggesting a cross talk between metabolism and the epitranscriptomic regulation of miRNA.

Concerning the m^6A readers, we distinguish the nuclear and the cytoplasmic readers. HNRNPA2B1 (Heterogeneous nuclear ribonucleoproteins A2/B1) (Alarcón et al. 2015a, b) and NKAP (NF-κB associated protein) are nuclear readers of the RGm^6AC sequence in the pri-miR (Zhang et al. 2019) and are involved in facilitating the first step of miRNA maturation by interacting with DGCR8 in the microprocessor complex. Cytoplasmic readers belong to the YTH domain family: YTHDF1–3 and YTHDC1–2. The roles of these proteins are well-known on mRNA (Liao et al. 2018) but not yet described in miRNA.

3.2 *Cytosine Methylation of MicroRNAs*

Regarding the miRNA m^5C methylation, there is no data available concerning eraser or reader. No publication demonstrates it clearly, but it is suggested that NSUN2 causes m^5C mark in tRNAs and miRNAs. Recently, we found the complex DNMT3A/AGO4 to be a m^5C writer of miRNAs (Cheray et al. 2020). Knockdown of various proteins involved in miRNA biogenesis shows that DNMT3A or AGO4 downregulations cause a decrease in the level of m^5C of the mir-181a-5p. Co-immunoprecipitation, proximity ligation in situ assay experiments, and measurement of radiolabeled methyl group incorporation in synthetic mir-181a-5p after knockdown of DNMT3A or AGO4 confirmed that DNMT3A and AGO4 are associated in a complex able to cytosine-methylate miRNAs. Erasers and readers of miRNA m^5C are not determined yet.

3.3 *Guanosine Methylation of MicroRNAs*

Concerning m^7G in miRNAs, METTL1 is the principal mRNA methyltransferase and has also been identified as a writer of miRNA by Pandolfini and colleagues (Pandolfini et al. 2019). According to their study conducted on the miRNA let-7,

Fig. 2 (continued) the mature miRNA and to inactive it. Literature showed that some methylated miRNAs are anti-tumoral and they are normally hypomethylated in normal cells. On the opposite, some methylated miRNAs are pro-tumoral and found hypomethylated in normal cells. In cancer cells, the expression in miRNA methylation players could explain the modification of target/balance of pro- or anti-tumoral miRNA methylation

m^7G occurs preferentially on the G11 of the pri-let-7e miRNA. G11 is located on a 16-nt-long G-rich sequence G2 + N4G2 + N4G2 + N4G2 (N = any bases) included in a G-quadruplex-stem loop structure. Some studies have investigated METTL1-binding motif in RNAs. Thus, PAR-CLIP sequencing showed that the "CUCUUCG" and "GGUUCGA" sequences are the most enriched sequence recognized by METTL1 in non-tRNAs and tRNAs, respectively (Bao et al. 2018). Although distinct, these two sequences contain a similar core sequence: UUCG. According to the published data, this core sequence can be extended to UUBD. In parallel, m^7G TRACseq analysis identified the RAGGU sequence as being the most frequently guanosine methylated sequence in tRNA (methylation occurring o, the second guanosine) (Lin et al. 2018). Erasers and readers of miRNA m^7G are not determined yet.

4 Impact of Methylations on MicroRNA Functionality

Like the enzymes producing epimarks, the effects of these chemical modifications on miRNAs are still little studied, both in the physiological and pathological context. However, it is highly possible that methylations are major regulators of miRNAs functionality. The biogenesis of miRNAs is a long and crucial step dependent on a very large number of proteins. It determines the expression level of mature miRNAs and consequently of their targets. In this context, several laboratories have hypothesized that the miRNA maturation depends in part on the presence or absence of methylation marks. Alarcon and colleagues (Alarcón et al. 2015a, b), Peng and colleagues (Peng et al. 2019), and Zhang and colleagues (Zhang et al. 2019) have concluded that pre-miRNA maturation is regulated by m^6A and m^7G methylation/demethylation.

Zhang and colleagues observed that METTL3 overexpression induced an increase of m^6A methylation of pri-miRNA-25 (Zhang et al. 2019). METTL3 overexpression also modified the balance between pri-miR-25 and more mature forms (pre-miR-25 and miR-25-3p). In this case, the level of pri-miR-25 is decreased and other forms are increased. METTL3 knockdown created the opposite effect suggesting that m^6A methylation is essential for pri-miRNA maturation. Furthermore, the increase in miR-25-3p maturation leads to the expression inhibition of its mRNA target PHLPP2 (PH domain leucine-rich repeat protein) and results in the activation of AKT-p70S6K signaling pathway. Peng and colleagues (Peng et al. 2019) have also explored the role of METTL3 in miRNAs. They showed that in colorectal cancers, METTL3 participates in promoting the maturation of miR-1246, causing then an inhibition of its mRNA target, the tumor suppressor SPREED2, and reactivating the MAPK pathway. Alarcon et al. also investigated the biological role of m^6A in miRNA. After METTL3 depletion, they observed a reduction of primary miRNAs in the DGCR8 microRNA processing complex that leads to unprocessing of pri-miRNAs and their accumulation in the nucleus (Alarcón et al. 2015a, b). They hypothesized the existence of nuclear m^6A readers promoting the recruitment of the

microprocessor complex on the pri-miRNAs marked by m^6A. As mentioned in the previous part, they identified the protein HNRNPA2B1 as a nuclear reader of m^6A. They showed that HNRNPA2B1 interacts with the DGCR8 protein, and this way recruits the miRNA processing machinery to the pri-miRNAs.

On their study, Berulava and colleagues observed that the steady state levels of almost all potentially methylated miRNAs are reduced after the invalidation of FTO (Berulava et al. 2015). Thus, they concluded that adenosine methylation of miRNAs affects their biogenesis and/or stability. However, Berulava and colleagues nuanced their conclusion since i) they did not observe significant changes of the primary transcripts of several miRNAs tested, except for pri-let-7e, and ii) many miRNAs are upregulated after FTO knockdown.

According to Pandolfini and colleagues, the m^7G in position 11 in pri-let-7e is likely to be involved in the maturation process (Pandolfini et al. 2019). This epimark seems to enable the destructuration of the G-quadruplex-stem loop and allows the binding of DROSHA to start the miRNA maturation. The m^7G would make the miRNA readable in modifying its secondary structure. Some maturation proteins would interact with the miRNA thanks to the newly opened structure rather than as m^7G readers (illustrate here with the maturation system DROSHA).

Recent studies all show that miRNA methylation promotes miRNA maturation. However, as many other epigenetic modifications, it is possible that miRNA methylation is double-edged. The functional consequences of miRNA methylations might be dependent on the cell type, the base carrying the methylation or even the methylation combination present on the same miR. Methylations could have different functional effects depending on their position on the ribonucleic sequence. The results of Konno and colleagues support this idea (Konno et al. 2019). They demonstrated that the fully m^6A methylated synthetic miR-200c has no longer inhibitory effect on its mRNA targets. The non-methylated and fully m^5C methylated forms continue to have a RNA silencing effect. They also examined the structural interactions between miRNAs and AGO by molecular simulations. The results suggest that the presence of m^5C at position 9 of miR-200c enhances van der Waals interactions with AGO protein but does not alter the conformation of the RNA recognition site. In miR-17-5p and let-7a-5p, m^6A methylation at position 13 and 19, respectively, causes a global conformation change affecting the RNA binding site which could lead to an inability of miRNA to inhibit the mRNA translation.

With a different approach, a very recent study from our laboratory has led to results corroborating the idea that a miRNA can carry different combinations of methylations, regulating each other (Sérandour et al. *Submitted*). In an in vitro sequential methylation assay, we showed that in absence of cytosolic extract, all sequential methylation can be possible, but when we added cytosol extract, the primary methylation by DNMT3A inhibits the secondary methylation by METTL1 or METTL3. In the same way, a primary methylation by METTL3 inhibits a secondary methylation by METTL1 but not totally in the case of DNMT3A. These results can be explained by the position of methylation sites in our miRNA of interest miR-105-p: for DNMT3A, there are two methylation positions C5 and C13, for METTL3 there is A16, and for METTL1 there is G20. C13, A16, and G20

are in very close proximity. We hypothesized that the methylated A16 is recognized by a cytosolic reader that prevents C13 and G20 from being methylated but not C5.

5 MicroRNAs Methylations Investigation Methods

The study of miRNA methylation, and in particular the N^6-methyladenosine (m^6A), 5-methylcytidine (m^5C), and the N^7-methylguanine (m^7G), is not a trivial task because of the lack of adapted method for miRNA methylation analysis. In a general manner, RNA methylation is not a well-understood mechanism. On the contrary, DNA methylation can be studied with a lot of established methods and bioinformatics tools. Since the development of next-generation sequencing and the growing understanding of the role of RNA methylation, some methods have recently been developed to study this biological process. Most of these methods try to reproduce the ones established for DNA methylation. In order to study RNA methylation, the reverse transcription step is problematic because the RNA conversion in cDNA results in the loss of the base modification information. The RNA methylation methods avoided this issue by using immunoprecipitation of methylated nucleotides or by chemical treatment leaving the epigenetic information on the cDNA that is detected by sequencing.

Concerning miRNA methylation analysis, methods need some adaptations because of its short length and its lower expression compared to other RNAs such as tRNAs and mRNAs. In order to study efficiently miRNA methylation, the first step must be the isolation of the miRNAs apart from the other RNAs.

The goal of chemical treatment of miRNA is to keep the epigenetic information in the cDNA. The chemical used is specific to each methylation. To study m^7G, the AlkAniline-Seq can be used (Marchand et al. 2018). It is based on alkaline hydrolysis followed by aniline cleavage which leads to a cut of the RNA at the m^7G location. The m^5C methylation can be detected using a bisulfite treatment leading to a change of the non-methylated cytosine to a thymidine on the final cDNA product, while the methylated cytosine remains a cytosine in the cDNA (Squires et al. 2012). Both of these methylation methods can then be followed by next-generation sequencing, after which well-established bioinformatics tools are used to analyze the raw data.

However, chemical approaches are not sufficient. For example, no chemical allowing to detect m^6A has been developed. Similarly to ChIP-seq and MeDIP, the MeRIP-seq/m^6A-seq uses the immunoprecipitation method to capture the miRNAs containing m^6A (Dominissini et al. 2013). The subsequent library is then sequenced using next-generation sequencing. This method needs a control miRNA-seq. Because of its similarity with the DNA equivalent enrichment method MeDIP, some bioinformatics tools can be used to analyze the data. A recent software called MoAIMS has been developed for the detection of enriched regions from the MeRIP-seq data (Zhang and Hamada 2020).

Another developed technique for the detection of m^6A methylation is based on the high-resolution melting (HRM) which was originally used for the detection of mutation or epimutations in DNA samples (Golovina et al. 2014). It is based on the property of the melting temperature of double-stranded nucleotide chains. For the detection of m^6A methylation in RNA, a complementary probe of the RNA is used to form a double-stranded chain; the presence of m^6A methylation reduces the melting temperature of the complex. Since it is an already known technique, software for analysis already exists. The HRM needs complementary probes of the RNA to study and can be adapted for miRNAs.

Detection of m^6A in miRNAs has also been done using MALDI-TOF as described by Konno et al. (2019). By isolating miRNAs, they were able to detect and quantify m^6A methylation levels of miRNAs in cancer samples. This technique might be adaptable for other types of methylations, and since MALDI-TOF is a well-known method, a lot of software exists to analyze the data (Konno et al. 2019).

Immuno-Northern blotting (INB) technique has also been developed to detect global RNA modifications using antibodies specific against modified bases (Mishima et al. 2015). The INB cannot be used for precise localization of modified bases in the RNA. However, the antibodies anti-m^6A and anti-m^5C used in this study could be used similarly to the MeRIP-Seq presented earlier.

Nevertheless, all the presented methods are specific for only one type of RNA methylation. In order to be able to detect all the methylations of miRNAs using one experiment, a promising new technology called direct-RNA-seq using the sequencer from Oxford Nanopore Technologies seems to be suitable (Ozsolak et al. 2009). This technology is based on nanopores subjected to an electric tension that are able to identify the ribonucleotides passing through thanks to the electric perturbation signatures. Based on the succession of these perturbations, the RNA molecule can be sequenced. By directly sequencing the RNA molecule, the technology by-passes the RT and PCR steps that erase all epigenetic information. However the direct-RNA-seq requires long molecules of more than 100 nucleotides and therefore needs to be adapted for miRNA (epi)sequencing. This technology is promising for detecting several methylations on a single miRNA without any prior chemical treatment. Due to the novelty of this technology, data exploitation stays a hard task, especially for the analysis of methylation profile. Few tools exist for the study of methylation, such as Tombo (Stoiber et al. 2016) or MINES (Lorenz et al. 2020), but they are only focused on the m^5C and m^6A methylation of RNAs.

The study of miRNAs methylation is for now a little explored area, but the growing interest about RNA methylation leads to technology development which can benefit the study of miRNA with some adaptations.

6 About the Main Posttranslational Modifications and Mutations Affecting the MicroRNA Methylation Enzymes in Cancer Cells

The methylation of adenosine, guanosine, and cytosine of miRNA is mainly catalyzed by five enzymes: METTL3, FTO, ALKBH5, METTL1, and DNMT3A (cf. Subsection 3). If a large number of articles report the aberrant expression or posttranslational modification (PTM) of these proteins in cancer context, the majority of these articles associate these modifications with modifications methylation profiles of DNA or other RNA subtypes than miRNAs. Despite this lack (certainly due to the novelty character of the study of methylations of miRNAs), the below subsection presents the main aberrant PTM and/or mutations of METTL1, METTL3, FTO, ALKBH5, and DNMT3A.

6.1 About METTL1

Okamoto and colleagues demonstrated that the expression and the Akt-mediated phosphorylation of METTL1 regulated the methylation of tRNA and determine the sensitivity to 5-fluorouracil in HeLa cells (Okamoto et al. 2014). Cartlidge and colleagues reported that PKB-mediated phosphorylation of METTL1 promotes its inactivation toward the tRNA methylation (Cartlidge et al. 2005), but no investigation reports the impact of this phosphorylation on the guanosine methylation of miRNAs.

6.2 About METTL3

Literature reports that METTL3 can be SUMOylated. Du and colleagues show that METTL3 is modified by SUMO1 mainly at lysine residues K177, K211, K212, and K215, which can be reduced by a SUMO1-specific protease SENP1. SUMOylation of METTL3 does not alter its stability, localization, and interaction with METTL14 and WTAP, but significantly represses its m6A methyltransferase activity resulting in the decrease of m6A levels in mRNAs (Du et al. 2018). Xu and colleagues indicated that the SUMO1-mediated SUMOylation of METTL3 promotes tumor progression via the regulation of Snail mRNA in hepatocellular carcinoma (Xu et al. 2020). None of these two articles has investigated the impact of the METLL3 SUMOylation of the adenosine methylation of miRNAs.

6.3 About FTO and ALKBH5

Several single nucleotide polymorphisms (SNP) affect FTO, and the FTO SNP has been associated with a cancer risk. For example, FTO rs1477196 and rs9939609 were associated with breast and prostate cancer risk (Lewis et al. 2010), respectively. Liu and colleagues report that FTO can be SUMOylated by RANBP2 and that this promotes the FTO degradation and the adenosine methylation of mRNA in hepatocellular carcinoma (Liu et al. 2020). None of these studies has investigated the impact of the FTO SNP or PTM on the generation of aberrant miRNA methylation profiles.

6.4 About DNMT3A

Several point mutations have been reported in DNMT3A in cancer and more particularly in hematological malignancies (Hoang and Rui 2020). Thus, Nguyen and colleagues reported that the R882H DNMT3A hot spot mutation stabilizes the formation of large DNMT3A oligomers with low DNA methyltransferase activity (Nguyen et al. 2019). Deplus and colleagues reported that the CK2-mediated phosphorylation of DNMT3A decreases the ability of this enzyme to methylate DNA (Deplus et al. 2014a, b). The PAD4-mediated citrullination of DNMT3A was described as a mechanism controlling the stability and the activity of DNMT3A (Deplus et al. 2014a, b). Literature also reports that DNMT3A can be SUMOylated by SUMO1 or CDX4 and that these posttranslational modifications affected its de novo methyltransferase activity and/or its capacity to repress transcription (Ling et al. 2004).

7 About the Oncogenic and Tumor Suppressor Role of the MicroRNA Methylation Enzymes

In this subsection, we discuss about the potential oncogenic or tumor suppressor role played by the miRNA methylation enzymes.

The overexpression of METTL1 has been correlated with poor prognosis in hepatocellular carcinoma by promoting cell proliferation and migration (Tian et al. 2019). The authors of this work also reported that METTL1 exerts an oncogenic activity via the suppression of PTEN signaling. But this work did not incriminate the modification of guanosine-methylation of miRNAs in these processes.

Zhang and colleagues show that the high expression of ALKBH5 is associated with a worse prognosis of GBM patients (Zhang et al. 2017). *Mechanistically, Zhang and colleagues revealed that the depletion of ALKBH5 suppresses* the proliferation

of patient-derived GSCs via the regulation of adenosine methylation level of FOXM1 transcripts.

Aberrant expression of FTO is frequently reported in cancer cells. However, its pro- and anticancer role remains controversial (Wang et al. 2020). This situation could be due to the fact that the FTO expression was not associated with a cancer risk, in contrast to FTO variants (see subsection 6).

Two studies evaluated the METTL3 enzymatic activity in cancer. In a letter published in 2017 in *Nature*, Barbieri and colleagues performed a CRISPR KO screen and identified METTL3 as being essential in acute myeloid leukemia (AML) (Barbieri et al. 2017). After further investigation, they showed that METTL3 is involved in AML cell growth by co-transcriptionally m^6A-methylating SP1 mRNA and so promoting the translation of this oncogene. At the opposite, Cui et al. showed that METLL3 overexpression or FTO inhibition suppresses glioblastoma stem cells self-renewal and tumorigenesis (Cui et al. 2016).

Aberrant expression, regulation, or mutation of DNMT3A is described as conferring an oncogenic or tumor suppressor role to DNMT3A.

As a de novo methyltransferase, DNMT3A acts like a real conductor of transcription and so plays a significant role in the maintenance of chromosomal homeostasis. Therefore, defective DNMT3A induce imbalances in DNA methylation, thus resulting in aberrant gene regulation. It is now admitted that the genomes of tumor cells often exhibit genome-wide hypomethylation and/or localized abnormal methylation patterns in particular genomic regions (Dawson and Kouzarides 2012). Those alterations can lead to increase oncogenes expression or on the contrary decrease tumor-protector genes expression (Zhang et al. 2020). More precisely DNMT3A has been identified as mutated in early stage of acute myeloid leukemia development, and its mutation is associated with poor prognosis in patients (Brunetti et al. 2017). DNMT3A is frequently induced overexpressed in hematologic malignancies (Venugopal et al. 2021). High DNMT3A levels have also been reported in lung cancers (Fabbri et al. 2007).

Logically, the loss or gain of methylation due to DNMT3A alterations may also be responsible for the silencing or overexpression of microRNAs. This is what Xiao and colleagues demonstrate in their study, DNMT3A-mediated hypermethylation is responsible for miR-639 expression inhibition, which promotes liver tumorigenesis (Xiao et al. 2020). Cheray and colleagues reported that the selective inhibition of DNMT3A/ISGF3 complex increased the sensitivity of glioblastoma cells to temozolomide and that the high level of presence of DNMT3A/ISGF3 complex is a molecular signature associated with a worse prognosis for GBM patients (Cheray et al. 2016). It is important to clarify that changes in methylation profiles following deregulation of DNMT3A are not always sufficient to cause carcinogenesis but can be considered as oncogenic "hits."

Recently, we highlighted that DNMT3a is not only a DNA methylation player but also a miRNA methylation player (Cheray et al. 2020). This multitasking increases the potential impact of DNMT3A deregulation in cancer development. Even if no study has been published on this field yet, it will not be surprising to highlight in the

future that DNMT3A is involved in tumorigenesis through atypical miRNA methylation (m^5C).

All these examples indicated that the miRNA methylation actors can act as oncogenes or tumor suppressors, but no example incriminates the modification of miRNA methylation in these roles.

8 Methylated MicroRNAs as Diagnostic/Prognostic Tools (Biomarkers)

It is well-known that miRNAs are involved in various mechanisms essential for cell life such as cell growth, metabolism, and apoptosis. A growing number of studies demonstrate the involvement of miRNAs in carcinogenesis. They suggest that alteration of miRNA expression can lead to tumorigenesis through the dysregulation of signaling pathways and transcription factors. These miRNAs could be designated as "oncomiRNAs" as they promote tumor formation by tumor suppressor inhibition (Zhang et al. 2007). For instance, the first validated oncomiRNAmir-17-92 cluster suppresses PTEN and Bim expression (Xiao et al. 2008). However, some miRNAs play a role of tumor suppressor by inhibiting oncogenes' expression. The expression of these miRNAs is found decreased in cancer. For example, Cimmino and colleagues have shown that the overexpression of the anti-apoptotic BCL2 is inversely correlated to miR-15a and miR-16-1 expression in chronic lymphocytic leukemia (CLL) (Cimmino et al. 2005).

In their recent review, Syeda and colleagues gathered current knowledge about the regulation mechanisms of miRNAs expression in cancer, but they only mention miRNA transcription dysregulation and alterations of genes involved in miRNA biogenesis (Ali Syeda et al. 2020). As mentioned in the introduction, a part of these dysregulations is assigned to well-described epigenetic mechanisms in cancer cells (Melo and Esteller 2011) such as cytosine hyper- and hypomethylation of miRNAs DNA loci (Yu et al. 2017), overshadowing the m^6A, m^7G, and m^5C methylations of miRNAs themself. This level of regulation should not be ruled out. These three poorly studied epimarks can impact miRNA functionality (cf. subsection 3), and it is assumed that a link exists between these epimarks and carcinogenesis. The miRNA epimethylome could change during carcinogenesis: by performing the RIP-Seq method, Konno and colleagues (Konno et al. 2019) identified 63 miRNAs commonly methylated in four cell lines of gastrointestinal cancers. They then investigated miRNA methylation levels in serum sample of patients with pancreatic cancers in comparison to healthy controls and found one methylated miRNA (miR-15-5p) specific to sick patients. Our recent study focused on glioblastoma multiforme (GBM) shows that presence of m^5C in miRNA-181a-5p results in the loss of its ability to interfere with BIM mRNA and is associated with poor prognostic in patients (Cheray et al. 2020). We were also able to separate the cohort into three groups: patients with a low level of unmethylated miR-181a-5p compared to the

median value of miR-181a-5p expression, patients with methylated miR-181a-5p, and the last category with a high level of unmethylated miR-181a-5p. Patients with a high level of unmethylated miRNA-181a-5p have a better survival prognostic than the two other groups. Ma and colleagues showed the critical role of m^6A in mRs in metastatic progression in hepatocellular carcinoma (HCC) (Ma et al. 2017). The aberrant decrease in m^6A level in HCC is concomitant with the decrease in METLL14 expression. Low m^6A level in miRNAs is strongly associated with HCC metastasis. Downregulation of METTL14 induces a decrease in miR-126 expression by disrupting pri-miRNA processing, but this study does not give any evidence about the downstream consequences on mRNA targets. Ma and colleagues discussed only the role of METTL14 as a prognostic tool and pointed out the important limitation of the lack of precision of their colorimetric method used for m^6A quantification. A more precise m^6A quantification method would be necessary to establish a direct link between the loss of methylation of miR-126 and tumor progression.

Sun and colleagues reported that the N^6-methyladenosine-dependent pri-miR-17-92 maturation suppresses PTEN/TMEM127 and promotes sensitivity to everolimus in gastric cancer (Sun et al. 2020). At mechanistic level, it appears that the adenosine methylation facilitated the processing of pri-miR-17-92 into the miR-17-2 cluster via a DGCR8-dependent manner.

Recently, Pan et al. (2021) reported that METTL3 promotes adriamycin resistance in MCF-7 breast cancer cells by accelerating pri-microRNA-221-3p maturation in a m6A-dependent manner (Pan et al. 2021).

Despite the limited data available, the involvement of miRNA methylations in cancer seems clear from the conclusions of these three studies. In the coming years, we expect the development of this kind of studies in the context of other cancer types. Currently, miRNAs are used as diagnostic and prognostic tools in various cancer types. We believe that miRNAs are ideal biomarkers because they can circulate freely or in microvesicles in human fluids in a stable and tissue-specific manner (Lan et al. 2020; Mitchell et al. 2008). For now, it consists in highlighting miR-polymorphisms, miRNA variants, and predictive miRNA signatures in patients (Mishra 2014). We can also imagine extending these studies to the miRNA epitranscriptome. The identification of new biomarkers can help to prevent cancer, to adapt treatment, and to ultimately reduce the associated morbidity. Earlier the diagnosis is and better is the survival. Konno and colleagues promote that methylated circulating miRNAs, free or in exosomes, can be very powerful biomarkers especially at early stage of cancer. To reach this objective, the next step is the development of accurate, qualitative, and quantitative technologies dedicated to the investigation of miRNA modifications, as mentioned in Subsection 4.

9 Therapeutic Potential

In this subsection, we will discuss about the idea of targeting the enzymes governing the methylation status of miRNA and of using use methylated miRNA as therapeutic agents.

9.1 Therapeutic Agents Targeting the MicroRNA Methylation Enzymes

Similar strategies than the ones used to develop epigenetic drugs can be used to develop inhibitors of miRNA methylation actors. Thus, it is reasonable to assume that inhibitors of miRNA methylation actors could be developed by trapping the miRNA methyltransferases (by analogy with the 5aza-2-deoxycytidine and DNMT (Stresemann and Lyko 2008)), by inhibiting the catalytic domain of the miRNA methyltransferases or demethylases (by analogy with the procainamide and DNMT1 (Lee et al. 2005)), by competing with the SAM-binding domain of the miRNA methyltransferases (by analogy with other methyltransferases (Zhang and Zheng 2016)), by inhibiting the miRNA demethylases (by analogy with the inhibition of TETs (Xu et al. 2011)), and/or by interfering with the recruitment of methylated miRNA readers (by analogy with JQ1 that inhibit the BRD4 recruitment on DNA (Filippakopoulos et al. 2010)).

Based on these considerations, few articles report the development and/or the use of already-known compounds targeting the miRNA methylation actors. Thus, sinefungin, a well-known inhibitor of methyltransferases (Richon et al. 2011), can be used to decrease the adenosine, guanosine, and cytosine methylation levels of miRNA (Sérandour et al. submitted). The work of Huang and colleagues showed that acid meclofenamic can be used as a FTO inhibitor (Huang et al. 2015). 2-Hydroxyglutarate and IOX3 can be used to inhibit FTO (Su et al. 2018) (McMurray et al. 2015). Thus these three last compounds can be used to increase the m6A level of miRNAs. Recently, Bedi et al. have designed derivatives of the adenosine moiety of SAM that inhibit the m^6A writer METTL3 (Bedi et al. 2020). On contrary, Selberg and colleagues have discovered a small compound that activates the METTL3–14-WTAP complex (Selberg et al. 2019). Thus, this small compound could be used to increase the adenosine methylation of miRNAs. In 2020, the Storm Therapeutics company announced that STC-15, its first-in-class drug candidate targeting METTL3, has been selected for human clinical studies (acute myeloid leukemia (AML) and other solid and hematological cancers) (Table 2).

In theory, we can think that compounds having the ability to inhibit DNMT3A could be used to inhibit the cytosine methylation of miRNAs. Thus, several compounds such as theaflavin 3,3 digallate and thearubigins (compounds having a IC50

Table 2 Examples of inhibitors of microRNA methylation actors

Compound	Target/selectivity
Sinefungin	METTL3, METTL1, DNMT3A, and all methyltransferases
Meclofenamic acid	FTO
2-Hydroxyglutarate	FTO and ALKBH5and also TETs and certain KDMs
STIC-15	METTL3
Derivatives of the adenosine moiety of SAM (compound#7)	METTL3
Theaflavin 3,3 digallate and thearubigins	DNMT3A

values in the lower micro molar) appear to be interesting candidates (Rajavelu et al. 2011).

The selectivity of the compounds inhibiting the miRNA methylation actors is a dual problem. First, METTL3 and the other miRNA methylation enzymes governing the miRNAs methylation status regulate the methylation status of a wide range of miRNAs that can have pro- or anti-tumoral roles. Second, enzymes governing the methylation status of miRNAs also govern the methylation of other bases localized in other RNA subtypes or in DNA (DNMT3A is involved in the cytosine methylation of miRNA and DNA; METLL3 is involved in the adenosine methylation of DNA, miRNA, and mRNA as example). Besides this situation is also true for the enzymes governing the "non-base methylation" modifications of miRNA. Thus, repressing BCDIN3D to decrease the miRNA phospho-dimethylation also inhibits the methylation of cytoplasmic histidyl tRNA ($tRNA^{his}$) since BCDIN3D is also described as a cytoplasmic histidine transfer RNA methyltransferase (Martinez et al. 2017)**.** Despite being promising, the current lack of selectivity inhibitors targeting miRNA methylations actors could lead to major dysregulations and excessive side effects for the patients. However, this opens a new area for the development of new epitranscriptomic drugs that could be used to treat several diseases such as cancer.

9.2 Methylated MicroRNA as Therapeutic Agents

The other strategy, already partly explored, consists in using the miRNAs themselves as therapeutic agents. The development of these therapeutic miRNAs is a real challenge involving the selection of miRNAs without immunogenic effects, moderate toxicity, escaping degradation, and very specific to their mRNA targets together with delivery optimization (Rupaimoole and Slack 2017). Several clinical trials are in progress. Thus, miR-34a was administrated in patients with advanced solid tumor (Clinical trial registration: NCT01829971) (Hong et al. 2020). This phase 1 study was closed early due to serious immune-mediated AEs. However this study provides proof-of-concept for miRNA-based cancer therapy since a dose-dependent

modulation of target genes was observe din white blood cells of miR-34a-treated patients. A second phase 1 clinical trial (Clinical trial registration: NCT02369198) (van Zandwijk et al. 2017) was also performed to determine the acceptable safety profile and early signs of activity of TargomiRNAs in patients with malignant pleural mesothelioma in combination with chemotherapy or immune checkpoint inhibitors. Via these examples, it appears that the clinical used of miRNA in anticancer therapy is a promising ongoing subject.

In addition to the research of ideal manner to administer the miRNAs to patients, extensive research has the challenge to decrease the side effect of therapeutic miR. For this purpose, Briand *and colleagues* recently proposed to use the adenosine methylated form of miRNAs as prodrug of miRNAs in order to minimize the pleiotropic side effect of miRNAs (Briand et al. 2020). More particularly, Briand and colleagues demonstrated that the adenosine-methylated form of miR-200b-3p is demethylated in cancer cells and not in astrocytes, PBMCs, neurons, and hepatocytes since the balance between actors of adenosine methylation and adenosine methylation was in favor of the demethylation in cancer cells. Thus, the use of miRNA chemical modifications to limit their side effect opens a promising axis of research in therapeutics.

10 Conclusion and Future Prospects

With the technological development of sequencing methods, the last 10 years have opened a new area on the understanding on the epitranscriptomic regulation of miRNAs and particularly on the regulation of the base methylations of miRNAs. Despite its novelty and youngness character, the study of miRNAs methylations can build on strong knowledge on its main actors that are the METLL3, FTO, and DNMT3A enzymes. Besides, the fact that METTL3 was already known as a crucial epitranscriptomic actor of other RNAs than miRNA has strongly contributed to the development of METTL3 inhibitors. However, there are still many questions that need to be answered: are there distinct METTL3-including complexes catalyzing the adenosine methylation of mRNA and miRNAs? Answer to this question could permit to increase the selectivity of METTL3 inhibitors and reduce their side effects; are there molecular mechanisms erasing the guanosine and the cytosine methylations of miRNAs? Can several base methylations occur in the same miRNA? How to inhibit the DNMT3A-including complexes catalyzing the cytosine methylation of miRNA without inhibiting the DNMT3A-including complexes catalyzing the cytosine methylation of DNA? Thus, the future investigations in the field of the base methylations modifications of miRNAs look fertile and their therapeutic applications in cancer seem promising.

References

Alarcón CR, Goodarzi H, Lee H et al (2015a) HNRNPA2B1 is a mediator of m(6)A-dependent nuclear RNA processing events. Cell 162:1299–1308

Alarcón CR, Lee H, Goodarzi H et al (2015b) N6-methyladenosine marks primary microRNAs for processing. Nature 519:482–485

Ali Syeda Z, Langden SSS, Munkhzul C et al (2020) Regulatory mechanism of MicroRNA expression in cancer. Int J Mol Sci 21(5)

Bao X, Guo X, Yin M et al (2018) Capturing the interactome of newly transcribed RNA. Nat Methods 15:213–220

Barbieri I, Tzelepis K, Pandolfini L et al (2017) Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control. Nature 552:126–131

Bedi RK, Huang D, Eberle SA et al (2020) Small-molecule inhibitors of METTL3, the major human Epitranscriptomic writer. ChemMedChem 15:744–748

Bernstein E, Caudy AA, Hammond SM et al (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409:363–366

Berulava T, Rahmann S, Rademacher K et al (2015) N6-adenosine methylation in MiRNAs. PLoS One 10:e0118438

Boele J, Persson H, Shin JW et al (2014) PAPD5-mediated 3′ adenylation and subsequent degradation of miR-21 is disrupted in proliferative disease. Proc Natl Acad Sci U S A 111:11467–11472

Briand J, Sérandour AA, Nadaradjane A et al (2020) N6-adenosine methylation of miRNA-200b-3p influences its functionality and is a Theranostic tool. Mol Ther Nucleic Acids 22:72–83

Brunetti L, Gundry MC, Goodell MA (2017) DNMT3A in Leukemia. Cold Spring Harb Perspect Med 7(2)

Cartlidge RA, Knebel A, Peggie M et al (2005) The tRNA methylase METTL1 is phosphorylated and inactivated by PKB and RSK in vitro and in cells. EMBO J 24:1696–1705

Cheray M, Etcheverry A, Jacques C et al (2020) Cytosine methylation of mature microRNAs inhibits their functions and is associated with poor prognosis in glioblastoma multiforme. Mol Cancer 19:36

Cheray M, Pacaud R, Nadaradjane A et al (2016) Specific inhibition of DNMT3A/ISGF3γ interaction increases the Temozolomide efficiency to reduce tumor growth. Theranostics 6 (11):1988–1999

Cimmino A, Calin GA, Fabbri M et al (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A 102:13944–13949

Crick FH (1958) On protein synthesis. Symp Soc Exp Biol 12:138–163

Cui Q, Yang S, Ye P et al (2016) Downregulation of TLX induces TET3 expression and inhibits glioblastoma stem cell self-renewal and tumorigenesis. Nat Commun 7:10637

Dakhlallah D, Batte K, Wang Y et al (2013) Epigenetic regulation of miR-17~92 contributes to the pathogenesis of pulmonary fibrosis. Am J Respir Crit Care Med 187:397–405

Dawson MA, Kouzarides T (2012) Cancer epigenetics: from mechanism to therapy. Cell 150:12–27

Deplus R, Blanchon L, Rajavelu A et al (2014a) Regulation of DNA methylation patterns by CK2-mediated phosphorylation of Dnmt3a. Cell Rep 8:743–753

Deplus R, Denis H, Putmans P et al (2014b) Citrullination of DNMT3A by PADI4 regulates its stability and controls DNA methylation. Nucleic Acids Res 42:8285–8296

Dominissini D, Moshitch-Moshkovitz S, Salmon-Divon M et al (2013) Transcriptome-wide mapping of N6-methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing. Nat Protoc 8:176–189

Du Y, Hou G, Zhang H et al (2018) SUMOylation of the m6A-RNA methyltransferase METTL3 modulates its function. Nucleic Acids Res 46:5195–5208

Duforestel M, Briand J, Bougras-Cartron G et al (2020) Cell-free circulating epimarks in cancer monitoring. Epigenomics 12:145–155

Fabbri M, Garzon R, Cimmino A et al (2007) MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proc Natl Acad Sci U S A 104:15805–15810
Faehnle CR, Walleshauser J, Joshua-Tor L (2017) Multi-domain utilization by TUT4 and TUT7 in control of let-7 biogenesis. Nat Struct Mol Biol 24:658–665
Filippakopoulos P, Qi J, Picaud S et al (2010) Selective inhibition of BET bromodomains. Nature 468:1067–1073
Gerken T, Girard CA, Tung Y-CL et al (2007) The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science 318:1469–1472
Golovina AY, Dzama MM, Petriukov KS et al (2014) Method for site-specific detection of m6A nucleoside presence in RNA based on high-resolution melting (HRM) analysis. Nucleic Acids Res 42:e27–e27
Heo I, Ha M, Lim J et al (2012) Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell 151:521–532
Higuchi M, Maas S, Single FN et al (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. Nature 406:78–81
Hoagland MB, Stephenson ML, Scott JF et al (1958) A soluble ribonucleic acid intermediate in protein synthesis. J Biol Chem 231:241–257
Hoang N-M, Rui L (2020) DNA methyltransferases in hematological malignancies. J Genet Genomics 47:361–372
Hojo H, Yashiro Y, Noda Y et al (2020) The RNA-binding protein QKI-7 recruits the poly (a) polymerase GLD-2 for 3′ adenylation and selective stabilization of microRNA-122. J Biol Chem 295:390–402
Holley RW, Apgar J, Everett GA et al (1965) Structure of a ribonucleic acid. Science 147:1462–1465
Hombach S, Kretz M (2016) Non-coding RNAs: classification, biology and functioning. Adv Exp Med Biol 937:3–17
Hong DS, Kang Y-K, Borad M et al (2020) Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours. Br J Cancer 122:1630–1637
Huang Y, Yan J, Li Q et al (2015) Meclofenamic acid selectively inhibits FTO demethylation of m6A over ALKBH5. Nucleic Acids Res 43:373–384
Kawahara Y, Megraw M, Kreider E et al (2008) Frequency and fate of microRNA editing in human brain. Nucleic Acids Res 36:5270–5280
Knouf EC, Wyman SK, Tewari M (2013) The human TUT1 nucleotidyl transferase as a global regulator of microRNA abundance. PLoS One 8:e69630
Konno M, Koseki J, Asai A et al (2019) Distinct methylation levels of mature microRNAs in gastrointestinal cancers. Nat Commun 10:3888
Krol J, Loedige I, Filipowicz W (2010) The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet 11:597–610
Lan F, Yue X, Xia T (2020) Exosomal microRNA-210 is a potentially non-invasive biomarker for the diagnosis and prognosis of glioma. Oncol Lett 19:1967–1974
Lander ES, Linton LM, Birren B et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921
Landgraf P, Rusu M, Sheridan R et al (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 129:1401–1414
Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75:843–854
Lee BH, Yegnasubramanian S, Lin X et al (2005) Procainamide is a specific inhibitor of DNA methyltransferase 1. J Biol Chem 280:40749–40756
Lewis SJ, Murad A, Chen L et al (2010) Associations between an obesity related genetic variant (FTO rs9939609) and prostate cancer risk. PLoS One 5:e13485
Lewis CJT, Pan T, Kalsotra A (2017) RNA modifications and structures cooperate to guide RNA-protein interactions. Nat Rev Mol Cell Biol 18:202–210

Liang H, Jiao Z, Rong W et al (2020) 3'-terminal 2'-O-methylation of lung cancer miR-21-5p enhances its stability and association with Argonaute 2. Nucleic Acids Res 48(13):7027–7040
Liao S, Sun H, Xu C (2018) YTH domain: a family of N6-methyladenosine (m6A) readers. Genomics Proteomics Bioinformatics 16:99–107
Lin S, Liu Q, Lelyveld VS et al (2018) Mettl1/Wdr4-mediated m7G tRNA Methylome is required for Normal mRNA translation and embryonic stem cell self-renewal and differentiation. Mol Cell 71 e5:244–255
Ling Y, Sankpal UT, Robertson AK et al (2004) Modification of de novo DNA methyltransferase 3a (Dnmt3a) by SUMO-1 modulates its interaction with histone deacetylases (HDACs) and its capacity to repress transcription. Nucleic Acids Res 32:598–610
Liu X, Liu J, Xiao W et al (2020) SIRT1 regulates N6 -Methyladenosine RNA modification in Hepatocarcinogenesis by inducing RANBP2-dependent FTO SUMOylation. Hepatology. https://doi.org/10.1002/hep.31222
Liu R, Wang X, Chen GY et al (2007) The prognostic role of a gene signature from tumorigenic breast-cancer cells. N Engl J Med 356:217–226
Lorenz DA, Sathe S, Einstein JM et al (2020) Direct RNA sequencing enables m6A detection in endogenous transcript isoforms at base-specific resolution. RNA 26:19–28
Ma J-Z, Yang F, Zhou C-C et al (2017) METTL14 suppresses the metastatic potential of hepatocellular carcinoma by modulating N6 -methyladenosine-dependent primary MicroRNA processing. Hepatology 65:529–543
Marchand V, Ayadi L, Ernst FGM et al (2018) AlkAniline-Seq: profiling of m 7 G and m 3 C RNA modifications at Single nucleotide resolution. Angew Chem Int Ed 57:16785–16790
Martinez A, Yamashita S, Nagaike T et al (2017) Human BCDIN3D monomethylates cytoplasmic histidine transfer RNA. Nucleic Acids Res 45:5423–5436
McMurray F, Demetriades M, Aik W et al (2015) Pharmacological inhibition of FTO. PLoS One 10:e0121829
Melo SA, Esteller M (2011) Dysregulation of microRNAs in cancer: playing with fire. FEBS Lett 585:2087–2099
Mishima E, Jinno D, Akiyama Y et al (2015) Immuno-northern blotting: detection of RNA modifications by using antibodies against modified nucleosides. PLoS One 10:e0143756
Mishra PJ (2014) MicroRNAs as promising biomarkers in cancer diagnostics. Biomark Res 2:19
Mitchell PS, Parkin RK, Kroh EM et al (2008) Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A 105:10513–10518
Nguyen T-V, Yao S, Wang Y et al (2019) The R882H DNMT3A hot spot mutation stabilizes the formation of large DNMT3A oligomers with low DNA methyltransferase activity. J Biol Chem 294:16966–16977
Okamoto M, Fujiwara M, Hori M et al (2014) tRNA modifying enzymes, NSUN2 and METTL1, determine sensitivity to 5-fluorouracil in HeLa cells. PLoS Genet 10:e1004639
Ozsolak F, Platt AR, Jones DR et al (2009) Direct RNA sequencing. Nature 461:814–818
Pan X, Hong X, Li S et al (2021) METTL3 promotes adriamycin resistance in MCF-7 breast cancer cells by accelerating pri-microRNA-221-3p maturation in a m6A-dependent manner. Exp Mol Med
Pandolfini L, Barbieri I, Bannister AJ et al (2019) METTL1 promotes let-7 MicroRNA processing via m7G methylation. Mol Cell 74:1278–1290.e9
Park J-E, Heo I, Tian Y et al (2011) Dicer recognizes the 5′ end of RNA for efficient and accurate processing. Nature 475:201–205
Peng W, Li J, Chen R et al (2019) Upregulated METTL3 promotes metastasis of colorectal cancer via miR-1246/SPRED2/MAPK signaling pathway. J Exp Clin Cancer Res 38:393
Ping X-L, Sun B-F, Wang L et al (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res 24:177–189
Rajavelu A, Tulyasheva Z, Jaiswal R et al (2011) The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols. BMC Biochem 12:16

Richon VM, Johnston D, Sneeringer CJ et al (2011) Chemogenetic analysis of human protein methyltransferases. Chem Biol Drug Des 78:199–210
Rupaimoole R, Slack FJ (2017) MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. Nat Rev Drug Discov 16:203–222
Selberg S, Blokhina D, Aatonen M et al (2019) Discovery of small molecules that activate RNA methylation through cooperative binding to the METTL3-14-WTAP complex active site. Cell Rep 26:3762–3771.e5
Shoshan E, Mobley AK, Braeuer RR et al (2015) Reduced adenosine-to-inosine miR-455-5p editing promotes melanoma growth and metastasis. Nat Cell Biol 17:311–321
Squires JE, Patel HR, Nousch M et al (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40:5023–5033
Stoiber M, Quick J, Egan R et al (2016) De novo identification of DNA modifications enabled by genome-guided Nanopore signal processing. Bioinformatics
Stresemann C, Lyko F (2008) Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. Int J Cancer 123:8–13
Su R, Dong L, Li C et al (2018) R-2HG exhibits anti-tumor activity by targeting FTO/m6A/MYC/CEBPA signaling. Cell 172:90–105.e23
Sun Y, Li S, Yu W et al (2020) N6-methyladenosine-dependent pri-miR-17-92 maturation suppresses PTEN/TMEM127 and promotes sensitivity to everolimus in gastric cancer. Cell Death Dis 11:836
Tian Q-H, Zhang M-F, Zeng J-S et al (2019) METTL1 overexpression is correlated with poor prognosis and promotes hepatocellular carcinoma via PTEN. J Mol Med (Berl) 97:1535–1545
Tomaselli S, Galeano F, Alon S et al (2015) Modulation of microRNA editing, expression and processing by ADAR2 deaminase in glioblastoma. Genome Biol 16:5
Torres AG, Batlle E, Ribas de Pouplana L (2014) Role of tRNA modifications in human diseases. Trends Mol Med 20:306–314
van der Kwast RVCT, van Ingen E, Parma L et al (2018) Adenosine-to-inosine editing of MicroRNA-487b alters target gene selection after ischemia and promotes neovascularization. Circ Res 122:444–456
van Zandwijk N, Pavlakis N, Kao SC et al (2017) Safety and activity of microRNA-loaded minicells in patients with recurrent malignant pleural mesothelioma: a first-in-man, phase 1, open-label, dose-escalation study. Lancet Oncol 18:1386–1396
Venugopal K, Feng Y, Shabashvili D et al (2021) Alterations to DNMT3A in hematologic malignancies. Cancer Res 81:254–263
Wang J-Y, Chen L-J, Qiang P (2020) The potential role of N6-Methyladenosine (m6A) demethylase fat mass and obesity-associated gene (FTO) in human cancers. Onco Targets Ther 13:12845–12856
Wang Q, Khillan J, Gadue P et al (2000) Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. Science 290:1765–1768
Wang Y, Xu X, Yu S et al (2017) Systematic characterization of A-to-I RNA editing hotspots in microRNAs across human cancers. Genome Res 27:1112–1125
Xhemalce B, Robson SC, Kouzarides T (2012) Human RNA methyltransferase BCDIN3D regulates microRNA processing. Cell 151:278–288
Xiao J, Liu Y, Wu F et al (2020) miR-639 expression is silenced by DNMT3A-mediated Hypermethylation and functions as a tumor suppressor in liver cancer cells. Mol Ther 28:587–598
Xiao C, Srinivasan L, Calado DP et al (2008) Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes. Nat Immunol 9:405–414
Xu C, Liu K, Tempel W et al (2014) Structures of human ALKBH5 demethylase reveal a unique binding mode for specific single-stranded N6-methyladenosine RNA demethylation. J Biol Chem 289:17299–17311

Xu H, Wang H, Zhao W et al (2020) SUMO1 modification of methyltransferase-like 3 promotes tumor progression via regulating snail mRNA homeostasis in hepatocellular carcinoma. Theranostics 10:5671–5686
Xu W, Yang H, Liu Y et al (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer Cell 19:17–30
Yamashita S, Nagaike T, Tomita K (2019) Crystal structure of the Lin28-interacting module of human terminal uridylyltransferase that regulates let-7 expression. Nat Commun 10:1960
Yang W, Chendrimada TP, Wang Q et al (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. Nat Struct Mol Biol 13:13–21
Yang F, Jin H, Que B et al (2019) Dynamic m6A mRNA methylation reveals the role of METTL3-m6A-CDCP1 signaling axis in chemical carcinogenesis. Oncogene 38:4755–4772
Yao L, Chi Y, Hu X et al (2016) Elevated expression of RNA methyltransferase BCDIN3D predicts poor prognosis in breast cancer. Oncotarget 7:53895–53902
Yi R, Qin Y, Macara IG et al (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev 17:3011–3016
Yi-Brunozzi HY, Easterwood LM, Kamilar GM et al (1999) Synthetic substrate analogs for the RNA-editing adenosine deaminase ADAR-2. Nucleic Acids Res 27:2912–2917
Yu Q, Xiang L, Yin L et al (2017) Loss-of-function of miR-142 by hypermethylation promotes TGF-β-mediated tumour growth and metastasis in hepatocellular carcinoma. Cell Prolif 50(6)
Yuan S, Tang H, Xing J et al (2014) Methylation by NSun2 represses the levels and function of microRNA 125b. Mol Cell Biol 34:3630–3641
Zhang J, Bai R, Li M et al (2019) Excessive miR-25-3p maturation via N6-methyladenosine stimulated by cigarette smoke promotes pancreatic cancer progression. Nat Commun 10:1858
Zhang Y, Hamada M (2020) MoAIMS: efficient software for detection of enriched regions of MeRIP-Seq. BMC Bioinformatics 21:103
Zhang B, Pan X, Cobb GP et al (2007) microRNAs as oncogenes and tumor suppressors. Dev Biol 302:1–12
Zhang J, Yang C, Wu C et al (2020) DNA methyltransferases in cancer: biology, paradox, aberrations, and targeted therapy. Cancers (Basel) 12(8)
Zhang S, Zhao BS, Zhou A et al (2017) m6A demethylase ALKBH5 maintains Tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. Cancer Cell 31:591–606.e6
Zhang J, Zheng YG (2016) SAM/SAH analogs as versatile tools for SAM-dependent methyltransferases. ACS Chem Biol 11:583–597
Zhao Y, Mo B, Chen X (2012) Mechanisms that impact microRNA stability in plants. RNA Biol 9:1218–1223
Zhu D-Q, Lou Y-F, He Z-G et al (2014) Nucleotidyl transferase TUT1 inhibits lipogenesis in osteosarcoma cells through regulation of microRNA-24 and microRNA-29a. Tumour Biol 35:11829–11835

Mass Spectrometry-Based Methods for Characterization of Hypomodifications in Transfer RNA

Priti Thakur, Manasses Jora, Ruoxia Zhao, Gwenn Parungao, Scott Abernathy, Patrick A. Limbach, and Balasubrahmanyam Addepalli

Contents

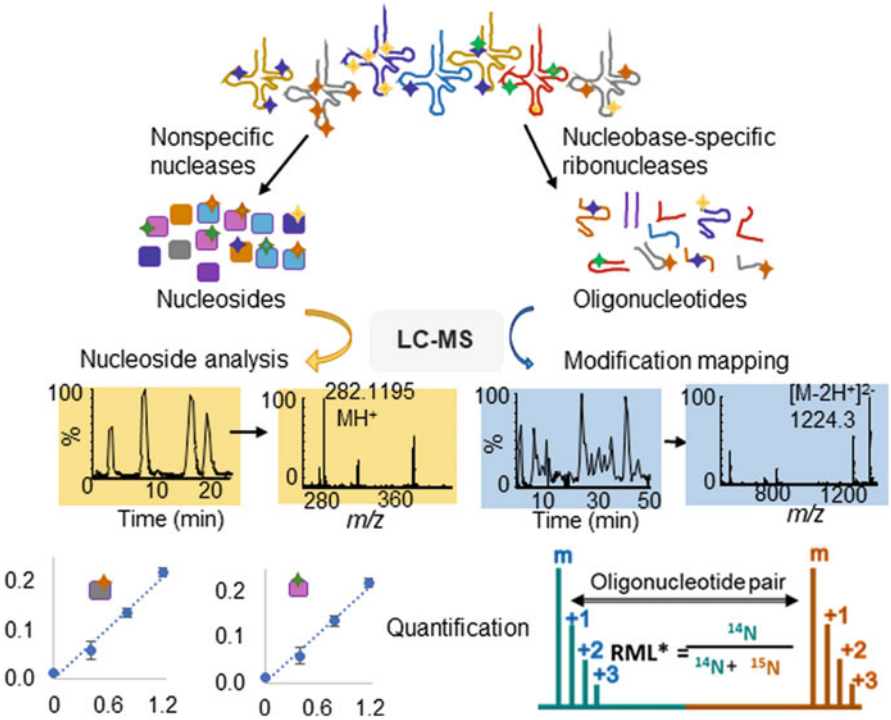
Abstract Quantitative characterization of the identity and location of nucleoside modifications (added posttranscriptionally) in RNA molecules is the first step for associating them to biochemical functions and disease. This chapter examines the

P. Thakur · M. Jora · R. Zhao · G. Parungao · S. Abernathy · P. A. Limbach · B. Addepalli (✉)
Rieveschl Laboratories for Mass spectrometry, University of Cincinnati, Cincinnati, OH, USA
e-mail: balasual@ucmail.uc.edu

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_21

ways to reduce the RNA complexity for accurate identification and quantification of modifications in transfer RNA (tRNA). These analytical methods are organized under three subsections. (1) *RNA purification*: We will contrast the classical and the affinity-based purification methods for tRNA enrichment. (2) *Modification detection and abundance*: Liquid chromatography coupled with mass spectrometry (LC-MS) methods involving high-energy collision dissociation and chromatographic retention time-based assignment will be compared. Ways to compute the relative abundance and application of multiple mass filters to document absolute amounts with external vs internal calibration for computing changes will be discussed. (3) *Site-specific quantification*: Relative quantification of modifications in the sequence context through isotope-labeling and signature oligonucleotide digestion products during RNA modification mapping will be examined. Thus, we will discuss a range of methods suitable for characterization of hypomodifications in tRNA (or any other RNA) in this chapter.

Graphical Abstract



Keywords LC-MS/MS · Posttranscriptional nucleoside modifications · Relative or absolute quantification · Oligonucleotide analysis · RNA modification mapping

1 Introduction

Transfer RNA (tRNA) serves as an adaptor molecule during translation through base-pairing interactions between mRNA codons and corresponding anticodons, so that the genetic code is matched with the right amino acid. These adaptor molecules contain 70–90 nucleotides with a characteristic cloverleaf secondary structure consisting of amino acid acceptor stem, D stem loop, T stem loop, variable loop, and the anticodon stem loop. These molecules assume an L-shaped tertiary structure, with an amino acid attached at 3′-end of sequence while decoding the genetic message through the anticodons (Fig. 1). The four building blocks of tRNA – A, G, C, and U – are posttranscriptionally modified with a myriad of chemical groups that are synthesized by site-specific enzymes (Boccaletto et al. 2018; McCown et al. 2020). These chemically diverse posttranscriptional nucleoside modifications (PTMs) include isomerization, single atom substitution, simple methylation, and complex hypermodifications catalyzed by multi-enzymatic biosynthetic pathways.

The research interest in chemical modifications on RNA has been labeled as epitranscriptomics which imparts regulation of gene expression at posttranscriptional level. The field of epitranscriptomics aims to identify the functionally relevant changes to chemical modifications and editing events in transcriptome. The marks of epitranscriptome may be categorized based on reference ribonucleotide and stages of modifications (Mathlin et al. 2020). Linking the biochemical effect to the ribonucleoside modification requires determination of the chemical identity and location in

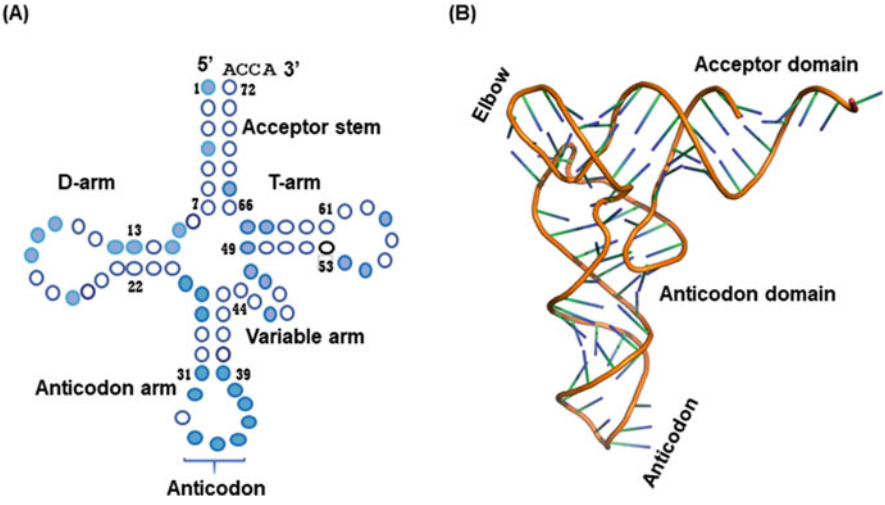


Fig. 1 Transfer RNAs exhibit higher density of posttranscriptional nucleoside modifications. (**a**) Clover-leaf model of tRNA representing various positions of the sequence that are posttranscriptionally modified across all tRNAs of *Saccharomyces cerevisiae*. Each circle represents a nucleoside, and filled circles denote the location of posttranscriptional modification. (**b**) L-shaped tertiary structure of yeast tRNAPhe (adapted from protein database entry PDB:1EHZ)

the RNA sequence. Chemical modifications found in the junctions between stems and loops of tRNA cloverleaf structure affect structural stability, aminoacylation, or fragmentation (Motorin and Helm 2010; Väre et al. 2017). Hypomethylation of cytosine (leading to loss of m^5C) in *NSun2* knockout mice exhibited higher levels of angiogenin-mediated tRNA cleavage, 5- tRNA fragment accumulation, and enhanced sensitivity to oxidative stress (Blanco et al. 2014). A wide diversity of modified nucleosides is found in the anticodon loop (positions 34 and 37). Their modification status can alter the aminoacylation specificity, codon recognition efficiency, and translation fidelity (Agris et al. 2017; Björk and Hagervall 2014), thereby controlling gene expression at mRNA translation.

Eukaryotic cells contain tRNAs transcribed from a large number of tRNA encoding genes ranging from 275 to 416 in yeast and humans, respectively (Chan and Lowe 2016). The tRNA molecules that specify the same amino acid but differ in anticodon sequence are referred to as isoacceptors. The tRNAs that have same anticodon but differ elsewhere in the primary nucleotide sequence are referred to as isodecoders. The diversity of isoacceptors and isodecoders leads to complex tRNA profiles, potential to alter translation function (Geslain and Pan 2010) and generate points for tissue-specific regulation in metazoans. Such an effect is manifested by increasing the translation rates of specific transcripts thereby increasing the protein amounts under stress conditions (Torrent et al. 2018) and even influencing the rates of mRNA decay (Hanson and Coller 2018). A model based on "translational control" of protein synthesis or adaptive translation was proposed where coordinated interactions between tRNA modifications or tRNA epitranscriptome and a select set of transcripts can schedule the synthesis of stress response proteins in prokaryotes and eukaryotes (Chan et al. 2018).

Consequences of PTM defects are prominent in mitochondria. For example, lack of carboxymethylamiomethyl $(cmnm)^5s^2U$ in mt-$tRNA^{Lys}$, mt-$tRNA^{Glu}$, and mt-$tRNA^{Gln}$ alters mitochondrial biogenesis and respiration (Wang et al. 2010). Similarly, lowered levels of taurinomethyluridine due to defective enzymes at position 34 of mt-$tRNA^{Leu}$ led to reduced translation, mitochondrial dysfunction, cardiomyopathy, and lactic acidosis (Kazuhito and Wei 2020). Likewise, several diseases are associated with the modifications in cytosolic tRNA (Chen et al. 2021; Haruehanroengra et al. 2020). However, out of 260 plus unique cytosolic tRNA sequences only 36 are characterized for their modifications in human cells. Recent improvements in RNA purification and analytical technologies are aimed at bridging this wide gap. In this chapter, we examine various approaches for characterization of hypomodifications or variable degree of modifications in tRNA through improved developments at multiple stages of analysis including sample preparation and LC-MS/MS data acquisition. We will discuss the suitability and potential limitations of these approaches for documenting modification levels.

2 RNA Purification

2.1 Total RNA Purification

RNA research kick started to high gear with the development of a single step extraction employing guanidine thiocyanate-phenol-chloroform mixture even from ribonuclease-rich sources such as the pancreas (Chomczynski and Sacchi 2006). This extraction solution is commercially available under specific trademarks such as TRIzol Reagent (Thermo Fisher) and TRI Reagent (Sigma-Aldrich). The purified total RNA consists of ribosomal RNA, tRNA, mRNA, snRNA, miRNA, etc. Characterization of resident PTMs in transfer RNA requires its further fractionation.

2.2 Transfer RNA Purification

Transfer RNA was referred to as soluble RNA because of its solubility in cold NaCl solution (1 M), but not the high molecular weight RNA (Zubay 1962). Alternatively, RNA with $<$100 nucleotides (that includes tRNA) are selectively separated by precipitation of large RNA ($>$100 nucleotides) with lithium chloride or PEG/sodium chloride (Nilsen 2012) or salt-based fractionation (Lee et al. 2007). Ion exchange chromatography is another option to isolate the tRNA pool from the cellular RNA mixture (Russell and Limbach 2013, Suzuki and Suzuki 2014). Another method of choice is size exclusion chromatography which allows size fractionation of noncoding classes of RNA (Chionh et al. 2013).

2.3 Enrichment of Specific tRNA Species

2.3.1 General Chromatography Methods

To obtain quantitative data of the PTMs, enrichment of target RNA is desirable. This is achieved by methods like countercurrent distribution based on differential solubility in two phases (Apgar et al. 1962), column chromatography involving ion exchange resin, partitioning between biphasic solvent system (Muench and Berg 1966), etc. Presence of hydrophobic base wybutosine facilitated single-step purification of yeast $tRNA^{Phe}$ through hydrophobic and ionic interactions with benzoyl groups of column material (Wimmer et al. 1968). Similarly, hydroxyapatite (Yamakawa et al. 1990) and reversed phase column chromatography methods (Thompson et al. 1983) were helpful in resolving up to 14 different species of tRNA. Resolution was improved further by two-dimensional electrophoresis where 47 different tRNA species of *E. coli* were resolved (Berg and Kurland 1997). However, this method required radiolabeling and may not have the capacity

to generate enough sample quantities in a suitable format for determining the modified sequence.

2.3.2 Affinity Chromatography Methods

Initial affinity chromatography methods were used to isolate uncharged or charged (with specific amino acid) tRNA species. A monoclonal anti-AMP antibody affinity matrix was used to isolate native *uncharged* tRNAs by affinity purification (Zhu et al. 1987). Elongation factor Tu (EF-Tu) forms a stable ternary complex with tRNA in the presence of GTP. EF-Tu-GTP readily recognizes and binds aminoacylated tRNA for subsequent separation by gel filtration (Klyde and Bernfield 1973) or after immobilization to a cyanogen bromide-activated Sepharose 4B (Fischer et al. 1982) or crosslinking to an Ni-nitroloacetic acid agarose (Ribeiro et al. 1995). Such approaches help isolate tRNA for differential characterization of uncharged or charged (aminoacylated) tRNA modification profiles.

Use of DNA-based affinity probes: Hybridization of tRNA sequence with a complementary DNA probe is another attractive affinity purification method. Success of these probes requires exclusion of anticodon region, modifications that interfere with the Watson and Crick base pairing, and long pyrimidine stretches. To expand the capability of this approach, chaplet column chromatography (CCC) was developed where multiple tRNA species were isolated in a single step. Here, each hybridization probe is immobilized on a solid support and packed in separate columns that are connected in series. The bulk tRNA preparation was passed through these columns one after another to isolate specific tRNA from each column (Suzuki and Suzuki 2007). This method led to the complete isolation of all 22 *Bos taurus* mitochondrial tRNAs for characterization of all the modifications (Suzuki and Suzuki 2014).

In general, the chaplet column chromatography approach is associated with high column back pressure which is a consequence of connecting multiple columns in series (Miyauchi et al. 2007). Therefore, a parallel and automated affinity chromatography called reciprocal circulating chromatography (RCC) was developed that retains the same capability of CCC (Miyauchi et al. 2007). Both CCC and RCC techniques were utilized in the isolation of all human mitochondrial tRNAs and subsequent characterization (Suzuki et al. 2020). However, its applicability for analysis of large tRNA population (>45 in *E. coli* or $>$ 260 in human cells) is unknown. Nevertheless, biotinylated deoxyoligonucleotide probes are highly suitable for small scale tRNA isolation, and use of tetraalkylammonium salts improves the efficiency further (Yokogawa et al. 2010). For samples containing >40 different tRNA species, separation of bulk tRNA into distinct groups based on their differential binding to the column stationary phase can be an attractive option. Here, each group can be analyzed for the census and the location of modifications in the nucleotide sequences.

3 Computing the Census of Posttranscriptional Nucleoside Modifications

Non-specific nucleases (e.g., nuclease P1, nuclease S1, phosphodiesterase I, benzonase) and phosphatases (e.g., alkaline phosphatase, antarctic phosphatase) are used to digest the RNA to ribonucleosides. While nucleases hydrolyze the phosphodiester bonds generating nucleotides, phosphatases remove the phosphate groups to yield ribonucleosides. However, modifications can be detected (confirming the presence) either as nucleotides or nucleosides. By virtue of the additional chemical group, the [^{32}P]-radiolabeled modified and canonical nucleotides can be separated by retardation factors in two-dimensional thin-layer chromatography (2D-TLC). By this technique, 70 different modifications are identified and R_f (retention factor) values are cataloged to serve as references for subsequent investigations (Grosjean et al. 2007).

3.1 Liquid Chromatography

High-performance liquid chromatography (HPLC or LC) is an effective technique for the separation of non-volatile canonical nucleosides and PTMs with adequate selectivity and resolution. Historically, reversed-phase liquid chromatography (RP-LC) methods (Pomerantz and McCloskey 1990, Werner 1993) have been used to resolve nucleosides based on their differential interaction with the C18 stationary phase (column material) under defined mobile (liquid solvent(s)) phase conditions. In general, polar species, such as pseudouridine (Ψ) and dihydrouridine (D), exhibit shorter retention times, whereas the hydrophobic species, such as 2-methylthio-N6-isopentenyladenosine (ms^2i^6A) and 5-methylaminomethyl-2-geranylthiouridine (mnm^5ges^2U), exhibit longer retention times during RP-LC separations. On the other hand, hydrophilic interaction liquid chromatography (HILIC) can improve the detection of compounds that are retained for shorter times on RP columns. Here, the polar compounds such as uridine and their derivatives can exhibit increased interaction with the hydrophilic solid phase in the presence of higher content of organic mobile phase thereby improving the retention and resolution (Zhou et al. 2014).

The resolution and analysis speed can be improved further with ultra-performance liquid chromatography (UPLC) and smaller sample volumes. UPLC uses smaller particle (2–1.7μm compared to >3μm in HPLC columns) in the column stationary phase (packing material) and operates at pressures 6000–15,000 psi (compared to 2000 to 4000 psi). They provide reduced band broadening, increased sensitivity, and higher chromatographic peak resolution, thus enhancing separation performance (resolution between individual peaks) or peak capacity (maximum theoretical number of peaks resolved within the gradient time) (Gilar and Neue 2007).

3.2 Mass Spectrometry-Based Detection

Liquid chromatography when coupled with mass spectrometry (LC-MS) provides improved detection as it complements the retention time with the mass values. Modified ribonucleosides exhibit a mass shift (Fig. 2) reflecting the additional chemical group attached to the nucleobase or ribose sugar compared to the canonical ribonucleosides (C, U, A, and G). For instance, in case of methylation, a + 14 Da shift is observed due to the replacement of hydrogen (-H) by a -CH_3 group. Thus, measurement of mass values by mass spectrometry (MS) allows a direct readout of any modifying group added to the canonical ribonucleoside (Pomerantz and McCloskey 1990). In fact, almost all the >170 known PTMs were either discovered

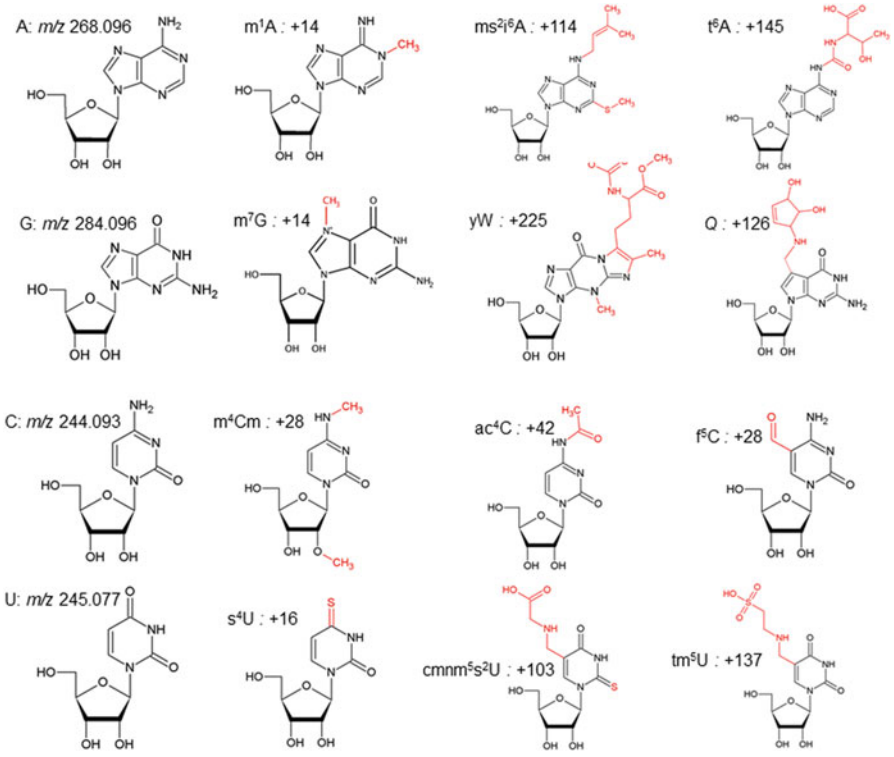


Fig. 2 A few examples of posttranscriptional nucleoside modifications and their corresponding mass shifts when compared against their canonical ribonucleosides. A, adenosine; m^1A, 1-methyladenosine; ms^2i^6A, 2-methylthio-N^6-isopentenyladenosine; t^6A, N6-threonylcarbamoyladenosine; G, guanosine; m^7G, 7-methylguanosine; yW, wybutosine; Q, queuosine; C, cytidine; m^4Cm, N^4, 2'-O-dimethylcytidine; ac^4C, N^4-acetylcytidine; f^5C, 5-formylcytidine; U, uridine; s^4U, 4-thiouridine; $cmnm^5s^2U$, 5-carboxymethylaminomethyl-2-thiouridine; tm^5U, 5-taurinomethyluridine. The modifying groups are shown in red-colored font

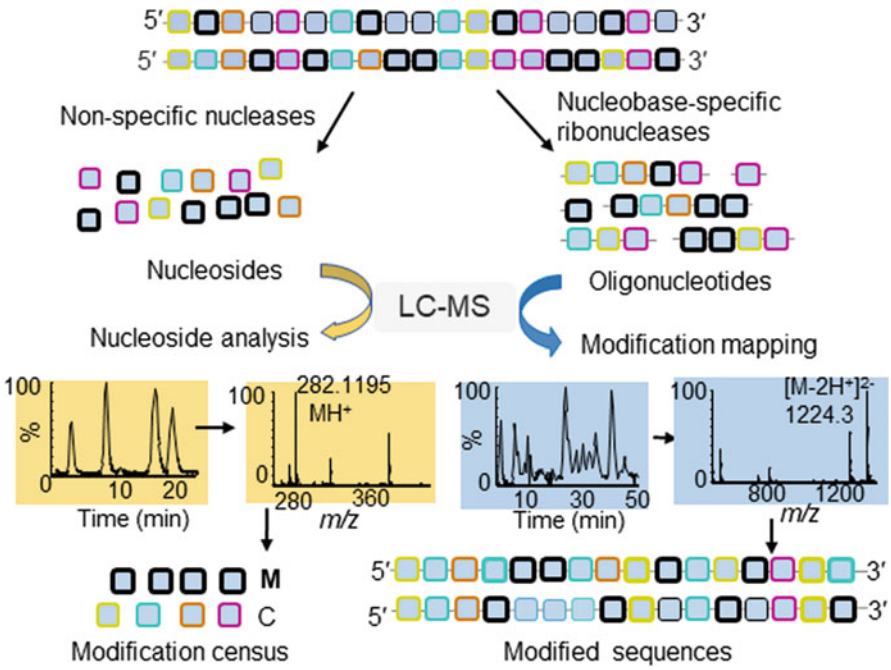


Fig. 3 Workflow for characterization of posttranscriptional nucleoside modifications in RNA. Canonical ribonucleosides, A, G, C, and U are represented by color-coded rectangles (C) and the modified nucleosides indicated with thick border (M). Characterization of modified RNA involves two types of analysis. Ribonucleoside analysis provides information about the types and abundance of modifications. Oligonucleotide analysis locates the position and frequency of modification in the sequence

or structurally characterized for their identity and sequence location by LC-MS. A general workflow of ribonucleoside modification analysis is shown in Fig. 3**.**

Following chromatographic separation, protonated nucleosides in the liquid phase are transferred to the gas phase for mass analysis through electrospray ionization (ESI) (Kebarle and Tang 1993). The nucleoside ions [MH^+] in the gas phase are transmitted to the mass spectrometer, and their mass-to-charge (*m/z)* values are recorded for their relative intensities in positive ion mode. The mass spectrum is depicted as a plot between *m/z* values and their relative intensity (or abundance) at given time. Different mass analyzers with different mass resolution capabilities (ability to separate two adjacent narrow mass spectral peaks) can be used (Haag 2016). Low-resolution instruments, such as ion-traps or triple quadrupole (QqQ) mass spectrometers, provide unit resolution and provide average mass values of nucleoside (weighted average calculated by multiplying the relative isotope abundance by the atomic masses and summing the products). On the other hand, high-resolution instruments, such as orbitrap, time-of-flight (TOF), and Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers can resolve ions made from

monoisotopic atoms (molecular ion that is made from the most abundant isotopic atoms for each element in the compound) from those with the naturally occurring heavy isotopes. Besides mass resolution, high-resolution instruments also provide high mass accuracy (defined as the difference between the theoretical and measured *m/z* and is measured in mDa (u or mTh) or ppm) (Brenton and Godfrey 2010).

Tandem mass spectrometry: The presence of a modifying chemical group is reflected on the mass shift of molecular ion (MH^+) compared to its canonical nucleoside (Fig. 2). Modifications that exist as isomers—such as Ψ and U, as well as 2'-O- (Cm), 3- (m^3C), N^4- (m^4C), and 5-methylcytidine (m^5C)—require additional criteria for their unambiguous identification. One of these criteria is the observed retention time during chromatographic separation (Boccaletto et al. 2018). However, since this is an error-prone extrinsic feature, tandem mass spectrometry (MS/MS) is traditionally used as well. During MS/MS analysis, a specific *m/z* value is selected for fragmentation by a process called collision-induced dissociation (CID). Under CID, the molecular ion (MH^+) is accelerated by an electric field so that bond breakage can occur due to collisions with a neutral gas (generally He, Ar, or N_2). With a few exceptions (e.g., Ψ), CID-based fragmentation of ribonucleosides result in the nucleobase ion (referred to as BH_2^+) following cleavage of the *N*-glycosidic bond that connects the ribose sugar and nucleobase. This characteristic fragmentation helps differentiate ribonucleosides from background ions, as well as delineating the site of modification to the ribose or nucleobase, owing to the neutral loss of a modified or unmodified ribose sugar (Jora et al. 2019; Kellner et al. 2014a, b; Pomerantz and McCloskey 1990). Thus, the three detected features, i.e., retention time, acquired *m/z* values of MH^+, and BH_2^+ ions, serve as analytical figures of merit for assigning the LC-MS signal to RNA modification (Pomerantz and McCloskey 1990; Russell and Limbach 2013). Indeed, the LC-MS/MS is considered as the gold standard for analytical characterization of modified nucleosides (Basanta-Sanchez et al. 2016, Heiss et al. 2017; Russell and Limbach 2013; Santos and Brodbelt 2021; Su et al. 2014; Thüring et al. 2016).These approaches can be tailored for both *targeted* (selected set) and *untargeted* (global profiles) characterization.

3.3 Positional Isomer Detection

Even though most ribonucleosides are characterized by the above approach, the unambiguous characterization of a few positional isomers (including isomers that exhibit same mass but differ in the position of modification group on nucleobase) may still be challenging. Figure 4 illustrates the three chromatographic peaks observed for *m/z* 258.109 (methylcytidines) during the analysis of *S. cerevisiae* tRNA on an orbitrap mass spectrometer. While the peak at 5.0 min may be characterized as Cm due to its characteristic CID-based MS/MS spectrum (neutral loss of 146 Da for 2'-O-methylribose sugar), other two peaks at 2.5 and 3.7 min exhibit identical *m/z* pattern for both MH^+ and BH_2^+ ions. Application of retention

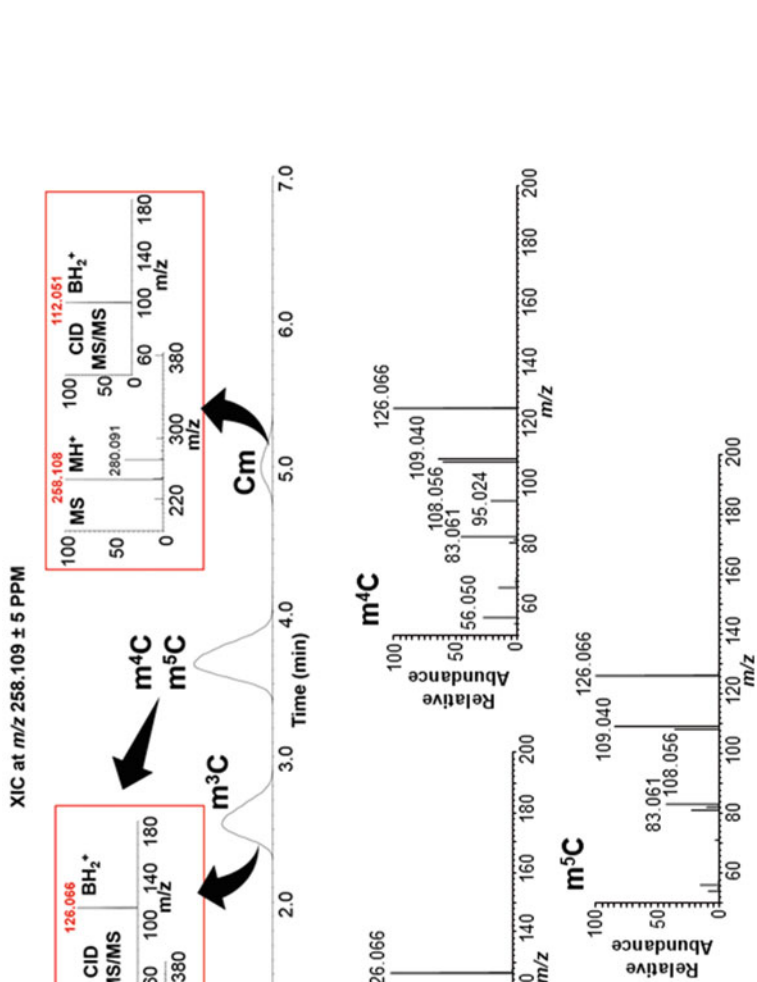


Fig. 4 HCD-MS/MS-based differentiation of m^4C and m^5C positional isomers. (**a**) Reversed phase liquid chromatographic elution pattern of methylated cytidines. (**a**) While m^3C and Cm are resolved from other isomers of methylated cytidine and distinguished by traditional LC-MS/MS through collision-induced dissociation, m^4C and m^5C coeluted under a single peak and indistinguishable by CID-MS/MS due to identical m/z patterns of molecular ion (MH^+) and nucleobase ions (BH_2^+). (**b**) Higher energy collision dissociation-based differentiation of m^4C and m^5C through altered abundance of fragment ions (*m/z* 108.056 vs *m/z* 109.040) and differences in unique fragment ion abundance (*m/z* 95.024 for m^4C but not m^5C)

time criterion would denote the peak at 2.5 min to m^3C. However, both m^4C and m^5C exhibit the same retention time under these HPLC conditions, necessitating additional experiments to confirm the nature of the third peak. Selective chemical derivatization with osmium (Os) (Meng-Dan Lan 2019) or alternative chromatographic methods such as HILIC (Zhao et al. 2013) or multistage fragmentation (MS/MS/MS or three stage MS analysis) (Jensen et al. 2007) or cyclic ion mobility (cIM)-based resolution (Kenderdine et al. 2020) could be required to identify the chemical nature. The ion mobility fractionates gas phase ions in a drift tube (generally fitted between quadrupole and time-of-flight mass analyzers) based on their size, charge, and shape (Gabelica and Marklund 2018). The position of modification within the nitrogenous base could alter the shape of the molecule resulting in its separation from its isomeric compounds. Differential ion mobility can also be achieved by exposing ions to high and low electric fields in gas flow at atmospheric pressure (field asymmetric waveform ion mobility spectrometry—FAIMS) (Kolakowski and Mester 2007).

While CID is the most traditional tool used for MS/MS studies, higher-energy collisional dissociation (HCD) can also be employed for structural characterization. As its name suggests, higher energy imparted on the precursor ion can also dissociate the nucleobase ion (BH_2^+) in the same MS/MS analysis without requiring third stage. Thus, information generally obtained in MS3 (MS/MS/MS) can be obtained in MS2 (MS/MS) analysis. The hardware associated with HCD can trap and detect low molecular weight fragment ions that enable the differentiation of positional isomers (Jora et al. 2018). As illustrated in Fig. 4, in addition to the characteristic BH_2^+ at *m/z* 126.066 corresponding to methylated cytosine, presence of additional unique fragment ions and their abundance can serve as fingerprint to distinguish positional isomers, m^3C, m^4C, and m^5C. Thus, cases of chromatographic co-elution (m^4C and m^5C) can be resolved through the HCD MS/MS spectra of fragment ions revealing the nature of coeluting species. The positional isomer, m^3C, although chromatographically resolved, exhibits a different pattern of MS/MS spectra or fingerprint. In other words, this could serve as additional criterion for confirming its presence and reduce the retention time dependence for assignment.

3.4 Discovery of New Modifications

Observation of an unknown retention time and/or mass value in the LC-MS analysis can represent the existence of a new modification. However, signal from contaminants in the sample or chromatographic solvents must be ruled out to prove its presence and identity. Moreover, artificial adducts can also be generated during sample preparation (Jora et al. 2020); therefore, systematic characterization of detected signals is required to confirm the existence of a new modification (Dal Magro et al. 2018). Metabolic labeling of RNA through stable isotope labeling is one sample preparation technique that can resolve the ambiguity and confident assignment (Kellner et al. 2014a, b). Comparative genome sequencing can also be

considered along with the LC-MS/MS analysis for identification of novel modifications and modifying enzymes (Zallot et al. 2017). Moreover, the fingerprinting nature of HCD has opened the possibility of finding similarities in fragmentation patterns to identify novel modifications in other organisms (Yu et al. 2019).

4 Quantification of PTMs

Quantitative information determines the modification status of a tRNA species whether complete or incomplete or in other words hypomodification. The extent of tRNA modification in all copies of the same sequence is not constant. It can vary in response to growth, nutrient or oxygen levels, cellular, and environmental factors (Buck and Ames 1984; Chionh et al. 2016; Emilsson et al. 1992; Sakai et al. 2016; Shepherd and Ibba 2015). Lack of the 5-methylaminomethyl-2-thiouridine (mnm^5s^2U) modification at wobble position of $tRNA^{Glu}$, $tRNA^{Lys}$, and/or $tRNA^{Gln}$ caused suppressor phenotype of temperature-sensitive *E. coli* (Isak and Rydén-Aulin 2009). *Loss of ASL modification, 5-methoxycarbonylmethyl-2-thiouridine, induced starvation responses in exponential growth phase of Saccharomyces cerevisiae (Bruch* et al. 2020*).* Similarly, queuosine containing tRNAs in cancer cells were incompletely modified due to lower activity of TGTase which catalyzes the base exchange from guanine to queuine (Pathak et al. 2005). However, Q modification in human cells also protects $tRNA^{His}$ and $tRNA^{Asn}$ against cleavage by ribonucleases (Wang et al. 2018). Hypoxia increased wobble modification, cmo^5U, levels in tRNAThr(UGU), thereby, enabling the translation of ACG-codon enriched transcripts in mycobacteria (Chionh et al. 2016). *Thus, it is of biochemical interest to monitor the levels of modifications to understand their links to the metabolic status of cell.*

Quantification criteria: Measurements of relative changes generate useful information for intra-lab comparative analysis (e.g., gene knockouts, cellular stress). Computing absolute amounts provides a definitive insight into the molecular dynamics of modifications. Such strategies include (a) stable isotope labeling-based approaches (precise, but experimentally challenging) and (b) label-free methods (easier to be implemented, but with a larger associated error). Compared to other quantification method like UV or fluorescence, response factors (a measurement where a change in signal reflects a change in quantity) in mass spectrometry can be less robust. Factors that can affect MS response include sample preparation (matrix effects caused by salt load, contamination, coelution of unknown components, etc.), instrumental settings, and physicochemical parameters of the analyte/solvent (Kellner et al. 2014a, b). The quantification methods need to account for these factors to achieve high precision (reproducibility of measurements in replicates) and accuracy (observed measurement closer to the true value). The analytical figures of merit for quantification include defining the set of nucleoside analytes that are being measured, detection specificity (ability to distinguish analyte from the rest in

the sample), limits of detection (LOD), limits of quantification (LOQ), and a validation that a change in signal indeed reflects the change in quantity.

4.1 Relative Quantification Strategies

For relative quantification, the peak area of each modified nucleoside ion (MH^+) (referred to as extracted ion chromatogram (XIC) for specific *m/z* value) and base ion (BH_2^+) is compared in different samples. A study from our laboratory showed that the average percent RSD (relative standard deviation or how data is spread out from the mean) in peak areas was 5.9% and the variability ranged from 1.0 to 12.4% (Russell and Limbach 2013). To overcome the signal fluctuations caused by instrument parameters over time, target signal abundance is normalized against that of canonical nucleoside (i.e., m^5C to C) along with the amount of input RNA (Laxman et al. 2013). Normalization of the target modification with the spiked internal standard (IS) can also be used to compare the fold changes in modification levels between treated and untreated controls. For example, [^{15}N]5-deoxyadenosine as internal standard to quantify reprogramming of tRNA modifications under different environmental stresses was employed (Chan et al. 2010, 2012; Su et al. 2014). It would also give instant information while screening samples for enzymes knockout mutants (Kang et al. 2017; Meyer et al. 2020) or dose-dependent responses (Sun et al. 2018).

4.2 Absolute Quantification Strategies

Measurement of the exact levels of modification among several samples or in each RNA mixture or quantification of relative occupancy at specific positions of pure tRNA species can be denoted in absolute amounts (Kellner et al. 2014a, b). These measurements yield moles of modification per mole of RNA species. For this type of analysis, a triple quadrupole mass spectrometer is the most widely used instrument. They operate at <4000 *m/z* with scan speeds reaching 1000 *m/z* or more but exhibit unit mass resolution with mass accuracy around 0.1 *m/z* unit. This mass analyzer consists of three sets of four parallel metal rods (Q1, Q2, and Q3), where the first and third quadrupoles (Q1 and Q3) allow transmission of a narrow band of specific *m/z* values along the axis. The Q2 maintains low pressure of inert gas to induce collision-induced dissociation of the ions transmitted by Q1. The Q3 would monitor for the presence of specific product ions; thus, panels of precursor/product ion pairs can be created to detect many targets by a process referred to as multiple reaction monitoring (MRM) (Dalluge et al. 1996). In the MRM, two mass filters one at the precursor ion level and the other at the detection of select fragment ions ion transitions (T1, T2. . .Tn) are applied (Fig. 5). These double mass filters reduce the background noise and increase sensitivity and selectivity, thereby achieving lower limits of detection

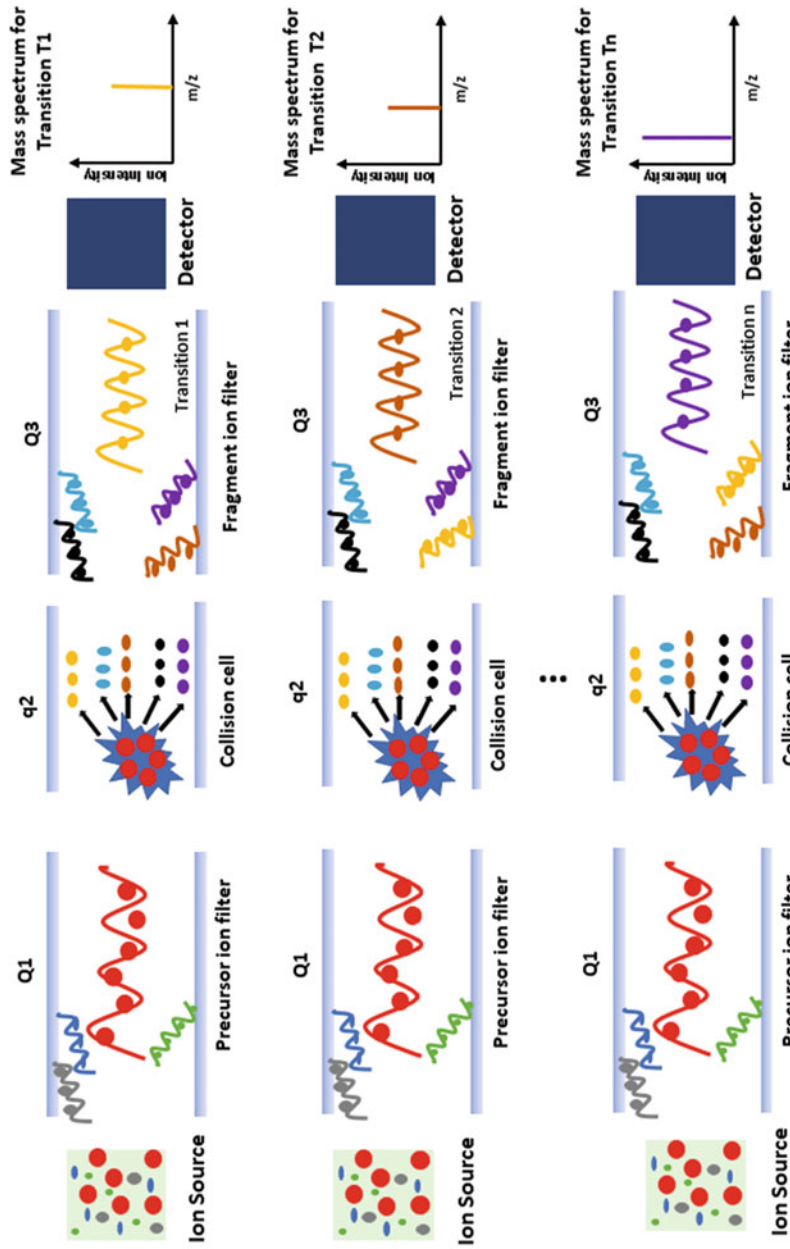


Fig. 5 Scheme of multiple reaction monitoring (MRM) mode on triple quadrupole mass spectrometer. The first quadrupole (Q1) is set to monitoring ions of a specific m/z value (precursor ion). The selected precursor ion is then fragmented in the second quadrupole (q2, collision cell). The third quadruple (Q3) scan for the presence of selected fragment ion(s). Depending on the experimental setup, one or more transitions (T1, T2, ...Tn) for a given precursor ion can be sequentially monitored

(LLOD), lower limits of quantification (LLOQ), and wider linear range for quantification (Pitt 2009).

The MS response may vary from one modification type to another based on their chemical nature. In such cases, normalization by a different molecular ion may not be ideal. In such cases, calibration curve for each modification become essential to correlate the MS signal intensity to the exact amount of target in the sample. Therefore, the biggest challenge in absolute quantification is the generation of appropriate calibration curve reflecting the quantitative response by MS. Standard solutions of a synthetic and highly pure modified nucleosides are employed to make either external calibration (Chen et al. 2013a, b) or spike-in measurements (Contreras-Sanz et al. 2012).

4.2.1 Quantification Using External Calibration

An external calibration curve can be plotted from signal response against the serial dilutions of defined concentration injected into the LC-MS system. The absolute amount of each modified nucleoside can be determined from the linear dynamic range of the calibration curve generated from diluted standards. For best reproducibility, technical replicates of the calibration curve need to be performed: one set before, and another set after the sample data acquisition. If a large set of samples are being analyzed, another set of calibration should be performed during the middle of the runs. Use of response average from technical replicates can reduce the errors of quantification from instrument fluctuations (Traube et al. 2019).

Apart from MRM mode, measurement of HCD-fragment ion abundance ratios could also be used for quantification of coeluting ions that exhibit identical mass values (e.g., m^4C and m^5C). A calibration curve can be prepared using mixtures of known amounts of each nucleoside in a fixed total amount (e.g., 25 ng total amount consisting of m^4C and m^5C varying from 10:0 to 0:10 ratios, respectively). Here, the observed ratio of fragment ion abundances (*m/z* 109.039 and 108.055) is plotted against the amount of each isomer used for each mixture (Fig. 6). Similarly, the abundance of unique fragment ion *m/z* 95.023 in the MS/MS spectra can also be used to prepare a calibration curve reflecting the amount of m^4C in the sample. This would provide the response factor for quantification of coeluting isomers. Use of external calibration is relatively simple and less expensive. This advantage, however, is offset by the requirement of several sets of data acquisition increasing the analysis time. This also does not account for the sample background (or matrix effects) that influences competitive ionization and ion suppression. Such matrix effects can produce varied signal response between external standard and sample analyte resulting in misleading interpretations (Kimura et al. 2020).

4.2.2 Quantification Using Internal Calibration

One way to develop internal calibration is spiking of unnatural modification into the sample to use it as an internal standard (IS). The MS response of standard analyte

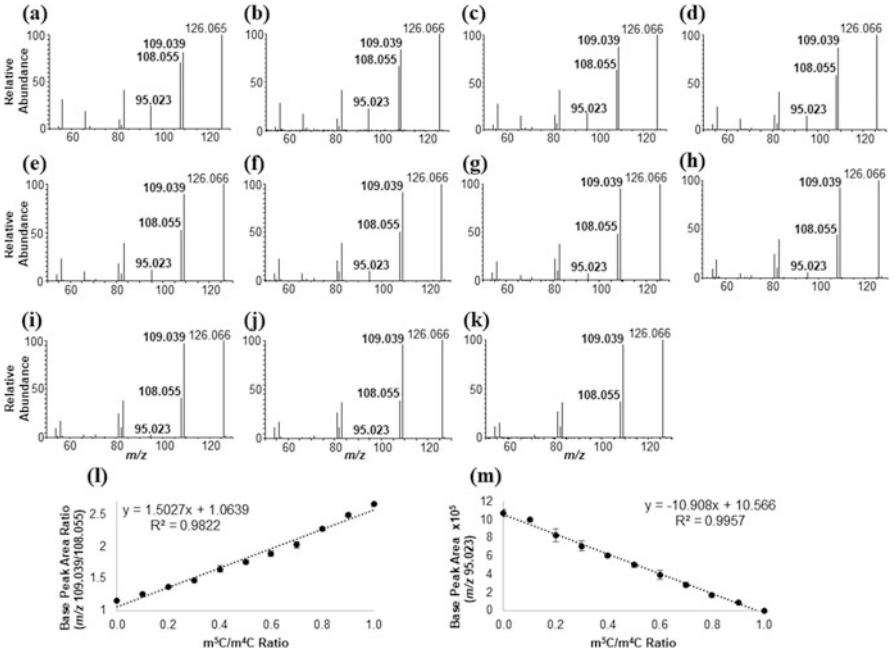


Fig. 6 Generation of response factor calibration curves for quantification of coeluting positional isomers. Representative HCD MS/MS (acquired at collision energy 80 a.u.) for m^4C:m^5C mixture with a fixed total amount of 25 ng is shown. They contained the following compositions: (**a**) 25.0, (**b**) 22.5:2.5, (**c**) 20:5, (**d**) 17.5:7.5, (**e**) 15:10, (**f**) 12.5:12.5, (**g**) 10:15, (**h**) 7.5:17.5, (**i**) 5:20, (**j**) 2.5:22.5, (**k**) 0:25 ng. Calibration curves generated with base peak areas for the ions at *m/z* 109.039 and 108.055 (bold font) (**l**), and the base peak area for m4C-specific unique fragment ion at *m/z* 95.023 (bold font) are shown (**m**)

dilutions were normalized against fixed amount of IS added to each dilution. This normalized response of each dilution was plotted against the concentration to generate the calibration curve. For example, 32 individual nucleosides were quantified by making each calibration with fixed amount of isotopically labeled [^{13}C][^{15}N]-G added as an IS (Basanta-Sanchez et al. 2016). By normalizing the external calibration solutions and samples spiked with IS, the instrument fluctuations and signal suppression were removed. Thus, use of single nucleoside species as internal standard could quantify multiple modifications in a given sample (Su et al. 2014; Yan et al. 2013). However, accurate response factors can only be derived from isotopomers of each specific modification (Dalluge et al. 1996). The isotopomer will have similar response factor and fragmentation pattern. Therefore, an impressive number of stable isotope-labeled internal standards (SIL-IS) were developed for quantification of corresponding nucleoside modifications (Brückl et al. 2009; Thumbs et al. 2020) through a technique called isotope dilution MS. The recommended stable isotopes for labeling are [^{13}C] or [^{15}N] because no retention time (RT) shifts were observed for ^{12}C/^{13}C and ^{14}N/^{15}N substitutions. By using this

approach, they were able to quantify modified tRNA nucleosides of *E. coli*, pork liver, HeLa, and HCT-116 cell lines (Brückl et al. 2009), murine and porcine tissues (Brandmayr et al. 2012), and distribution of hypermodified Q-nucleosides in mouse tissues (Thumbs et al. 2020).

4.2.3 Biosynthetic Isotopomers as Internal Standard (IS)

Preparation of isotope labeled nucleoside standards requires a significant amount of organic synthetic work, and it is not possible to synthesize all naturally occurring nucleoside modifications (Boccaletto et al. 2018). Thus, the lack of available stable isotope-labeled internal standard (SIL-IS) can be a big limitation for most laboratories. To overcome these limitations, stable heavy isotopes were incorporated into RNA through metabolic labeling, and the quantities of modifications were determined by a combination of external calibration of unlabeled standard and spiking of the biosynthetic SIL-IS. This method was highly precise with a broader linear dynamic range (Kellner et al. 2014a, b) with ability to overcome detector saturation at higher concentrations. A modification of this approach enabled differentiation of metabolic status through inclusion/exclusion of isotope labels for both bacteria and eukaryotes, an approach referred to as NAIL-MS (nucleic acid isotope labeling coupled mass spectrometry) (Borland et al. 2019). Application of this type of pulse-chase NAIL-MS experiments enabled them to study the fate of RNA modifications (Heiss et al. 2017, Reichle et al. 2019) including hypomodifications.

5 Mapping PTMs to RNA Sequences

During hydrolysis of RNA to nucleosides, information about their location and co-occurrence in the sequence is lost. However, that information can help identify the modifying enzyme, develop deeper understanding of the biochemical function, and understand the phenotypic effect such as molecular pathology (Haruehanroengra et al. 2020; Lant et al. 2019). Approaches to determine sequence location should retain sufficient sequence context to position the modification to a specific nucleotide in the sequence or limit the region where modification can be found (Limbach and Paulines 2017). Two major approaches have been used to address this challenge. One is next-generation sequencing-based technique referred to as RNA-seq, which is a high-throughput approach, and another is LC-MS/MS for obtaining both qualitative (identity and location of modification) and quantitative (is specific modification present in every copy of tRNA) information.

5.1 High-Throughput Methods

These methods are capable of global or transcriptome wide detection of modifications in the sequence. RT-PCR (reverse transcriptase reaction combined with polymerase chain reaction) is central to NGS-based RNA-seq technology where the location of modification was identified as "RT stops" or base misincorporation (Motorin et al. 2007). The quantification is done by documenting the number of PCR cycles required to reach certain amount of amplified modified RNA-specific product (also referred to as qRT-PCR). Quantitative RT-PCR can be combined with other sample preparation treatment such as treatment of RNA with γ-toxin from yeast, *Kluyveromyces lactis*, to quantitate the number of tRNA transcripts that contained mcm^5s^2U modification (Lentini et al. 2018). However, this analysis exhibited high variability in its output signal.

RT-PCR in combination with next-generation sequencing was developed for locating N6-methyladenosine or m^6A (Meyer et al. 2012), 1-methyladenosine or m^1A (Dominissini et al. 2016), pseudouridine (ψ) (Zhang et al. 2019), 5-methylcytidine (m^5C) (Squires et al. 2012), 5-hydroxymethylcytidine (hm^5C) (Delatte et al. 2016), ribose methylations (Nm) (Birkedal et al. 2015), N4-acetylcytidine (ac^4C) (Thomas et al. 2019), 7-methylguanosine (m^7G) (Enroth et al. 2019), 3-methylcytidine (m^3C) (Marchand et al. 2018), and inosine (I) (Cattenoz et al. 2013) either through RT stops or specific chemical reactivity or both. Thus, a few common themes (Fig. 7) were identified for global profiling of modifications. They include enhanced RT stops through selective chemical treatment, differential sample analysis, and enrichment step involving antibodies for background reduction, where targeted modification can be characterized individually. Similarly, small groups of modifications (m^5C, ψ and m^1A) can also be quantified by bisulfite sequencing at single base resolution (Khoddami et al. 2019).

RNA-seq technologies in general are not very well compatible with highly modified sequences such as tRNA, where reverse transcriptase can terminate due to inherent structure. Employment of enzymes that remove methylation (m^3C, m^1A, m^1G and $m^{2,2}G$) marks before RT reaction could improve the sequencing (Cozen et al. 2015) and quantification of methylations (Clark et al. 2016). This method uses dealkylating enzyme *A*lkB to *r*emove the *m*ethylation marks (ARM-seq) before reverse transcription and library preparation for sequencing. Comparison of treated and untreated samples could provide information about the differential abundance of methylations. ARM-seq greatly increased the abundance and diversity of reads from tRNAs in widely divergent model organisms. An extension of demethylation RNA-seq was used to identify the fraction of tRNAs that are charged with cognate amino acid (Evans et al. 2017). Alternate approach to overcome the methylation block is to use manganese (Mn^{2+}) ions in RT reaction, which increases nucleotide skipping events at modified positions (m^1A, $m^{2,2}G$, m^1G, and m^3C) providing enhanced read-through capability in RNA-seq (Kristen et al. 2020). Nevertheless, not all modifications are removed by demethylases, and the strategies to select seed

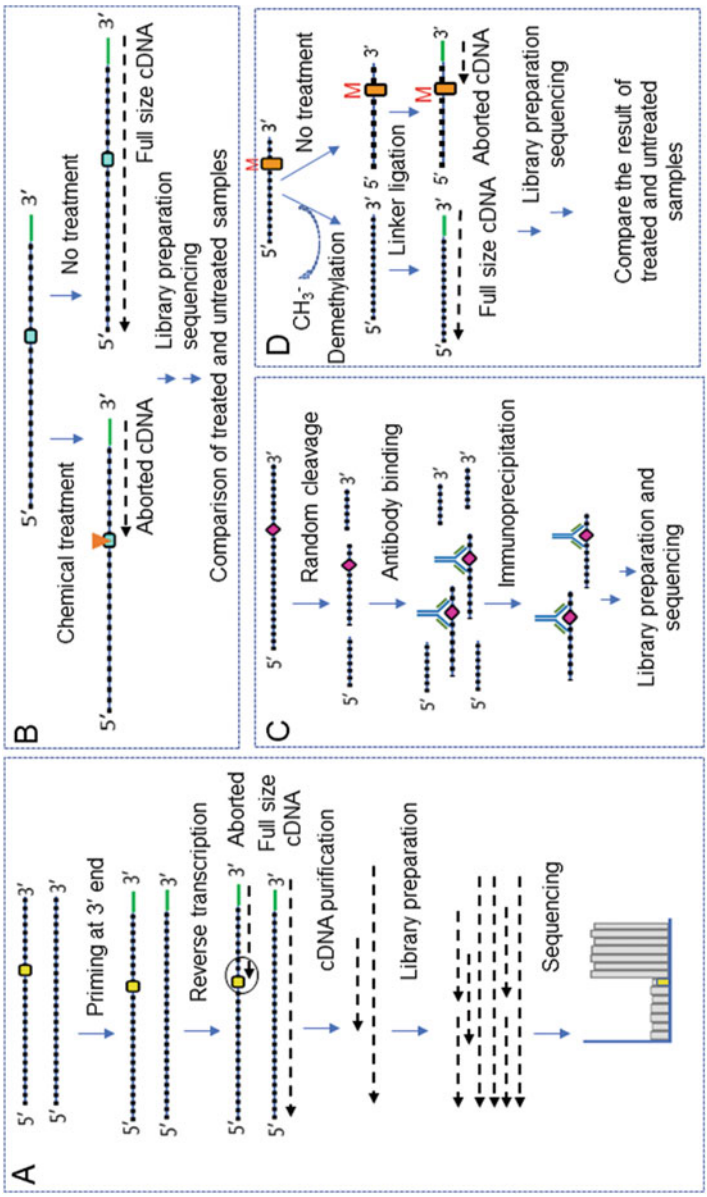


Fig. 7 Schematic representation of RNA-sequencing approaches based on next-generation sequencing (NGS). (**a**) Outline of reverse transcription-based next-generation sequencing of RNA. (**b**) RNA-seq following modification-specific chemical treatment of RNA pool. C- RNA-seq of the RNA pool enriched for specific modified nucleoside using modification specific antibody. D—ARM-Seq approach for identification of select set of methylations

sequences for modification analyses could influence the de novo identification of tRNA modification sites, thereby introducing another layer of bias in RNA-seq data.

Sequence similarity between different tRNA species, modified nucleotides, and difficulty to distinguish precursor and mature tRNA can also result in quantification biases (Zheng et al. 2015). A bioinformatic pipeline was developed to overcome biases for isodecoder-specific tRNA gene contribution in tRNA gene expression analysis (Torres et al. 2019). This approach referred to as quantitative mature tRNA sequencing (QuantM-tRNA seq) could identify CNS (central nervous system)-specific expression pattern of isodecoder tRNAs in mouse tissues (Pinkard et al. 2020). Lack of specificity for antibody can lead to false positives and negatives due to systematic errors (repeatable error associated with faulty experimental design or equipment or calibration) and base-calling (assigning nucleobases to chromatogram peaks). Moreover, only a handful modifications can be detected and quantitated, that too mostly one at a time, by these high-throughput methods.

5.2 *Sequence Profiling to Identify Modification Location by LC-MS*

Mass spectrometry-based sequencing of nucleic acids was enabled with the development of new ionization techniques—matrix-assisted laser desorption/ionization (MALDI) (Douthwaite and Kirpekar 2007). MALDI-MS uses time-of-flight mass analyzer and is highly suitable for detecting analytes with high *m/z* values. However, it has limited applicability due to its inability to couple with chromatography. Electrospray ionization developed in the 1980s (Fenn et al. 1989) can be easily coupled with chromatography and is the widely used approach for locating the modifications in the sequence. Sequencing is accomplished by bottom-up approaches where RNA is digested with nucleobase-specific ribonuclease and the resulting oligonucleotide digestion products are sequenced to locate the modifications. This is generally referred to as RNA modification mapping (Kowalak et al. 1993).

Liquid chromatography: To facilitate the liquid chromatographic separation of oligonucleotides, an alkylamine such as triethylamine (TEA) is used to mask negative charge on phosphodiester backbone generating a pseudo-neutral and its retention on a C18-reverse phase column. The longer the oligonucleotides, the greater the retention, thereby requiring higher organic content for elution. Although TEA addition ensures chromatographic retention, its removal requires addition of another organic modifier such as 1,1,1,3,3,3-hexafluoro-2-isopropyl alcohol (HFIP). Upon electrospray, the droplet evaporates rapidly due to the low boiling point and high Henry's law constant of HFIP. This leads to localized increase of pH in the droplet due to TEA presence. When the pH reaches pKa of alkylamine, the TEA dissociates from phosphate allowing the eluted oligonucleotide ions to be sampled into the mass spectrometer. This type of separation also referred to as ion-pairing

reversed phase liquid chromatography (IP-RP-LC) is widely used for separation of oligonucleotides (Apffel et al. 1997). Use of ion-pairing agents provides high chromatographic performance (Gilar et al. 2002), reduction of charge state (number of charges) (Chen et al. 2013a, b), and alkali metal adducts (Greig and Griffey 1995), but adversely affecting MS sensitivity or causing ion suppression. Hydrophilic interaction chromatography uses hydrophilic stationary phase to separate oligonucleotides without the need of ion-pairing reagents (Alpert 1990). Recently, polymer-based stationary phase containing modified diol phase showed promise for separation of oligonucleotides by HILIC chromatography (Lobue et al. 2019a, b). Other types of chromatography and mobile phase additives for improved separation of oligonucleotides are reviewed elsewhere (Santos and Brodbelt 2021).

Just like nucleoside analysis, two stages of mass analysis are performed (referred to as tandem mass spectrometry – MS/MS) to locate the modification in the sequence. The first stage involves selection of the oligonucleotide anion $[M\text{-}nH^+]^{n-}$ (n is the number of deprotonations, also known as precursor ion), and the second stage involves imparting energy and dissociation into fragment ions due to collisions with neutral gas. Dissociation occurs due to cleavage of phosphodiester bond or glycosidic bonds (between a nucleobase and ribose sugar). Out of 4 possible positions of cleavage on diester bond (Fig. 8a), cleavage of P-O bond is observed resulting in sequence-informative fragment ions, c- and y- ion series. The c-fragment ion series contain the 5′-end, while the y-fragment ion series share the 3′-end of the oligonucleotide precursor ion. Just like the mass shift of canonical nucleoside in nucleoside analysis, oligonucleotides and specific set of fragment ions exhibit corresponding mass shift compared to the unmodified counterparts. One such example of mass shift due to methylation is shown in Fig. 8b in comparison with the unmodified oligonucleotide. Once the mass shift is registered at one specific fragment ion, then it would be consistently reflected with the subsequent fragment ion series exhibiting the same mass shift in *m/z* values (c1 mass shift is continued for c2, c3, and so on in Fig. 8b), thereby identifying the modification location.

5.3 Sequencing tRNA Pool vs Single tRNA Species

RNA modification mapping has broad applicability (all RNA types and modification types), high accuracy (mass measurement with high precision), and well suited for identification of modifying enzymes (qualitative analysis) while screening gene knockout mutants (Addepalli and Limbach 2016; Gaston and Limbach 2014; Wetzel and Limbach 2016; Wong et al. 2013). The limitations include serial analysis of one RNA at a time, poor sensitivity, and requirement of technical expertise. The modified sequence determination methods for single RNA can be extended further to sequencing multiple tRNA species in a tRNA pool through approaches built on successes in proteomics. One such approach is comparison of the digestion products from multiple RNases such as T1, V1, and U2. Such an approach resulted in systematic identification of tRNAome from *Lactobacillus lactis* through LC-MS/

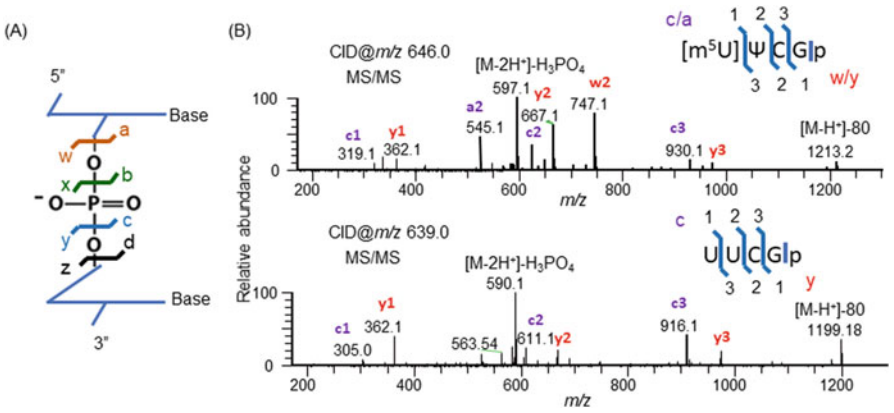


Fig. 8 Identification of modified oligonucleotide. (**a**) Potential cleavage sites on the phosphodiester bond of RNA oligonucleotide through collision-induced dissociation. (**b**) Mass shift of fragment ions denoting the position of modification. The tandem mass spectra of modified ([m^5U]ψCGp—*m/z* 646.07 (z = −2)) and unmodified (UUCGp—*m/z* 639.07) oligonucleotide precursor ions are shown on top and bottom panels, respectively. Both oligomers differ by 14 Da equivalent to methyl group (-CH3 replacing –H) and the c-fragment ion series (c1, c2, and c3) also differ by 14 units. Note the mass shift of c fragment ion series, but not the y-fragment ion series indicating that the methylation is at the 5′ end

MS analysis (Puri et al. 2014). Utilization of complementary nucleobase-specific ribonucleases improves the coverage further through generation of overlapping digestion products (Thakur et al. 2020). Combining the oligonucleotide sequences derived from multiple endonucleases can map the full sequence of modified RNA. This in turn would allow for the identification of target RNA through database searches (Matthiesen and Kirpekar 2009), de novo sequencing (Nakayama et al. 2009), or even deciphering the RNA sequence from ribonucleoprotein complexes (Taoka et al. 2009). Another approach to detect the presence of tRNA in a given sample is to detect the presence of signature digestion products that are unique to specific tRNA in a tRNA pool (Hossain and Limbach 2007; Hossain and Limbach 2009, Thakur et al. 2020).

5.4 Quantification Through Stable Isotope Label Incorporation in RNase Digests

Quantification interferences that affect the data quality such as matrix (sample background) effects can be addressed by new developments in RNA modification mapping. One way to normalize the MS response is standard addition (Rose et al. 2015). In this process, the target compound of interest gets added to a sample in increment amounts, and the concentration is estimated by extrapolating the signal onto the linear regression line. It requires synthetic oligonucleotides or the prior

knowledge of nucleotide sequence in the oligomer. If standards are not available as is the case of tRNAs where hypermodifications exist, stable isotope incorporation into the oligonucleotides to conduct comparative analysis of ribonucleic acid digests (CARD) (Li and Limbach 2012, 2013) is an attractive option. Here the oligonucleotide digestion products are labeled with stable isotopes, ^{18}O and ^{16}O, at 3′-terminal phosphate during enzymatic digestion, so that the uncharacterized RNA (C-candidate) can be compared against reference RNA (R-reference) whose modified sequence information is known. Here, the reference RNA is labeled with ^{16}O, and the candidate RNA is labeled with ^{18}O. The two digests are combined and subjected to LC-MS/MS analysis. The digestion products that share the same sequence within reference and candidate will appear as doublets separated by 2 Da, but any sequence or modification differences would generate singlets for further characterization (Fig. 9). This approach can also be extended to differentiate the isomeric sequences (e.g., CCCAAUAGp vs CAAA[ψ]CCGp) through y-fragment ion series. The isomeric sequences exhibit identical *m/z* values but differ in their sequence (Li and Limbach 2014). This approach is highly suitable for

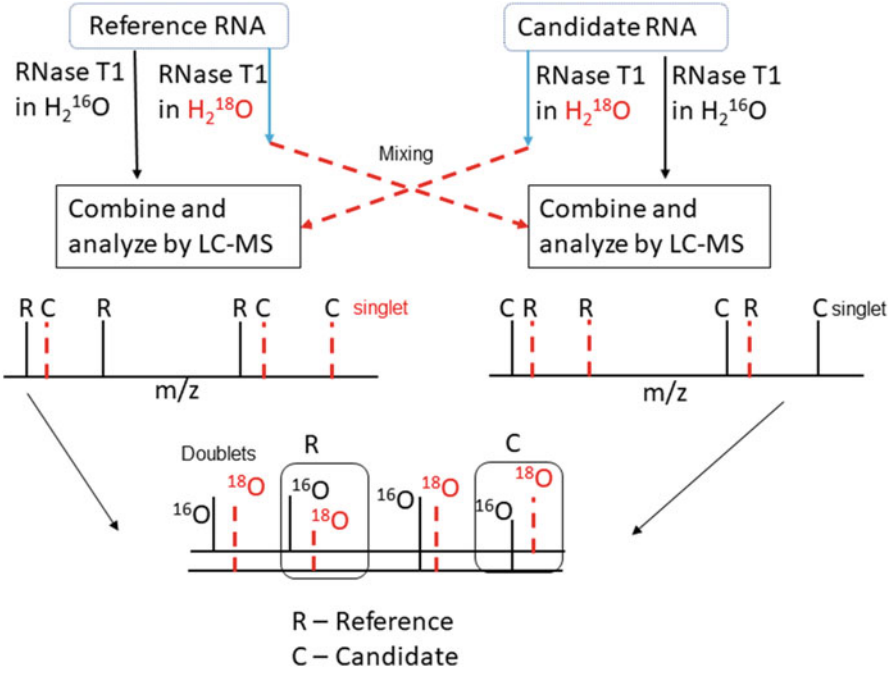


Fig. 9 LC-MS-based comparative sequencing of tRNA population through isotope labeling. Here the known tRNA sequences serves as reference and the query tRNA sample as candidate date are digested with stable isotope-labeled water so that they get incorporated in the phosphate termini. The digestion products that are equivalent will appear as doublets separated by 2 Da, and the digestion products that are different will appear as singlets adapted from Li and Limbach 2012

comparison of single or total tRNA, gene deletion mutants, and modification placement in unknown organisms.

Apart from the identification, stable isotope labeling can be used for relative quantification of RNAs (Meng and Limbach 2005). Detection of unique oligonucleotide sequences arising from single enzymatic digestion of tRNA can be another approach for global identification and potential quantification of tRNA isoacceptors (Hossain and Limbach 2007; Hossain and Limbach 2009; Wetzel and Limbach 2013). If those unique digestion products, also referred to as signature digestion products (SDPs), contain PTMs, monitoring both modified and unmodified sequences can validate the degree of modification or epitranscriptomics changes. Incorporating novel nucleobase-specific enzymes such as RNases MC1 and cusativin (Addepalli et al. 2015; Addepalli et al. 2017) can expand the number of SDPs (Thakur et al. 2020) and increase the number of potential opportunities for quantification of site-specific modifications. A combination of stable isotope labeling and SDPs can facilitate quantitative analysis of individual tRNAs in the mixture. The analytical figures of merit for this type of LC-MS method, establishment of quantitative signature digestion products (qSDP) were documented for *E. coli* tRNAs as a proof-of-concept application (Castleberry and Limbach 2010). Thus, the quantification of modified digestion products can be scalable to multiple tRNAs for relative quantification. However, this method requires conversion of 3′-cyclic phosphates to linear phosphate, which is generally a slower step in enzymatic hydrolysis by endonuclease. Treatment with protein phosphatase can overcome this limitation (Houser et al. 2015). The qSDP approach is limited by the coelution of multiple digestion products and challenging data analysis owing to spectral congestion by naturally occurring isotopes of carbon (^{13}C) and nitrogen (^{15}N). Increased charge state of oligonucleotides also makes it difficult to distinguish doublets separated by only 2 Da difference of ^{16}O vs ^{18}O isotopes. This becomes prominent while dealing with longer oligonucleotides (8–20 nt). Moreover, back exchange between two labeled oxygen atoms of digestion products can complicate the analysis even further.

5.5 *Stable Isotope Labeled In Vitro Transcript Addition*

Another approach for quantification is addition of ^{13}C-labeled in vitro transcript (as standard) to the sample for subsequent digestion with RNase and LC-MS analysis. The modified oligonucleotides of sample RNA can be quantified by comparing the intensities of doublet and singlet peaks and their abundance ratios (Taoka et al. 2015). Here, a ^{13}C and/or ^{15}N-labeled in vitro transcript can be used as reference to identify the modification containing oligonucleotide singlet peaks. Doublets are recognized by the increased mass shifts of identical sequence due to the ^{13}C and/or ^{15}N incorporation. Use of multiple stable isotopes in the reference RNA increases the mass shift of the reference compared to the sample RNA, thus simplifying the comparative analysis (Paulines and Limbach 2017) for position-

specific quantification. Adapting ^{13}C and/or ^{15}N-labeled labeled internal standard to the CARD strategy (SIL-CARD) is highly suitable for tRNA analysis, as only genomic information of RNA of interest is required to be known to generate isotopically labeled transcripts.

5.6 Quantification Through Metabolic Labeling

Metabolic labeling of RNA with two different isotopes is another approach for assuring reliable assignment of nucleolytic fragments including nucleosides (Sun et al. 2018) or oligonucleotides (Popova and Williamson 2014). Here, the reference or standard is produced by cells through metabolic labeling. The Williamson laboratory used ^{15}N media in *E. coli* cell cultures to generate ^{15}N-labeled 23S or 16S rRNA from their subunits. Mixture of ^{14}N- and ^{15}N-labeled RNA (1:1 ratio) was digested with ribonucleases, and the pairs of observed oligonucleotides were compared against the theoretically expected masses. The extracted LC-MS peak profiles and isotope distributions are fitted using least-square Fourier transform (LS-FTC) convolution (Popova and Williamson 2014; Sperling et al. 2008). Fourier transform convolution can calculate labeled isotope distributions and least squares fit for quantitative comparison with experimental peaks. This would allow determination of fractional atomic and residue label as integrated intensity of all isotopomers in the isotope distributions from experimental peaks (Sperling et al. 2008). The fractional (f) RNA values were calculated from the amplitudes of the ^{14}N and ^{15}N peaks (A_{14} and A_{15}) resulting from the least-square fit, where $f = A_{14}/(A_{14} + A_{15})$ (Fig. 10). The values were found to be tightly clustered around 0.5 ($<f> = 0.52 \pm 0.02$) indicating accurate and reproducible quantification. Altered levels of modification in the sequence would be reflected on the f value. This approach can also be adapted to a variety of other isotope labeling strategies (Sperling et al. 2008). This kind of isotope dilution LC-MS was also used for quantification of the RNA cap epitranscriptome in cellular and viral RNA (Wang et al. 2019). Although this is most rigorous approach for sensitivity, specificity, and quantitative accuracy, about 20% of all modifications were not reported by this analysis for rRNA indicating the need for identification and development of sequence-specific nucleases and chromatographic procedures that improve the separation of digestion products (Popova and Williamson 2014). Although similar applications were not demonstrated yet for tRNA, it could be a viable approach for quantification of subtle changes in modification profiles or epitranscriptomics changes in tRNAs.

5.7 Quantification of Mass Silent Modification Pseudouridine

Incorporation of additional metabolic labels such as 5,6-D-uracil enables detection of pseudouridine (isomer of uridine) modification because the 5-deuteron is

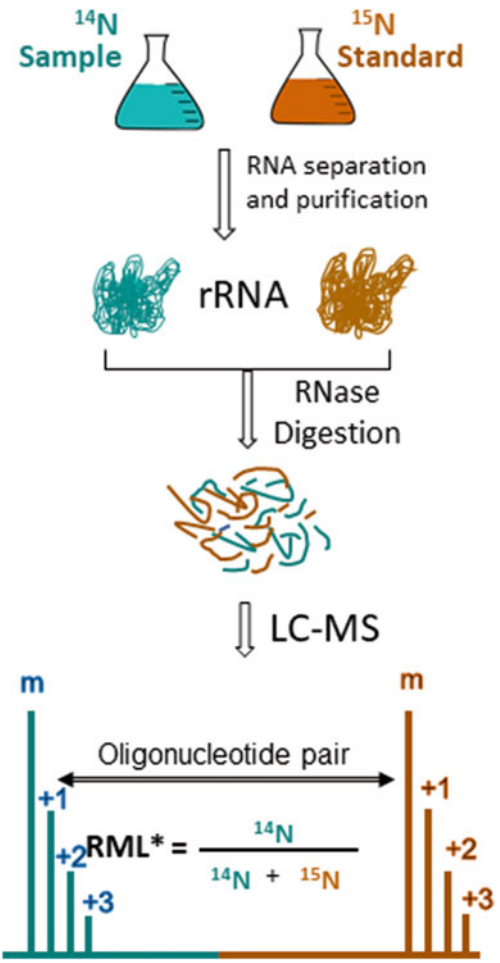


Fig. 10 Quantification of modified oligonucleotide levels through metabolic labeling and mass spectrometry. The quantitative mass spectrometry (qMS) workflow starts with the purification of RNA sample from cells grown with light and heavy isotope such as ^{14}N- and ^{15}N-labeled culture media. This isolated RNA is mixed in equal amounts and digested with RNase and subjected to LC-MS analysis. The mass spectra shows the pairs of light and heavy isotope-labeled ion peaks and their natural isotopic distributions such as monoisotopic mass (m), one (+1) or two (+2) or three (+3) heavy isotope atoms of oligonucleotide molecular ions derived from each RNA source. Following detection of specific masses of ^{14}N and ^{15}N peaks, the experimental peaks are fitted to the theoretical isotope distributions to compute the amplitudes. The peak's amplitudes are used to calculate the relative amount of the modifications (relative modification level or RML) present in the ^{14}N sample (adapted from Popova and Williamson 2014)

exchanged with the solvent providing a convenient −1 Da mass shift for pseudouridine (instead of expected +2 Da shift for uridine) detection and quantification (Popova and Williamson 2014). Such a combination of stable isotope labeling

could also distinguish compositional isomers such as AU[m^2G] and [m^2A]ψG and isobaric fragments such as [m^2G][m^5C] from G[m^4Cm] in 16S rRNA. The use of 5,6-D-uracil was combined with CRISPR-Cas9 technique to turn off uridine synthesis, so that complete labeling of uracil with uridine-5,6-D_2 was achieved to enable direct detection of pseudouridines in human RNA (Yamaki et al. 2020).

Pseudouridines could also be detected through chemical derivatization with *N*-cyclohexyl-*N*′-β-(4-methylmorpholinium) ethyl carbodiimide *p*-tosylate (CMCT) or acrylonitrile leading to mass shift of 251 Da or 53 Da for each pseudouridine. The derivatized oligomers could subsequently be detected by gel electrophoresis or LC-MS/MS analysis (Addepalli and Limbach 2016; Mengel-Jorgensen and Kirpekar 2002; Ofengand et al. 2001). The absolute amounts of pseudouridine-containing oligonucleotides could also be quantitated through selected reaction monitoring (SRM) assays where the double mass filters (operated at selection of precursor and product ions following dissociation of oligomer) improve the selectivity and increase the signal to noise ratio (Pomerantz and McCloskey 2005). The quantification is achieved through preparation of calibration curve by use of standard mixtures of ψ-containing and ψ-lacking oligomers (Addepalli and Limbach 2011). This approach does not require chemical derivatization or isotope labeling procedures.

5.8 Other Strategies

Other approaches that improve the signal for modified oligonucleotides could potentially be used for quantification, for example, mass exclusion list, where defined *m/z* values corresponding to unmodified oligonucleotides are excluded from sequence informative fragmentation pathways. Such a methodology improved the detection (and potential quantification) of modified oligonucleotides by 20% in data-dependent acquisition-based tandem mass spectrometry methods (Cao and Limbach 2015). The mass exclusion list can be altered to include both modified and unmodified versions, so that their sequence can be confirmed by MS/MS methods.

Alternate MS and MS/MS methods: Development of "top-down" MS for intact tRNAs is another attractive option for characterization of PTMs (Breuker et al. 2008; Huang et al. 2010). These approaches require uniquely high ($>10^5$) resolving power and accuracy (<1 ppm error) and is facilitated by Fourier transform mass spectrometer. However, this is currently limited to the highly purified single species of RNA molecules at this time. Additional MS/MS methods that allow alternate fragmentation methods such as electron detachment dissociation (EDD MS/MS) that yield noncomplementary *d* and *w* fragments, compared to *c* and *y* fragments in CID MS/MS, can improve the RNA sequencing data quality. Analysis of two spectra instead of single one could improve de novo sequencing of unknown modified RNA (Micura et al. 2013) for potential quantification. These fragmentation pathways were combined in radical transfer dissociation (RTD) where use of cobalt(III)hexamine

($[Co^{III}(NH_3)_6]^{3+}$) served as a reagent for production of RNA radical ions that dissociate into both *c, d, y*, and *w* fragments. Any mass spectrometer equipped with an ESI source and a CID collision cell can accomplish this task. By this technique, up to 39 nt RNA with highly labile modifications such as 5-formylcytidine and 5-hydroxymethylcytidine was successfully sequenced without the need for specialized instrumentation (Calderisi et al. 2020).

6 Data Analysis

The discussion on data analysis is limited to the LC-MS methods in this chapter. In general, sequencing of oligonucleotides depends on the generation of sequence informative product ions following CID. These product ion series share either 5′ or 3′ termini and the nomenclature, *a-B*, *c-*, and *w-*, *y-*, respectively (McLuckey et al. 1992). The mass differences in a particular ion series type (*c-*, *y-*, *w-*, *a-B*) allows determination of exact sequence including the location of modification within the oligonucleotide. However, this data interpretation is not high throughput as it requires assignment of individual precursor *m/z* values to each oligonucleotide and assignment of many product ions in the MS/MS spectra. Afterward, the identified sequence along with modification is mapped back to the overall sequence of transcript. However, the complexity and sheer number of LC-MS/MS spectra pose a significant challenge for such exercise.

For increased throughput and automated data interpretation, development of new computational methods is essential. Toward this end, a number of software solutions have been offered starting with simple oligonucleotide sequencer (SOS) (Rozenski and McCloskey 2002), Ariadne (Nakayama et al. 2009), OMA/OPA (Nyakas et al. 2013), Robooligo (Sample et al. 2015), RNAModMapper (Lobue et al. 2019a, b; Yu et al. 2017), and recently NucleicAcidSearchEngine (NASE) (Wein et al. 2020). SOS allowed ab initio sequencing in an interactive software environment from MS/MS spectra. Ariadne was the first computational platform catered toward determining the location of select modifications for a limited number of organisms using a web-based database search engine. The OMA and OPA analyze the spectra of oligonucleotides, their derivatives, and adducts with metal ions. RNA mass mapping (RMM) is another program that uses both RNA sequence databases and genome searches for mapping based on peptide mass fingerprinting (Matthiesen and Kirpekar 2009). Robooligo provided a handle for both manual and automated de novo analysis of modified tRNA but limited to single sequences (Sample et al. 2015).

RNAModMapper is a stand-alone program that can interpret large number of LC-MS/MS spectra to map oligonucleotides back onto the sequences. The user can choose the fixed or variable position mapping modes depending on the amount of known modification profiles of target sequence. To ensure accuracy of spectral interpretation, two-component scoring function with user-defined scoring thresholds were implemented (Yu et al. 2017). The approaches were further optimized for

analysis of each of two LC-MS platforms (Orbitrap-CID or HCD vs Synapt G2S time-of-flight beam-type CID) (Lobue et al. 2019a, b). Recently developed NucleicAcidSearchEngine (NASE) is a scalable database-matching tool implemented within the OpenMS framework for processing tandem MS/MS data. NASE also has advance features like false discovery rate (FDR) estimation, precursor mass correction, and support for salt adducts. The NASE could match sequences of 5–21 nt length, support automated data analysis workflows, and potentially label-free quantification of RNA oligonucleotides (Wein et al. 2020) with high-resolution MS data. This analysis also considers the instrument selection of higher isotopologue peaks for CID and provides relatively faster performance (23-modification searches of tRNA dataset takes ~30 min per file using 40 parallel threads). Future developments could focus on comparative quantification across multiple samples through streamlining of the OpenMS tools and algorithms.

7 Future Technologies for Quantification of Modifications in tRNA

Nanopore sequencing is a promising orthogonal technique which works directly with a single RNA molecule. It differs from NGS and qPCR as it does not require the synthesis of cDNA intermediate by reverse transcriptase, thereby removing amplification bias (Garalde et al. 2018; Workman et al. 2019). This kind of direct RNA-seq approach yields information for a full-length transcript instead of short reads and enables detection of nucleotide analogs or modifications in RNA. It uses electrophoresis to move the nucleic acid sequence through the nanoscale pore to measure the change in applied potential. When modified RNA passes through the pore, it causes temporary blockage of current enabling detection of 5-methylcytosine (m^5C), *N*6-methyl adenosine (m^6A), inosine, pseudouridine, and 7-methylguanosine (m^7G) in the sequence (Garalde et al. 2018; Liu et al. 2019; Smith et al. 2019; Viehweger et al. 2019, Workman et al. 2019).

The nanopore technology is still under development and further improvements such as pore chemistry changes and streamlined algorithm to interpret the changes in electrical potential readouts are required so that the accuracy can be improved (Kono and Arakawa 2019). In this direction, a software tool (ELIGOS) and bioinformatic software package was developed, which can predict the presence of modified bases from the background error data of in vitro transcript. This tool provided accurate information of epitranscriptional landscape (~90% for m^6A) in model organisms and human cells (Jenjaroenpun et al. 2021). However, errors are still being encountered with certain base modifications, homopolymers, strand damage, or even certain structural features. Nevertheless, ELIGOS could serve as a diagnostic tool for improvement of base-calling.

8 Conclusions

With direct RNA-seq still in its infancy, the LC-MS/MS approaches are highly relevant for characterization of modified sequences for hypomodifications in various RNA types. LC-MS/MS could also serve as an independent and orthogonal technique to validate multiple modifications found in RNA-seq or nanopore sequencing approaches. In combination with relative and absolute quantification methods, this platform should provide middle ground between RNA-seq- and qRT-PCR-based biochemical approaches.

References

Addepalli B, Lesner NP, Limbach PA (2015) Detection of RNA nucleoside modifications with the uridine-specific ribonuclease MC1 from Momordica charantia. RNA 21:1746–1756

Addepalli B, Limbach PA (2011) Mass spectrometry-based quantification of pseudouridine in RNA. J Am Soc Mass Spectrom 22:1363–1372

Addepalli B, Limbach PA (2016) Pseudouridine in the anticodon of Escherichia coli tRNATyr (QPsiA) is catalyzed by the dual specificity enzyme RluF. J Biol Chem 291:22327–22337

Addepalli B, Venus S, Thakur P et al (2017) Novel ribonuclease activity of cusativin from Cucumis sativus for mapping nucleoside modifications in RNA. Anal Bioanal Chem 409:5645–5654

Agris PF, Narendran A, Sarachan K et al (2017) The importance of being modified: the role of RNA modifications in translational Fidelity. Enzymes, vol 41. Elsevier, Oxford, UK, pp 1–50

Alpert AJ (1990) Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. J Chromatogr A 499:177–196

Apffel A, Chakel JA, Fischer S et al (1997) Analysis of oligonucleotides by HPLC-electrospray ionization mass spectrometry. Anal Chem 69:1320–1325

Apgar J, Holley RW, Merrill SH (1962) Purification of the alanine-, valine-, histidine-, and tyrosine-acceptor ribonucleic acids from yeast. J Biol Chem 237:796–802

Basanta-Sanchez M, Temple S, Ansari SA et al (2016) Attomole quantification and global profile of RNA modifications: Epitranscriptome of human neural stem cells. Nucleic Acids Res 44:e26

Berg OG, Kurland CG (1997) Growth rate-optimised tRNA abundance and codon usage. J Mol Biol 270:544–550

Birkedal U, Christensen-Dalsgaard M, Krogh N et al (2015) Profiling of ribose methylations in RNA by high-throughput sequencing. Angew Chem Int Ed Engl 54:451–455

Björk G, Hagervall T (2014) Transfer RNA modification: presence, synthesis, and function. Eco Sal Plus

Blanco S, Dietmann S, Flores JV et al (2014) Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. EMBO J 33:2020–2039

Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307

Borland K, Diesend J, Ito-Kureha T et al (2019) Production and application of stable isotope-labeled internal standards for RNA modification analysis. Genes 10:26

Brandmayr C, Wagner M, Brückl T et al (2012) Isotope-based analysis of modified tRNA nucleosides correlates modification density with translational efficiency. Angew Chem Int Ed Engl 51:11162–11165

Brenton AG, Godfrey AR (2010) Accurate mass measurement: terminology and treatment of data. J Am Soc Mass Spectrom 21:1821–1835

Breuker K, Jin M, Han X et al (2008) Top-down identification and characterization of biomolecules by mass spectrometry. J Am Soc Mass Spectrom 19:1045–1053
Bruch A, Laguna T, Butter F et al (2020) Misactivation of multiple starvation responses in yeast by loss of tRNA modifications. Nucleic Acids Res 48:7307–7320
Brückl T, Globisch D, Wagner M et al (2009) Parallel isotope-based quantification of modified tRNA nucleosides. Angew Chem Int Ed Engl 48:7932–7934
Buck M, Ames BN (1984) A modified nucleotide in tRNA as a possible regulator of aerobiosis: synthesis of cis-2-methyl-thioribosylzeatin in the tRNA of Salmonella. Cell 36:523–531
Calderisi G, Glasner H, Breuker K (2020) Radical transfer dissociation for De novo characterization of modified ribonucleic acids by mass spectrometry. Angew Chem Int Ed Engl 59:4309–4313
Cao X, Limbach PA (2015) Enhanced detection of post-transcriptional modifications using a mass-exclusion list strategy for RNA modification mapping by LC-MS/MS. Anal Chem 87:8433–8440
Castleberry CM, Limbach PA (2010) Relative quantitation of transfer RNAs using liquid chromatography mass spectrometry and signature digestion products. Nucleic Acids Res 38:e162–e162
Cattenoz PB, Taft RJ, Westhof E et al (2013) Transcriptome-wide identification of a > I RNA editing sites by inosine specific cleavage. RNA 19:257–270
Chan CTY, Dyavaiah M, DeMott MS et al (2010) A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. PLoS Genet 6:1–9
Chan PP, Lowe TM (2016) GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. Nucleic Acids Res 44:D184–D189
Chan CT, Pang YL, Deng W et al (2012) Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. Nat Commun 3:937
Chan C, Pham P, Dedon PC et al (2018) Lifestyle modifications: coordinating the tRNA epitranscriptome with codon bias to adapt translation during stress responses. Genome Biol 19:228
Chen B, Mason SF, Bartlett MG (2013a) The effect of organic modifiers on electrospray ionization charge-state distribution and desorption efficiency for oligonucleotides. J Am Soc Mass Spectrom 24:257–264
Chen ML, Shen F, Huang W et al (2013b) Quantification of 5-methylcytosine and 5-hydroxymethylcytosine in genomic DNA from hepatocellular carcinoma tissues by capillary hydrophilic-interaction liquid chromatography/quadrupole TOF mass spectrometry. Clin Chem 59:824–832
Chen K, Song B, Tang Y et al (2021) RMDisease: a database of genetic variants that affect RNA modifications, with implications for epitranscriptome pathogenesis. Nucleic Acids Res 49(D1): D1396–D1404
Chionh YH, Ho CH, Pruksakorn D et al (2013) A multidimensional platform for the purification of non-coding RNA species. Nucleic Acids Res 41:e168
Chionh YH, McBee M, Babu IR et al (2016) tRNA-mediated codon-biased translation in mycobacterial hypoxic persistence. Nat Commun 7:13302
Chomczynski P, Sacchi N (2006) The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. Nat Protoc 1:581–585
Clark WC, Evans ME, Dominissini D et al (2016) tRNA base methylation identification and quantification via high-throughput sequencing. RNA 22:1771–1784
Contreras-Sanz A, Scott-Ward TS, Gill HS et al (2012) Simultaneous quantification of 12 different nucleotides and nucleosides released from renal epithelium and in human urine samples using ion-pair reversed-phase HPLC. Purinergic Signal 8:741–751
Cozen AE, Quartley E, Holmes AD et al (2015) ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. Nature Meth 12:879
Dal Magro C, Keller P, Kotter A et al (2018) A vastly increased chemical variety of RNA modifications containing a Thioacetal structure. Angew Chem Int Ed Engl 57:7893–7897

Dalluge JJ, Hashizume T, McCloskey JA (1996) Quantitative measurement of dihydrouridine in RNA using isotope dilution liquid chromatography-mass spectrometry (LC/MS). Nucleic Acids Re 24:3242–3245
Delatte B, Wang F, Ngoc LV et al (2016) RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. Science 351:282–285
Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S et al (2016) The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. Nature 530:441–446
Douthwaite S, Kirpekar F (2007) Identifying modifications in RNA by MALDI mass spectrometry. Methods Enzymol 425:3–20
Emilsson V, Näslund AK, Kurland CG (1992) Thiolation of transfer RNA in Escherichia coli varies with growth rate. Nucleic Acids Res 20:4499–4505
Enroth C, Poulsen LD, Iversen S et al (2019) Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing. Nucleic Acids Res 47:e126
Evans ME, Clark WC, Zheng G et al (2017) Determination of tRNA aminoacylation levels by high-throughput sequencing. Nucleic Acids Res 45:e133–e133
Fenn JB, Mann M, Meng CK et al (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246:64–71
Fischer W, Derwenskus KH, Sprinzl M (1982) On the properties of immobilized elongation factor Tu from Thermus thermophilus HB8. Eur J Biochem 125:43–149
Gabelica V, Marklund E (2018) Fundamentals of ion mobility spectrometry. Curr Opin Chem Biol 42:1–59
Garalde DR, Snell EA, Jachimowicz D et al (2018) Highly parallel direct RNA sequencing on an array of nanopores. Nat Methods 15:201–206
Gaston KW, Limbach PA (2014) The identification and characterization of non-coding and coding RNAs and their modified nucleosides by mass spectrometry. RNA Biol 11:1568–1585
Geslain R, Pan T (2010) Functional analysis of human tRNA isodecoders. J Mol Biol 396:821–831
Gilar M, Fountain KJ, Budman Y et al (2002) Ion-pair reversed-phase high-performance liquid chromatography analysis of oligonucleotides: retention prediction. J Chromatogr A 958:167–182
Gilar M, Neue UD (2007) Peak capacity in gradient reversed-phase liquid chromatography of biopolymers: theoretical and practical implications for the separation of oligonucleotides. J Chromatogr A 1169:139–150
Greig M, Griffey RH (1995) Utility of organic bases for improved electrospray mass spectrometry of oligonucleotides. Rapid Commun Mass Spectrom 9:97–102
Grosjean H, Droogmans L, Roovers M et al (2007) Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. Methods Enzymol 425(55):57–101
Haag AM (2016) Mass analyzers and mass spectrometers. Adv Exp Med Biol 919:157–169
Hanson G, Coller J (2018) Codon optimality, bias and usage in translation and mRNA decay. Nat Rev Mol Cell Biol 19:20–30
Haruehanroengra P, Zheng YY, Zhou Y et al (2020) RNA modifications and cancer. RNA Biol 17:1560–1575
Heiss M, Reichle VF, Kellner S (2017) Observing the fate of tRNA and its modifications by nucleic acid isotope labeling mass spectrometry: NAIL-MS. RNA Biol 14:1260–1268
Hossain M, Limbach PA (2007) Mass spectrometry-based detection of transfer RNAs by their signature endonuclease digestion products. RNA 13:295–303
Hossain M, Limbach PA (2009) Multiple endonucleases improve MALDI-MS signature digestion product detection of bacterial transfer RNAs. Anal Bioanal Chem 394:125–1135
Houser WM, Butterer A, Addepalli B et al (2015) Combining recombinant ribonuclease U2 and protein phosphatase for RNA modification mapping by liquid chromatography-mass spectrometry. Anal Biochem 478:52–58
Huang TY, Liu J, McLuckey SA (2010) Top-down tandem mass spectrometry of tRNA via ion trap collision-induced dissociation. J Am Soc Mass Spectrom 21:890–898

Isak G, Rydén-Aulin M (2009) Hypomodification of the wobble base in tRNAGlu, tRNALys, and tRNAGln suppresses the temperature-sensitive phenotype caused by mutant release factor 1. J Bacteriol 191:1604

Jenjaroenpun P, Wongsurawat T, Wadley TD et al (2021) Decoding the epitranscriptional landscape from native RNA sequences. Nucleic Acids Res 49:e7

Jensen SS, Ariza X, Nielsen P et al (2007) Collision-induced dissociation of cytidine and its derivatives. J Mass Spectrom 42:49–57

Jora M, Borland K, Abernathy S et al (2020) Chemical amination/imination of carbonothiolated nucleosides during RNA hydrolysis. Angew Chem Int Ed Engl

Jora M, Burns AP, Ross RL et al (2018) Differentiating positional isomers of nucleoside modifications by higher-energy collisional dissociation mass spectrometry (HCD MS). J Am Soc Mass Spectrom 29:1745–1756

Jora M, Lobue PA, Ross RL et al (2019) Detection of ribonucleoside modifications by liquid chromatography coupled with mass spectrometry. Biochim Biophys Acta Gene Regul Mech 1862:280–290

Kang BI, Miyauchi K, Matuszewski M et al (2017) Identification of 2-methylthio cyclic N6-threonylcarbamoyladenosine (ms2ct6A) as a novel RNA modification at position 37 of tRNAs. Nucleic Acids Res 45:2124–2136

Kazuhito T, Wei FY (2020) Posttranscriptional modifications in mitochondrial tRNA and its implication in mitochondrial translation and disease. J Biochem 168:435–444

Kebarle P, Tang L (1993) Fom ions in solution to ions in the gas phase. Anal Chem 65:972A–986A

Kellner S, Neumann J, Rosenkranz D et al (2014a) Profiling of RNA modifications by multiplexed stable isotope labelling. Chem Commun (Camb) 50:3516–3518

Kellner S, Ochel A, Thüring K et al (2014b) Absolute and relative quantification of RNA modifications via biosynthetic isotopomers. Nucleic Acids Res 42:e142

Kenderdine T, Nemati R, Baker A et al (2020) High-resolution ion mobility spectrometry-mass spectrometry of isomeric/isobaric ribonucleotide variants. J Mass Spectrom 55:e4465

Khoddami V, Yerra A, Mosbruger TL et al (2019) Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. Proc Natl Acad Sci U S A 116:6784–6789

Kimura S, Srisuknimit V, Waldor MK (2020) Probing the diversity and regulation of tRNA modifications. Curr Opin Microbiol 57:41–48

Klyde BJ, Bernfield MR (1973) Purification of chicken liver seryl transfer ribonucleic acid by complex formation with elongation factor EF-Tu:GTP. A general micromethod of aminoacyl transfer ribonucleic acid purification. Biochemistry 12:3752–3757

Kolakowski BM, Mester Z (2007) Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS). Analyst 132:842–864

Kono N, Arakawa K (2019) Nanopore sequencing: review of potential applications in functional genomics. Develop Growth Differ 61:316–326

Kowalak JA, Pomerantz SC, Crain PF et al (1993) A novel method for the determination of post-transcriptional modification in RNA by mass spectrometry. Nucleic Acids Res 21:4577–4585

Kristen M, Plehn J, Marchand V et al (2020) Manganese ions individually Alter the reverse transcription signature of modified Ribonucleosides. Genes (Basel) 11:950

Lant JT, Berg MD, Heinemann IU et al (2019) Pathways to disease from natural variations in human cytoplasmic tRNAs. J Biol Chem 294:5294–5308

Laxman S, Sutter BM, Wu X et al (2013) Sulfur amino acids regulate translational capacity and metabolic homeostasis through modulation of tRNA thiolation. Cell 154:416–416

Lee C, Kramer G, Graham DE et al (2007) Yeast mitochondrial initiator tRNA is methylated at guanosine 37 by the Trm5-encoded tRNA (guanine-N1-)-methyltransferase. J Biol Chem 282:27744–27753

Lentini JM, Ramos J, Fu D (2018) Monitoring the 5-methoxycarbonylmethyl-2-thiouridine (mcm5s2U) modification in eukaryotic tRNAs via the γ-toxin endonuclease. RNA 24:749–758

Li S, Limbach PA (2012) Method for comparative analysis of ribonucleic acids using isotope labeling and mass spectrometry. Anal Chem 84:8607–8613
Li S, Limbach PA (2013) Mass spectrometry sequencing of transfer ribonucleic acids by the comparative analysis of RNA digests (CARD) approach. Analyst 138:1386–1394
Li S, Limbach PA (2014) Identification of RNA sequence isomer by isotope labeling and LC-MS/MS. J Mass Spectrom 49:1191–1198
Limbach PA, Paulines MJ (2017) Going global: the new era of mapping modifications in RNA. Wiley Interdiscip Rev RNA 8
Liu H, Begik O, Lucas MC et al (2019) Accurate detection of m6A RNA modifications in native RNA sequences. Nat Commun 10:4079
Lobue PA, Jora M, Addepalli B et al (2019a) Oligonucleotide analysis by hydrophilic interaction liquid chromatography-mass spectrometry in the absence of ion-pair reagents. J Chromatogr A 1595:39–48
Lobue PA, Yu N, Jora M et al (2019b) Improved application of RNAModMapper–an RNA modification mapping software tool–for analysis of liquid chromatography tandem mass spectrometry (LC-MS/MS) data. Methods 156:128–138
Marchand V, Ayadi L, Ernst FGM et al (2018) AlkAniline-Seq: profiling of m7G and m3C RNA modifications at single nucleotide resolution. Angew Chem Int Ed Engl 57:16785–16790
Mathlin J, Le Pera L, Colombo T (2020) A census and categorization method of Epitranscriptomic Marks. Int J Mol Sci 21:4684
Matthiesen R, Kirpekar F (2009) Identification of RNA molecules by specific enzyme digestion and mass spectrometry: software for and implementation of RNA mass mapping. Nucleic Acids Res 37:e48
McCown PJ, Ruszkowska A, Kunkler CN et al (2020) Naturally occurring modified ribonucleosides. Wiley Interdisciplinary Reviews: RNA 11:e1595
McLuckey SA, Van Berker GJ, Glish GL (1992) Tandem mass spectrometry of small, multiply charged oligonucleotides. J Am Soc Mass Spectrom 3:60–70
Meng Z, Limbach PA (2005) Quantitation of ribonucleic acids using 18O labeling and mass spectrometry. Anal Chem 77:1891–1895
Meng-Dan Lan B-FY, Feng Y-Q (2019) Deciphering nucleic acid modifications by chemical derivatization-mass spectrometry analysis. Chinese Chem Lett 30:1–6
Mengel-Jorgensen J, Kirpekar F (2002) Detection of pseudouridine and other modifications in tRNA by cyanoethylation and MALDI mass spectrometry. Nucleic Acids Res 30:e135/131–e135/137
Meyer B, Immer C, Kaiser S et al (2020) Identification of the 3-amino-3-carboxypropyl (acp) transferase enzyme responsible for acp3U formation at position 47 in Escherichia coli tRNAs. Nucleic Acids Res 48:1435–1450
Meyer KD, Saletore Y, Zumbo P et al (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149:1635–1646
Micura R, Kreutz C, Breuker K (2013) A personal perspective on chemistry-driven RNA research. Biopolymers 99:1114–1123
Miyauchi K, Ohara T, Suzuki T (2007) Automated parallel isolation of multiple species of non-coding RNAs by the reciprocal circulating chromatography method. Nucleic Acids Res 35:e24
Motorin Y, Helm M (2010) tRNA stabilization by modified nucleotides. Biochemistry 49:4934–4944
Motorin Y, Muller S, Behm-Ansmant I et al (2007) Identification of modified residues in RNAs by reverse transcription-based methods. Methods Enzymol 425:21–53
Muench KH, Berg P (1966) Fractionation of transfer ribonucleic acid by gradient partition chromatography on Sephadex columns*. Biochemistry 5:970–981
Nakayama H, Akiyama M, Taoka M et al (2009) Ariadne: a database search engine for identification and chemical analysis of RNA using tandem mass spectrometry data. Nucleic Acids Res 37:e47

Nilsen TW (2012) Selective precipitation of large RNAs. Cold Spring Harbor Protocols 2012: pdb.prot072322
Nyakas A, Blum LC, Stucki SR et al (2013) OMA and OPA--software-supported mass spectra analysis of native and modified nucleic acids. J Am Soc Mass Spectrom 24:249–256
Ofengand J, Del Campo M, Kaya Y (2001) Mapping pseudouridines in RNA molecules. Methods 25:365–373
Pathak C, Jaiswal YK, Vinayak M (2005) Hypomodification of transfer RNA in cancer with respect to queuosine. RNA Biol 2:143–148
Paulines MJ, Limbach PA (2017) Stable isotope labeling for improved comparative analysis of RNA digests by mass spectrometry. J Am Soc Mass Spectrom 28:51–561
Pinkard O, McFarland S, Sweet T et al (2020) Quantitative tRNA-sequencing uncovers metazoan tissue-specific tRNA regulation. Nat Commun 11:4104
Pitt JJ (2009) Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. Clin Biochemist Rev 30:19–34
Pomerantz SC, McCloskey JA (1990) Analysis of RNA hydrolyzates by liquid chromatography-mass spectrometry. Methods Enzymol 193:796–824
Pomerantz SC, McCloskey JA (2005) Detection of the common RNA nucleoside pseudouridine in mixtures of oligonucleotides by mass spectrometry. Anal Chem 77:687–4697
Popova AM, Williamson JR (2014) Quantitative analysis of rRNA modifications using stable isotope labeling and mass spectrometry. J Am Chem Soc 136:2058–2069
Puri P, Wetzel C, Saffert P et al (2014) Systematic identification of tRNAome and its dynamics in Lactococcus lactis. Mol Microbiol 93:944–956
Reichle VF, Kaiser S, Heiss M et al (2019) Surpassing limits of static RNA modification analysis with dynamic NAIL-MS. Methods 156:91–101
Ribeiro S, Nock S, Sprinzl M (1995) Purification of aminoacyl-tRNA by affinity chromatography on immobilized Thermus thermophilus EF-Tu.GTP. Anal Biochem 228:330–335
Rose RE, Quinn R, Sayre JL et al (2015) Profiling ribonucleotide modifications at full-transcriptome level: a step toward MS-based epitranscriptomics. RNA 21:1361–1374
Rozenski J, McCloskey JA (2002) SOS: a simple interactive program for ab initio oligonucleotide sequencing by mass spectrometry. J Am Soc Mass Spectrom 13:200–203
Russell SP, Limbach PA (2013) Evaluating the reproducibility of quantifying modified nucleosides from ribonucleic acids by LC-UV-MS. J Chromatogr B Analyt Technol Biomed Life Sci 923-924:74–82
Sakai Y, Miyauchi K, Kimura S et al (2016) Biogenesis and growth phase-dependent alteration of 5-methoxycarbonylmethoxyuridine in tRNA anticodons. Nucleic Acids Res 44:509–523
Sample PJ, Gaston KW, Alfonzo JD et al (2015) RoboOligo: software for mass spectrometry data to support manual and de novo sequencing of post-transcriptionally modified ribonucleic acids. Nucleic Acids Res 43:e64
Santos IC, Brodbelt JS (2021) Recent developments in the characterization of nucleic acids by liquid chromatography, capillary electrophoresis, ion mobility and mass spectrometry (2010-2020). J Sep Sci 44:340–372
Shepherd J, Ibba M (2015) Bacterial transfer RNAs. FEMS Microbiol Rev 39:280–300
Smith AM, Jain M, Mulroney L et al (2019) Reading canonical and modified nucleobases in 16S ribosomal RNA using nanopore native RNA sequencing. PLoS One 14:e0216709
Sperling E, Bunner AE, Sykes MT et al (2008) Quantitative analysis of isotope distributions in proteomic mass spectrometry using least-squares Fourier transform convolution. Anal Chem 80:4906–4917
Squires JE, Patel HR, Nousch M et al (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40:5023–5033
Su D, Chan CT, Gu C et al (2014) Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. Nat Protoc 9:828–841
Sun C, Jora M, Solivio B et al (2018) The effects of ultraviolet radiation on nucleoside modifications in RNA. ACS Chem Biol 13:567–572

Suzuki T, Suzuki T (2007) Chaplet column chromatography: isolation of a large set of individual RNAs in a single step. Methods Enzymol 425:231–239
Suzuki T, Suzuki T (2014) A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. Nucleic Acids Res 42:7346–7357
Suzuki T, Yashiro Y, Kikuchi I et al (2020) Complete chemical structures of human mitochondrial tRNAs. Nat Commun 11:4269
Taoka M, Nobe Y, Hori M et al (2015) A mass spectrometry-based method for comprehensive quantitative determination of post-transcriptional RNA modifications: the complete chemical structure of Schizosaccharomyces pombe ribosomal RNAs. Nucleic Acids Res 43:e115
Taoka M, Yamauchi Y, Nobe Y et al (2009) An analytical platform for mass spectrometry-based identification and chemical analysis of RNA in ribonucleoprotein complexes. Nucleic Acids Res 37:e140
Thakur P, Estevez M, Lobue PA et al (2020) Improved RNA modification mapping of cellular non-coding RNAs using C- and U-specific RNases. Analyst 145:816–827
Thomas JM, Bryson KM, Meier JL (2019) Chapter three–nucleotide resolution sequencing of N4-acetylcytidine in RNA. Methods in enzymology, vol 621. Academic Press, Shukla, AK, pp 31–51
Thompson JA, Blakesley RW, Doran K et al (1983) Purification of nucleic acids by RPC-5 ANALOG chromatography: peristaltic and gravity-flow applications. Methods Enzymol 100:368–399
Thumbs P, Ensfelder TT, Hillmeier M et al (2020) Synthesis of galactosyl-queuosine and distribution of hypermodified Q-nucleosides in mouse tissues. Angew Chem Int Ed Engl 59:12352–12356
Thüring K, Schmid K, Keller P et al (2016) Analysis of RNA modifications by liquid chromatography-tandem mass spectrometry. Methods 107:48–56
Torrent M, Chalancon G, De Groot NS et al (2018) Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. Sci Signal 11:eaat6409
Torres AG, Reina O, Stephan-Otto Attolini C et al (2019) Differential expression of human tRNA genes drives the abundance of tRNA-derived fragments. Proc Natl Acad Sci U S A 116:8451–8456
Traube FR, Schiffers S, Iwan K et al (2019) Isotope-dilution mass spectrometry for exact quantification of noncanonical DNA nucleosides. Nat Protocols 14:283–312
Väre VYP, Eruysal ER, Narendran A et al (2017) Chemical and conformational diversity of modified nucleosides affects tRNA structure and function. Biomol Ther 7:29
Viehweger A, Krautwurst S, Lamkiewicz K et al (2019) Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. Genome Res 29:1545–1554
Wang J, Alvin Chew BL, Lai Y et al (2019) Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. Nucleic Acids Res 47:e130
Wang X, Matuszek Z, Huang Y et al (2018) Queuosine modification protects cognate tRNAs against ribonuclease cleavage. RNA 24:1305–1313
Wang X, Yan Q, Guan MX (2010) Combination of the loss of cmnm5U34 with the lack of s2U34 modifications of tRNALys, tRNAGlu, and tRNAGln altered mitochondrial biogenesis and respiration. J Mol Biol 395:1038–1048
Wein S, Andrews B, Sachsenberg T et al (2020) A computational platform for high-throughput analysis of RNA sequences and modifications by mass spectrometry. Nat Commun 11:926
Werner A (1993) Reversed-phase and ion-pair separations of nucleotides, nucleosides and nucleobases: analysis of biological samples in health and disease. J Chromatogr 618:3–14
Wetzel C, Limbach PA (2013) The global identification of tRNA isoacceptors by targeted tandem mass spectrometry. Analyst 138:6063–6072
Wetzel C, Limbach PA (2016) Mass spectrometry of modified RNAs: recent developments. Analyst 141:16–23

Wimmer E, Maxwell IH, Tener GM (1968) Simple method for isolating highly purified yeast phenylalanine transfer ribonucleic acid. Biochemistry 7:2623–2628

Wong SY, Javid B, Addepalli B et al (2013) Functional role of methylation of G518 of the 16S rRNA 530 loop by GidB in Mycobacterium tuberculosis. Antimicrob Agents Chemother 57:6311–6318

Workman RE, Tang AD, Tang PS et al (2019) Nanopore native RNA sequencing of a human poly (a) transcriptome. Nat Methods 16:1297–1305

Yamakawa Y, Miyasaka K, Ishikawa T et al (1990) High-performance liquid chromatography of transfer ribonucleic acids on spherical hydroxyapatite beads. J Chromatogr 506:319–326

Yamaki Y, Nobe Y, Koike M et al (2020) Direct determination of Pseudouridine in RNA by mass spectrometry coupled with stable isotope labeling. Anal Chem 92:11349–11356

Yan M, Wang Y, Hu Y et al (2013) A high-throughput quantitative approach reveals more small RNA modifications in mouse liver and their correlation with diabetes. Anal Chem 85:12173–12181

Yokogawa T, Kitamura Y, Nakamura D et al (2010) Optimization of the hybridization-based method for purification of thermostable tRNAs in the presence of tetraalkylammonium salts. Nucleic Acids Res 38:e89

Yu N, Jora M, Solivio B et al (2019) tRNA modification profiles and codon-decoding strategies in Methanocaldococcus jannaschii. J Bacteriol 201:e00690–e00618

Yu N, Lobue PA, Cao X et al (2017) RNAModMapper: RNA modification mapping software for analysis of liquid chromatography tandem mass spectrometry data. Anal Chem 89:10744–10752

Zallot R, Ross R, Chen WH et al (2017) Identification of a novel Epoxyqueuosine reductase family by comparative genomics. ACS Chem Biol 12:844–851

Zhang W, Eckwahl MJ, Zhou KI et al (2019) Sensitive and quantitative probing of pseudouridine modification in mRNA and long noncoding RNA. RNA 25:1218–1225

Zhao H-Q, Wang X, Li H-M et al (2013) Characterization of nucleosides and nucleobases in natural Cordyceps by HILIC-ESI/TOF/MS and HILIC-ESI/MS. Molecules (Basel, Switzerland) 18:9755–9769

Zheng G, Qin Y, Clark WC et al (2015) Efficient and quantitative high-throughput tRNA sequencing. Nat Methods 12:835–837

Zhou G, Pang H, Tang Y et al (2014) Hydrophilic interaction ultra-performance liquid chromatography coupled with triple-quadrupole tandem mass spectrometry (HILIC-UPLC-TQ-MS/MS) in multiple-reaction monitoring (MRM) for the determination of nucleobases and nucleosides in ginkgo seeds. Food Chem 150:260–266

Zhu RX, Ching WM, Chung HK et al (1987) Purification of individual tRNAs using a monoclonal anti-AMP antibody affinity column. Anal Biochem 161:460–466

Zubay G (1962) The isolation and fractionation of soluble ribonucleic acid. J Mol Biol 4:347–356

Experimental and Computational Methods for Guiding Identification and Characterization of Epitranscriptome Proteins

Matthew R. Burroughs, Juan C. Gonzalez-Rivera, Antonio Cordova, and Lydia M. Contreras

Contents

M. R. Burroughs · J. C. Gonzalez-Rivera
McKetta Department of Chemical Engineering, University of Texas at Austin, Austin, TX, USA
e-mail: mrburroughs@utexas.edu; jc.gonzalez@utexas.edu

A. Cordova
Department of Molecular Biosciences, University of Texas at Austin, Austin, TX, USA
e-mail: acordova4144@utexas.edu

L. M. Contreras (✉)
McKetta Department of Chemical Engineering, University of Texas at Austin, Austin, TX, USA

Department of Molecular Biosciences, University of Texas at Austin, Austin, TX, USA
e-mail: lcontrer@che.utexas.edu

S. Jurga, J. Barciszewski (eds.), *Epitranscriptomics*, RNA Technologies 12,
https://doi.org/10.1007/978-3-030-71612-7_22

Abstract Modifications to the four canonical RNA nucleotides decorate cellular RNAs, forming what is collectively known as the epitranscriptome. A few of the well-studied RNA modifications (e.g., m^6A) have been shown to be implicated in processes critical to RNA regulation, including RNA translation, stability, splicing, and nuclear export. Many additional RNA modifications exist which have yet to be characterized in such a manner. Proteins which interface with these RNA modifications via modification formation, binding, and removal (epitranscriptome proteins) are key elements required for a wholistic understanding of the modification's regulatory functionalities; at present, these epitranscriptome proteins are even less characterized than their cognate RNA modifications. In this chapter, we provide an analysis of the different techniques that are available for epitranscriptome protein identification and characterization. We discuss both experimental (wet-lab) and computational strategies for guiding these investigations. We also include a brief commentary on how knowledge of epitranscriptome proteins can be used to inform strategies for therapeutic intervention with respect to RNA modification-related medical ailments.

Keywords RNA modifications · Modified RNA-protein interactions · RNA modification readers · RNA modification writers · RNA modification erasers · Epitranscriptome proteins · Computational predictions

1 Introduction

From a colloquial vantage point, RNA molecules are biopolymers comprised of four different monomer subunits: adenosine (A), guanosine (G), cytidine (C), and uridine (U). However, the transcriptome is a much more diverse landscape than just these four canonical nucleotides; in fact, the transcriptome possesses a suite of over 160 unique modified nucleotides (RNA modifications) (Boccaletto et al. 2018). RNA modifications occur naturally in vivo, but they may also arise from exogenous cellular stressors. For example, our group has recently linked air pollution reactive oxygen species (ROS) exposure to increases in the levels of the 8-oxoG RNA modification within a human epithelial lung cell model (Baldridge et al. 2015; Contreras et al. 2020; Gonzalez-Rivera et al. 2020a, c). Using an 8-oxoG antibody-based pulldown approach, our group has determined that environmental air pollution exposure differentially impacts the oxidation levels of a select population of mRNA transcripts (Contreras et al. 2020; Gonzalez-Rivera et al. 2020a, c). As part of this study, the differentially oxidized farnesyl-diphosphate farnesyltransferase 1 (FDFT1) transcript, which encodes for squalene synthase (an important enzyme in the cholesterol biosynthesis pathway), was linked to a significant reduction in cell membrane integrity (Contreras et al. 2020; Gonzalez-Rivera et al. 2020a). This work demonstrates, in part, the notable downstream impacts that RNA modifications can have toward cellular homeostasis (Contreras et al. 2020; Gonzalez-Rivera et al.

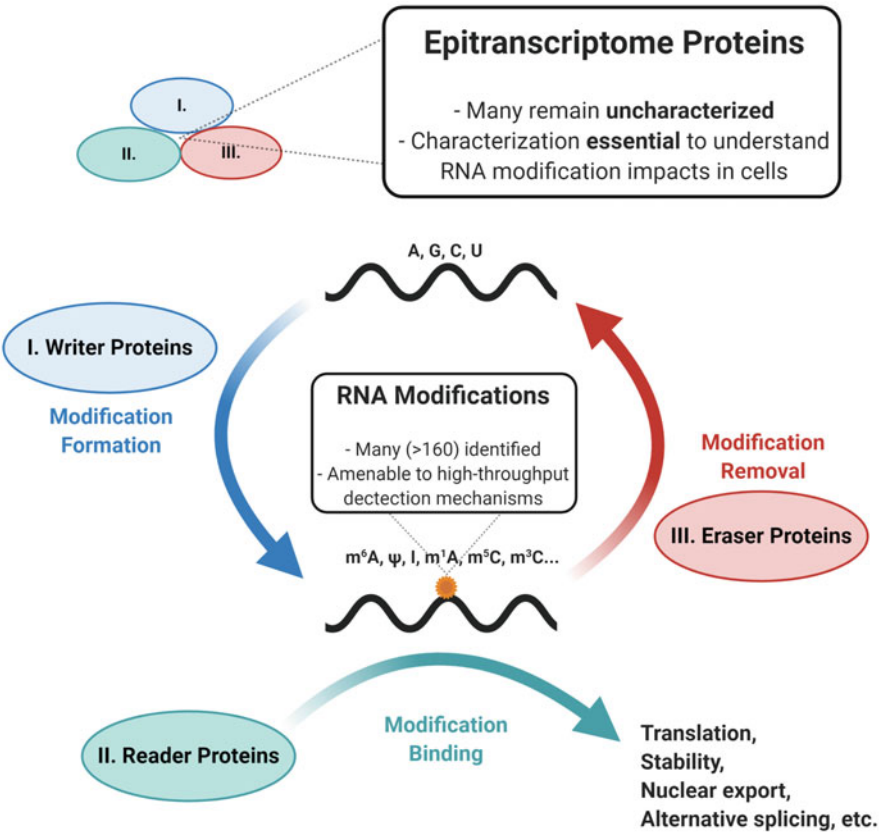


Fig. 1 The RNA modification landscape is regulated by a system of epitranscriptome writer (I), reader (II), and eraser (III) proteins. The identification and characterization of these epitranscriptome proteins is essential toward developing a better understanding of the RNA modifications that they regulate

2020a, c). Recent advances in high-throughput sequencing strategies have proven successful toward identifying the location and prevalence of a few of the most common RNA modifications; these strategies are reviewed extensively in Gilbert et al. (2016); Li et al. (2016b); Helm and Motorin (2017); and Anreiter et al. (2020). However, to date, many of the over 160 known RNA modifications have yet to be characterized in this way.

RNA modifications are largely non-static entities; rather, many well-studied RNA modifications (e.g., m^6A) have been shown to be regulated by an overarching network of epitranscriptome proteins (Zhao et al. 2016; Patil et al. 2018; Yang et al. 2018; Zaccara et al. 2019). These epitranscriptome proteins fall into three distinct classes: (I) writers, responsible for the formation of RNA modifications; (II) readers, responsible for binding to and acting upon RNA modifications; and (III) erasers, responsible for the removal of RNA modifications (Fig. 1). At present, these

epitranscriptome proteins are even less characterized than their cognate RNA modifications; yet, identifying and understanding the function of these proteins is needed for comprehensive understanding of the diverse roles of different RNA modifications. The lack of knowledge regarding epitranscriptome proteins exposes a fundamental gap in understanding the roles of RNA modifications, thereby sparking the need for additional research efforts to be directed toward these studies.

In this chapter, we review the current methods and techniques that are available for the identification and characterization of epitranscriptome proteins. We take a 2-pronged approach toward reviewing the state of this field, wherein we analyze both the experimental (wet-lab) and computational-based methods for aiding the discovery and classification of these proteins. We conclude the chapter with a brief commentary on the ways in which advancements in the field of epitranscriptome protein characterization are crucial toward the development of therapeutic strategies for mitigating the negative health impacts related to aberrant RNA modifications in vivo.

2 Experimental Approaches Used for Identifying and Characterizing Proteins Responsible for Epitranscriptome Activity

Here, we provide an analysis into the current methods that exist for experimental discovery and characterization of epitranscriptome readers, writers, and erasers. We begin with an examination of the many ways in which the abundance and landscape of specific RNA modifications can be controlled. We then detail established methods for the extraction and identification of proteins that interface with modified RNAs. Finally, we conclude this section by providing an overview of the techniques available for further characterization of nascent epitranscriptome proteins.

2.1 Techniques for Controlling the Landscape of Specific RNA Modifications

Studies of modified RNA bases and their cellular metabolism have revealed the importance of understanding the functions of proteins that regulate the epitranscriptome. Two examples of fundamental critical questions are: (i) How does an increase in the abundance of different types of RNA modifications impact the activity and expression of the proteome? And (ii) how does the depletion of specific proteins impact the global abundance of different types of RNA modifications?

At present, the synthetic biology toolbox includes many strategies for targeted perturbations to pathways involved in the intracellular regulation of the

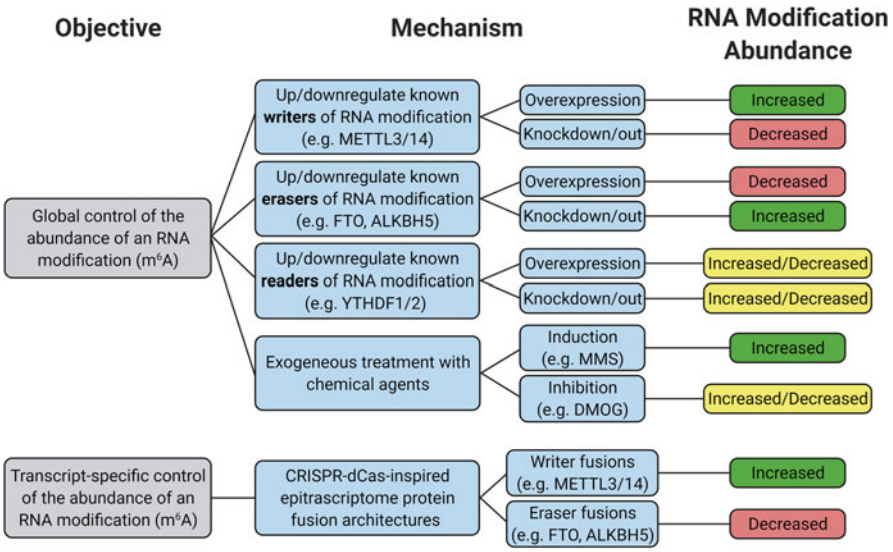


Fig. 2 Flowchart depicting available strategies for controlling the abundance of an RNA modification of interest, using the common m^6A modification as an example

epitranscriptome. These new approaches have enabled focused studies of the interplay between epitranscriptome proteins and their cognate modified RNA bases. Herein, we describe many of the common practices employed for targeted control of the epitranscriptome landscape; these techniques have proven to be valuable in uncovering the unique functional relevance of epitranscriptome proteins and of the few well-characterized RNA modifications (e.g., m^6A) (Fig. 2).

2.1.1 Genetic Deletions and Overexpression of Epitranscriptome Proteins for Tuning Global Abundance of RNA Modifications

When considering methods aimed at altering the relative cellular amounts of an RNA modification, the focus has often been directed toward modulating the activity of known modification writer/eraser proteins. This approach is highly limiting—major writers and erasers have only been identified for a few well-studied RNA modifications (e.g., m^6A); furthermore, mechanisms that regulate cellular abundance of these proteins are well-documented (Table 1). As a commonly applied strategy, functional depletion of major writers and erasers have served to modulate abundance of a specific RNA modification. For instance, knockdown of established m^6A, Ψ, m^1A, and m^5C writers, such as METTL3 (Dominissini et al. 2012; Liu et al. 2014, 2015; Zhao et al. 2014; Wang et al. 2014, 2015; Meyer et al. 2015; Edupuganti et al. 2017; Huang et al. 2018; Xiang et al. 2018; Song et al. 2019), TRUB1 (Safra et al. 2017), TRTM6 (Zheng et al. 2020), and NSUN2 (Yang et al. 2017), respectively,

Table 1 Endogenous methods for regulation of global RNA modification abundance

Modification	Protein class	Protein name	Regulation type	Technique	Cell type
N6-Methyladenosine (m^6A)	Writer	METTL3	Knockdown	siRNA transfection	3 T3-L1 (Zhao et al. 2014); HepG2 (Dominissini et al. 2012); Hela (Liu et al. 2014; Wang et al. 2014, 2015; Meyer et al. 2015; Edupuganti et al. 2017); HEK293T (Liu et al. 2015); HEK293FT (Liu et al. 2014); NMVC (Song et al. 2019)
				shRNA transduction	HepG2 (Huang et al. 2018); HeLa (Huang et al. 2018; Xiang et al. 2018); HEK293FT (Xiang et al. 2018); H9c2 (Song et al. 2019)
			Knockout	CRISPR-Cas9-mediated	Stra8-GFPCre Knockin Mouse Model (Lin et al. 2017)
			Overexpression	Mammalian expression vector transfection	H9c2 (Song et al. 2019)
		METTL14	Knockdown	siRNA transfection	HeLa, HEK293FT (Liu et al. 2014); HEK293T (Liu et al. 2015)

(continued)

Table 1 (continued)

Modification	Protein class	Protein name	Regulation type	Technique	Cell type
N6-Methyladenosine (m^6A)				shRNA transduction	HepG2, HEK293T (Huang et al. 2018); HeLa (Huang et al. 2018; Xiang et al. 2018); HEK293FT (Xiang et al. 2018)
			Knockout	CRISPR-Cas9-mediated	Stra8-GFPCre Knockin Mouse Model (Lin et al. 2017)
		METTL16	Knockdown	siRNA transfection	HeLa (Warda et al. 2017)
	Eraser	FTO	Knockdown	siRNA transfection	HeLa, HEK293FT (Jia et al. 2011); 3 T3-L1 (Zhao et al. 2014)
				shRNA transduction	MEF (Meyer et al. 2015); H9c2 (Song et al. 2019)
			Knockout	Double nickase plasmid transfection	HEK293 (Sindelar et al. 2019)
			Overexpression	Mammalian expression vector transfection	HeLa (Jia et al. 2011); HEK293 (Meyer et al. 2015)
				Mammalian expression vector transduction	MEF (Meyer et al. 2015)
			Chemical inhibition	2-HG inhibition of FTO	TF-1 (Sindelar et al. 2019)

(continued)

Table 1 (continued)

Modification	Protein class	Protein name	Regulation type	Technique	Cell type
N6-Methyladenosine (m^6A)		ALKBH5	Knockdown	siRNA transfection	HeLa (Zheng et al. 2013); 3 T3-L1 (Zhao et al. 2014); NMVC (Song et al. 2019)
				shRNA transduction	H9c2 (Song et al. 2019); HTR8/Svneo (Zheng et al. 2020)
			Overexpression	Mammalian expression vector transfection	HeLa (Zheng et al. 2013); H9c2 (Song et al. 2019)
				Mammalian expression vector transduction	NMVC (Song et al. 2019)
	Reader	YTHDF2	Knockdown	siRNA transfection	HeLa (Wang et al. 2014); HT-22 (Wu et al. 2019)
			Overexpression	Mammalian expression vector transfection	HeLa (Wang et al. 2014)
				Mammalian expression vector transduction	HEK293FT (Xiang et al. 2018)
		FMR1	Overexpression	CRISPR-Cas9-mediated	HeLa (Edupuganti et al. 2017)
		IGF2BP1/2/3	Knockdown	shRNA transduction	HepG2, HeLa, CD34+ (Huang et al. 2018)
			Knockout	CRISPR-Cas9-mediated	HepG2 (Huang et al. 2018)
			Overexpression	Mammalian expression vector transfection	HeLa, HEK293T (Huang et al. 2018)
		Prrc2a	Knockdown	siRNA transfection	HT-22, GL261 (Wu et al. 2019)

(continued)

Table 1 (continued)

Modification	Protein class	Protein name	Regulation type	Technique	Cell type
N6-Methyladenosine (m^6A)			Overexpression	Mammalian expression vector transduction	HT-22, GL261 (Wu et al. 2019)
		YTHDF1	Knockdown	siRNA transfection	HeLa (Wang et al. 2015)
			Overexpression	Mammalian expression vector transfection	HeLa (Wang et al. 2015)
		HNRNPC	Knockdown	siRNA transfection	HEK293T (Liu et al. 2015)
		HNRNPD	Knockdown	siRNA transfection	H9c2, NMVC (Song et al. 2019)
8-oxo-7,8-Dihydroguanosine (8-oxoG)	Reader	hPNP	Knockdown	siRNA transfection	HeLa (Hayakawa and Sekiguchi 2006; Wu and Li 2008)
			Overexpression	Mammalian expression vector transfection	HeLa (Wu and Li 2008)
		HNRNPC	Knockdown	siRNA transfection	HeLa MR (Hayakawa et al. 2010)
		DAZAP1	Knockdown	siRNA transfection	HeLa MR (Hayakawa et al. 2010)
		SF3B4	Knockdown	siRNA transfection	HeLa MR (Hayakawa et al. 2010)
		HNRNPD	Knockdown	siRNA transfection	HeLa MR (Hayakawa et al. 2010)
		PCBP1	Knockout	CRISPR-Cas9-mediated	HeLa S3 (Ishii et al. 2018)
Pseudouridine (Ψ)	Writer	TRUB1	Knockdown	siRNA transfection	HEK293T (Safra et al. 2017)
			Overexpression	Mammalian expression vector transfection	HEK293T (Safra et al. 2017)
		PUS7	Knockdown	siRNA transfection	HEK293T (Safra et al. 2017)

(continued)

Table 1 (continued)

Modification	Protein class	Protein name	Regulation type	Technique	Cell type
5-Methylcytosine (m^5C)	Writer	NSUN2	Knockdown	siRNA transfection	HeLa (Yang et al. 2017)
			Overexpression	Mammalian expression vector transfection	HeLa (Yang et al. 2017)
		TRM4B	Overexpression	Plant expression vector transfection	*Arabidopsis thaliana* (David et al. 2017)
	Reader	ALYREF	Knockdown	siRNA transfection	HeLa (Yang et al. 2017)
			Overexpression	Mammalian expression vector transfection	HeLa (Yang et al. 2017)
		YTHDF2	Knockout	CRISPR-Cas9-mediated	HEK293T (Dai et al. 2020)
N1-Methyladenosine (m^1A)	Writer	TRMT6	Knockdown	shRNA transduction	HTR8/Svneo (Zheng et al. 2020)
	Eraser	ALKBH3	Knockdown	siRNA transfection	HeLa (Seo and Kleiner 2020)
				shRNA transduction	HTR8/Svneo (Zheng et al. 2020)
		AlkB	Overexpression	Bacterial expression vector transformation	*Escherichia coli* (Ougland et al. 2004)
	Reader	YTHDF1	Knockdown	siRNA transfection	HeLa (Dai et al. 2018)
		YTHDF2	Knockdown	siRNA transfection	HeLa (Dai et al. 2018; Seo and Kleiner 2020)
		YTHDF3	Knockdown	siRNA transfection	HeLa (Dai et al. 2018)
				shRNA transduction	HTR8/Svneo (Zheng et al. 2020)
Inosine (I)	Writer	ADAR1	Overexpression	Mammalian expression vector transduction	HEK293FT (Xiang et al. 2018)

(continued)

Table 1 (continued)

Modification	Protein class	Protein name	Regulation type	Technique	Cell type
3-Methylcytidine (m^3C)	Writer	METTL8	Knockout	CRISPR-Cas9-mediated	HeLa S3, HCT116 (Xu et al. 2017)
	Eraser	AlkB	Overexpression	Bacterial expression vector transformation	*Escherichia coli* (Reichle et al. 2019)
		ALKBH1	Knockdown	siRNA transfection	HEK293T (Ma et al. 2019)
			Overexpression	Mammalian expression vector transfection	HEK293T (Ma et al. 2019)

have been performed to reduce the overall cellular abundance of these modifications. In each case, global reduction of the specific modification within the entire RNA pool has been confirmed by using semi-quantitative measurements such as mass spectrometry and *h*igh-*p*erformance *l*iquid *c*hromatography (HPLC) (Jia et al. 2011; Zheng et al. 2013; Liu et al. 2014; Wang et al. 2014; Yang et al. 2017). It is worth noting that while the overall pool of a specific type of modification can be successfully reduced by genetic knockdown of writers and erasers, this strategy does not enable targeted reduction of a specific modification within a specific RNA.

Various gene knockout approaches have been previously implemented to reduce intracellular abundance of modified RNA writers and erasers. siRNA transfection has been used to induce transient knockdown of the targeted protein (Dominissini et al. 2012; Liu et al. 2014, 2015; Zhao et al. 2014; Wang et al. 2014, 2015; Meyer et al. 2015; Edupuganti et al. 2017; Safra et al. 2017; Yang et al. 2017; Warda et al. 2017; Song et al. 2019), whereas more stable gene knockdowns have been achieved in cell cultures by means of shRNA transduction (Meyer et al. 2015; Huang et al. 2018; Xiang et al. 2018; Song et al. 2019; Zheng et al. 2020). Moreover, CRISPR-Cas9-mediated knockout of the writer/eraser encoding genes has also been performed to completely abolish expression of the protein from the organism of study and thereby achieve more significantly reduced abundance of the targeted RNA modification (Xu et al. 2017; Lin et al. 2017). Overexpression of known writers, often achieved through transfection of cell cultures with an expression vector encoding for the protein (Safra et al. 2017; Yang et al. 2017; David et al. 2017; Xiang et al. 2018; Song et al. 2019), has been shown to elicit the expected opposite impact on the transcriptome-wide abundance of the targeted modifications—resulting in global increases of the RNA modification. Moreover, targeting established eraser proteins for an RNA modification has proven to be another effective method for modulating RNA modification abundance. In this case,

Table 2 Exogenous methods for regulation of global RNA modification abundance

Modification	Technique	Reagent used	Cell type
8-oxo-7,8-Dihydroguanosine (8-oxoG)	Chemical induction	Hydrogen peroxide (H_2O_2)	HeLa (Hayakawa and Sekiguchi 2006; Wu and Li 2008; Hayakawa et al. 2010); HeLa S3 (Ishii et al. 2018)
		Menadione	HeLa (Hayakawa and Sekiguchi 2006; Wu and Li 2008)
General alkylation	Chemical induction	Methyl-methanesulfonate (MMS)	*Escherichia coli* (Reichle et al. 2018, 2019)
		Streptozotocin	*Escherichia coli* (Reichle et al. 2019)
N6-Methyladenosine (m^6A)	Chemical inhibition	Dimethyloxalylglycine (DMOG)	Jurkat (Perez-Perri et al. 2018)

functional knockdown/knockout of targeted eraser proteins, such as the m^6A eraser FTO, increases overall levels of the RNA modification of study while overexpression reduces overall abundance of the associated RNA modification (Ougland et al. 2004; Jia et al. 2011; Zheng et al. 2013, 2020; Zhao et al. 2014; Meyer et al. 2015; Ma et al. 2019; Sindelar et al. 2019; Song et al. 2019; Reichle et al. 2019; Seo and Kleiner 2020).

The impact of manipulating reader protein expression for altering global levels of intracellular RNA modification abundance is not as intuitive as that of manipulating writer and eraser protein expression. A major challenge present when manipulating expression of reader proteins to achieve predictable changes in overall abundance of specific types of RNA modifications is that this diverse class of proteins may differently impact their RNA targets. For example, the m^6A reader YTHDF2 is responsible for directing its target RNAs toward degradation (Wang et al. 2014; Edupuganti et al. 2017; Rauch et al. 2018). As a result, overexpression/depletion of this protein elicits a response similar to the one achieved from controlling the expression of eraser proteins (Wang et al. 2014). Conversely, other reader proteins (e.g., IGF2BP2) have been shown to be involved with increasing target RNA transcript stability and thus preserve RNA modification abundance (Huang et al. 2018).

Beyond genetically altering the levels of proteins that are known to be responsible for the formation/removal of RNA modifications endogenously, alternative chemical approaches have been utilized in shifting the abundance of RNA modifications within an RNA pool; this approach has been specifically useful for studies involving modifications that arise from exposure to environmental stressors (Table 2). As an example, culturing of cells with low concentrations of general alkylative agents, such as methyl-methane sulfonate (MMS) (Reichle et al. 2018, 2019) and streptozotocin (Reichle et al. 2019), has been used to induce the unspecific formation of myriad RNA methylation modifications (e.g., m^6A, m^1A, m^7G, m^3C, m^3U, etc.)

on *Escherichia coli* tRNA. Conversely, a reduction in overall methylation modifications has been achieved through incubation of cells with 0.5 mM dimethyloxalylglycine (DMOG), a general inhibitor of methyltransferase activity (Perez-Perri et al. 2018). Exogenous mechanisms for controlling RNA modification abundance are most helpful when studying modifications for which there are not well-characterized/known writer and eraser proteins. Notably, the abundance of 8-oxoG, the most widespread RNA modification resulting from cellular oxidative stress, can be increased through culturing of cells with low concentrations of oxidative agents such as hydrogen peroxide (H_2O_2) (Hayakawa and Sekiguchi 2006; Wu and Li 2008; Hayakawa et al. 2010; Ishii et al. 2018; Gonzalez-Rivera et al. 2020c) and menadione (Hayakawa and Sekiguchi 2006; Wu and Li 2008). Exogenous H_2O_2 treatments have been useful, in part, toward elucidating the contributions of various mutations to 8-oxoG readers (i.e., PNPase) under conditions of higher cellular oxidation (Gonzalez-Rivera et al. 2020c). A major limitation of these exogenous approaches resides in the lack of control over tuning the abundance of a specific RNA modification of interest.

An additional option available for regulating the levels of an RNA modification is through induced chemical inhibition of writers and erasers. As an example, FTO m^6A-demethylation activity can be inhibited in vivo with addition of D-2-hydroxyglutarate (2-HG), an FTO co-factor competitor (Sindelar et al. 2019). Such inhibition has been achieved by expressing mutant forms of the isocitrate dehydrogenase enzymes (IDH1/2) to generate high levels of 2-HG in cells for stable inhibition of FTO (Sindelar et al. 2019). This type of strategy avoids the need for siRNA transient transfection or other forms of specific epitranscriptome protein knockdown.

Overall, the methods outlined above present useful tools to adjust native global abundance of specific types of RNA modifications to begin to answer many fundamental questions regarding the function of epitranscriptome proteins. It is worth noting that in some cases, combinations of different methods have been used simultaneously for tuning intracellular abundance of RNA modifications in more extreme ways (e.g., double knockdown of writers) (Huang et al. 2018; Xiang et al. 2018).

2.1.2 Targeted Methods for the Formation/Removal of RNA Modifications Within Specific RNA Transcripts

Recently, CRISPR-dCas-based systems have been developed to facilitate targeted RNA modification formation and removal (Table 3). These systems use fusions of known epitranscriptome writers/erasers and catalytically dead Cas endonucleases (dCas). The resultant fusion proteins are capable of interfacing with CRISPR guide RNAs (gRNAs) for the delivery/removal of RNA modifications in a sequence

Table 3 Targeted methods for regulation of RNA transcript-specific RNA modification abundance

Modification	Fusion protein architecture	Effector protein class	Effector protein	Cell type
N6-Methyladenosine (m^6A)	CRISPR-dCas9-effector	Writer	METTL3/14	HeLa, MEF (Liu et al. 2019)
		Eraser	ALKBH5	HeLa (Liu et al. 2019)
			FTO	HeLa (Liu et al. 2019)
	CRISPR-dCas13-effector	Reader	YTHDF1	HEK293T, HeLa (Rauch et al. 2018)
			YTHDF2	HEK293T, HeLa (Rauch et al. 2018)
		Writer	METTL3	HEK293T (Wilson et al. 2020)
			METTL3/14	HEK293T (Wilson et al. 2020)
	CIRTS	Reader	YTHDF1	HEK293T (Rauch et al. 2019)
			YTHDF2	HEK293T (Rauch et al. 2019)
Inosine (I)	CRISPR-dCas13-effector	Writer	ADAR1	HEK293FT (Cox et al. 2017)
			ADAR2	HEK293FT (Cox et al. 2017)
	CIRTS	Writer	ADAR2	HEK293T (Rauch et al. 2019)

specific manner. These systems have been developed using dCas9 (Rau et al. 2019; Liu et al. 2019) and dCas13 (Cox et al. 2017; Wilson et al. 2020) endonucleases, both having demonstrated impressive specificity and selectivity toward target RNA sequences in vitro (Rau et al. 2019) and in vivo (Cox et al. 2017; Rau et al. 2019; Liu et al. 2019; Wilson et al. 2020). CRISPR-dCas based systems have thus far been predominately used for studying the site-specific formation and removal of the m^6A modification, for which the writer and eraser proteins are well-established (Liu et al. 2019; Wilson et al. 2020). In theory, such a system architecture has the capacity to deliver targeted formation/removal of any RNA modification which possesses a known writer or eraser protein. These systems have been employed across various human cell lines with similar success, providing evidence toward the robustness of this targeted approach (Cox et al. 2017; Liu et al. 2019). Moreover, while the effectiveness of using these systems for targeting non-coding RNAs (ncRNAs) has not yet been investigated, the demonstrated sequence-specific targeting of mRNAs suggests that these tools would be amenable for studying other classes of RNAs.

*C*RISPR-*i*nspired *R*NA *t*argeting *s*ystems (CIRTS) constructed entirely from human-derived protein parts have similarly been developed to deliver reader and writer protein functionalities in a manner that is specific to sequences within targeted transcripts (Rauch et al. 2019). This method substitutes the comprehensive dCas protein with a suite of human protein domains which mimic the dCas functionality:

(i) an RNA-binding domain which specifically recognizes a hairpin on the gRNA strand, (ii) a non-specific RNA-binding domain which protects the gRNA strand during transport, and (iii) an effector domain which possesses the desired epitranscriptome reader/writer/eraser functionality. CIRTS have successfully been developed to deliver m^6A readers and A-to-I writers to target mRNA transcripts (Rauch et al. 2019).

Furthermore, m^6A writers and erasers have been genetically fused to engineered RNA-binding Pumilio and FBF homology (PUF) proteins for the delivery of site-specific formation and removal of RNA modifications (Shinoda et al. 2020). PUF proteins possess mRNA-binding regions which, through rational mutagenesis, have been engineered to recognize specific mRNA sequence fragments (Zhao et al. 2018). Thus, genetic fusions of PUF proteins with epitranscriptome effector proteins allow for an alternative mechanism to specifically target mRNA transcripts—one that does not require the addition of gRNAs to guide delivery (Shinoda et al. 2020).

Having the capability to deliver targeted epitranscriptome regulation at the RNA transcript level allows more acute studies of the roles of the proteins associated with the RNA modification, as well as of the role of the RNA modification itself on the fate of the RNA. A major limitation to these targeted delivery approaches is their lack of high-throughput capacity. For example, CRISPR/CIRTS-based approaches require the transfection of a new gRNA template for each transcript being targeted. Likewise, PUF protein fusions require a unique engineered PUF protein construct for each targeted RNA. Also inherent to these sequence-specific targeting methods is the concern over off-target binding effects. It is important to interrogate these off-target impacts using high-throughput sequencing approaches during the development of these systems.

2.2 Techniques for Extraction and Identification of Proteins Which Interface with Modified RNAs

2.2.1 Affinity-Based Extraction and Identification of Epitranscriptome Reader Proteins Using Modified RNA Baits

Discovery of native proteins that interact with modified RNAs (or "reader" proteins) requires extraction of proteins that interface with a given RNA modification of interest as well as identification of the recovered proteins. Methods for extracting proteins that recognize and associate with modified RNAs largely follow a consistent workflow (Fig. 3a). These techniques involve affinity-based capture of proteins that bind to functionalized RNA bait sequences; these sequences (~10–50 nucleotides in length) are constructed synthetically to include RNA modifications of interest. In these protocols, cell cultures are grown and lysed to obtain a protein lysate solution. Following lysis, proteins are incubated with short RNA oligomers that contain either the RNA modification of interest or an unmodified version of the same sequence (Hayakawa et al. 2010; Dominissini et al. 2012; Arguello et al. 2017; Edupuganti

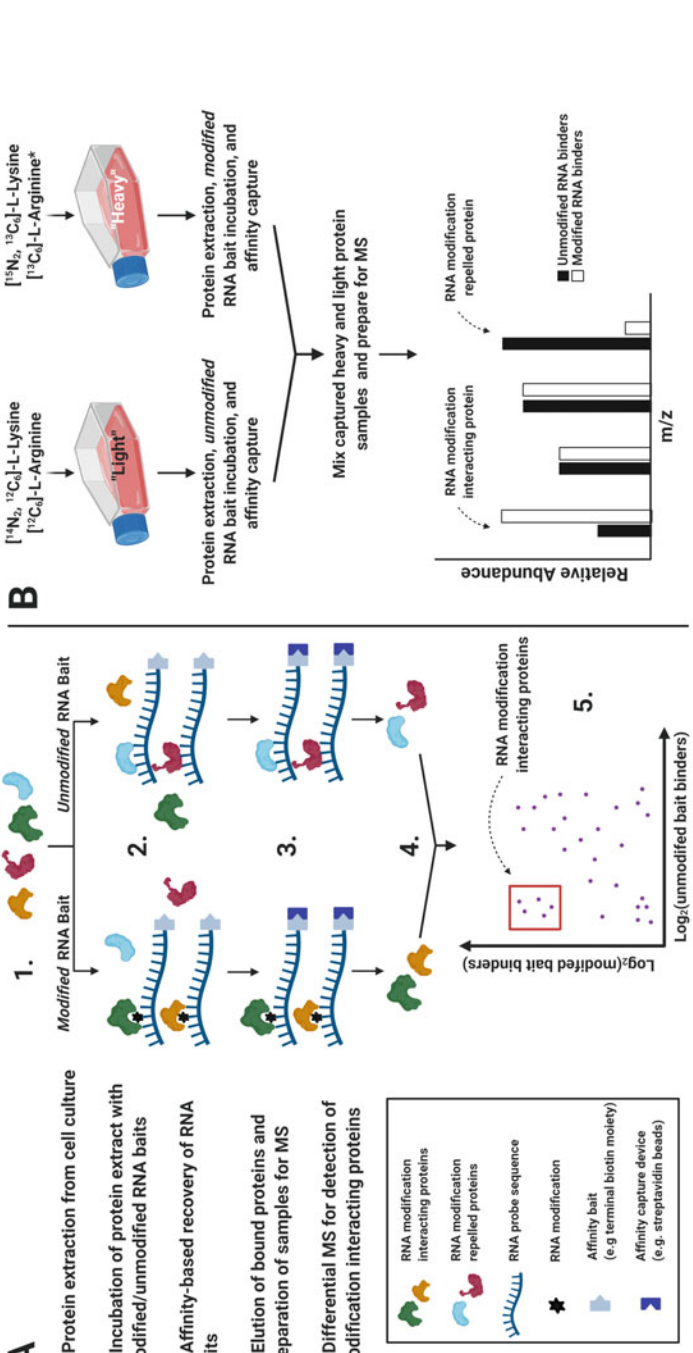


Fig. 3 Common methods for extraction and identification of RNA modification interacting proteins. (**a**) RNA bait affinity pulldown and mass spectrometry analysis. The square on the mass spectrometry plot indicates proteins which associate strongly with the modified RNA bait over the unmodified RNA bait. (**b**) SILAC approach for differential mass spectrometry analysis of proteins recovered from RNA bait affinity pulldowns. RNA modification interacting proteins are identified by comparing the relative abundance of protein fragments isolated from heavy versus light protein lysates. *[$^{13}C_6$, $^{15}N_4$]-L-Arginine can also be used in SILAC heavy media for isotopic labeling. MS = *mass spectrometry*

et al. 2017; Yang et al. 2017; Dai et al. 2018, 2020; Huang et al. 2018; Sajini et al. 2019; Wu et al. 2019; Seo and Kleiner 2020; Zheng et al. 2020). The design of these oligomers is dependent upon both the RNA modification studied and the desired level of binding protein specificity. For modifications that possess known consensus motifs (e.g., GG-(m^6A)-CU), oligomer sequences house the targeted modification within the associated motif (Arguello et al. 2017; Edupuganti et al. 2017; Huang et al. 2018; Wu et al. 2019; Seo and Kleiner 2020). Alternatively, surrounding the modification of interest with a randomized sequence can serve to identify proteins which bind independent of sequence context (Edupuganti et al. 2017). Other studies have taken an additional step in designing RNA oligomers that mimic endogenous transcripts known to consistently harbor the RNA modification (Dominissini et al. 2012; Dai et al. 2018, 2020; Sajini et al. 2019; Zheng et al. 2020). A common feature of many affinity-based methods to capture reader proteins has been the use of multiple instances of the RNA modification on the RNA oligomer sequence (e.g., overrepresenting the modification) as a strategy to increase protein recruitment. Increasing the amount of available modified RNA sites might also serve to increase the background RNA bait binding signal in subsequent mass spectrometry analysis steps. It is worth mentioning that a question raised by this affinity capture technique is whether or not recognition of the RNA modification by the reader protein is likely to occur in the context of native cellular RNAs, since modified bases such as m^6A and 8-oxoG have been found to exist at ratios of approximately 1 m^6A per 400 A mRNA nucleotides and 1 8-oxoG per 10^5 G mRNA nucleotides in mammalian cell lines (Shen et al. 2000; Shu et al. 2020).

An important characteristic of designed RNA baits is that they also must be functionalized to allow for affinity purification. Often, the addition of a terminal biotin moiety to the RNA oligomer has been used to allow affinity capture and pulldown with streptavidin-conjugated precipitation beads (Dominissini et al. 2012; Arguello et al. 2017; Edupuganti et al. 2017; Yang et al. 2017; Dai et al. 2018, 2020; Huang et al. 2018; Wu et al. 2019; Seo and Kleiner 2020; Zheng et al. 2020). Alternatively, these RNA baits may be covalently linked to dihydrazide agarose beads (Sajini et al. 2019) or may be captured through base pair annealing of poly (A) RNA bait tails with biotinylated oligo(dT) molecules (Hayakawa et al. 2010). For increased RNA oligomer binding retention, biotin-streptavidin affinity capture can be coupled with photo-activatable cross-linking of RNA baits to bound proteins by including a photo-activatable diazirine nucleoside flanking the RNA modification of interest (Arguello et al. 2017; Seo and Kleiner 2020). This coupling strategy enhances binding protein capture retention without sacrificing binding specificity between modified and unmodified bait sequences (Seo and Kleiner 2020). After affinity capture and pulldown of RNA bait sequences, bound protein readers are either directly subjected to elution and trypsin digestion for mass spectrometry analysis (Arguello et al. 2017; Edupuganti et al. 2017; Sajini et al. 2019; Seo and Kleiner 2020) or are first separated through gel electrophoresis and stained to identify concentrated protein clusters. In cases of the latter, slices of resolved gel that contain enriched protein samples are excised and subsequently digested with trypsin (Hayakawa et al. 2010; Dominissini et al. 2012; Yang et al. 2017; Dai et al.

2018, 2020; Huang et al. 2018; Wu et al. 2019; Zheng et al. 2020). Finally, after appropriate preparations, digested protein samples are subjected to mass spectrometry analysis to identify the proteins that preferentially associate with the modified RNA baits relative to those that associate with control unmodified RNA baits (Hayakawa et al. 2010; Dominissini et al. 2012; Arguello et al. 2017; Edupuganti et al. 2017; Yang et al. 2017; Dai et al. 2018, 2020; Huang et al. 2018; Sajini et al. 2019; Wu et al. 2019; Seo and Kleiner 2020; Zheng et al. 2020).

The modified RNA pulldown approaches described above have been beneficial to the discovery of epitranscriptome readers for the common m^6A (Dominissini et al. 2012; Arguello et al. 2017; Yang et al. 2017; Huang et al. 2018; Wu et al. 2019), m^5C (Yang et al. 2017; Sajini et al. 2019; Dai et al. 2020), m^1A (Dai et al. 2018; Seo and Kleiner 2020; Zheng et al. 2020), and 8oxoG (Hayakawa et al. 2010) modifications. ALKBH5, an m^6A eraser protein, has also been identified through an RNA modification affinity pulldown approach (Arguello et al. 2017); however, these techniques lack means for characterizing eraser and writer protein functionalities. While mass spectrometry is an effective approach for the high-throughput identification of proteins recovered from affinity-based RNA pulldowns, it has been shown to produce false positives (Hayakawa et al. 2010; Edupuganti et al. 2017) and thus requires a secondary confirmation; this confirmation has been achieved through complementary low-throughput techniques such as Western blotting (Hayakawa et al. 2010; Dominissini et al. 2012; Arguello et al. 2017; Yang et al. 2017; Huang et al. 2018) or *e*lectrophoretic *m*obility *s*hift *a*ssay (EMSA) analysis between the mass spectrometry-identified binding protein and the modified RNA bait used for extraction (Yang et al. 2017; Huang et al. 2018; Wu et al. 2019; Seo and Kleiner 2020). It is worth noting that, although these assays can collectively confirm modified RNA-protein binding pairs, they cannot determine the functional impact of the protein on the RNA.

A similar strategy that has been used for discovery of epitranscriptome reader proteins has utilized *s*table *i*sotope *l*abeling by *a*mino acids in *c*ell culture (SILAC) modified affinity capture. The SILAC approach closely aligns with the RNA bait affinity pulldown approaches described above while observing a few slight adjustments (Fig. 3b). SILAC approaches involve the use of cell culture media containing either normal (light) amino acids or media supplemented with isotopically labeled (heavy) lysine and arginine residues (Edupuganti et al. 2017; Dai et al. 2018, 2020; Sajini et al. 2019). Cell cultures are passaged extensively to ensure total adoption of heavy amino acid isotopes into the proteins expressed by cells cultured in the heavy media. Following lysis, light protein lysates are incubated with modified RNA baits, and heavy protein lysates are incubated with unmodified RNA baits or vice versa (Edupuganti et al. 2017; Dai et al. 2018, 2020; Sajini et al. 2019). Following preparation of protein samples using on-bead (Edupuganti et al. 2017; Sajini et al. 2019) or in-gel (Dai et al. 2018, 2020) trypsin digestion, proteins can be identified by mass spectrometry. The peptide fragments extracted from heavy cell culture, having a unique isotopic mass spectrometry signature, are easily distinguished from peptide fragments extracted from light cell culture. As a result, the specific protein abundance in modified vs. unmodified RNA bait pulldowns can be directly compared as

part of the same mass spectrometry run sample. To elucidate any potential binding bias that might result from the isotopic labeling of protein samples, an additional experimental replicate is typically performed wherein the heavy and light protein lysates are swapped and incubated with the opposite RNA bait variety (modified/unmodified) (Edupuganti et al. 2017; Dai et al. 2018, 2020). As is the case for the aforementioned RNA bait affinity pulldown assays, further binding confirmation experiments are required to demonstrate preferential binding between the mass spectrometry-identified protein and the modified RNA oligomer (Edupuganti et al. 2017; Dai et al. 2018, 2020; Sajini et al. 2019).

An additional, less investigated method for epitranscriptome protein discovery is the *e*nhanced *R*NA *i*nteractome *c*apture (eRIC) strategy (Perez-Perri et al. 2018); this method has been applied to one study thus far. Unlike the two methods mentioned above for epitranscriptome protein extraction, eRIC does not rely on the incubation of protein lysates with modified/unmodified RNA oligomers. Instead, this method relies on altering the global landscape of the RNA modification of interest by exogenous chemical treatment to observe changes in the levels of proteins which bind RNA. In this approach, proteins are first cross-linked to their cognate RNAs in vivo through UV irradiation (Perez-Perri et al. 2018). Following cell lysis, proteins attached to poly(A) RNAs are extracted via capture with LNA2.T-coupled magnetic beads. The stable base-pairing interactions of LNA2.T with poly(A) RNA allows for more stringent washing and elution conditions than would be otherwise permitted by oligo(dT)-coupled beads (Perez-Perri et al. 2018). Following washing and protein digestion, samples are analyzed using mass spectrometry to determine proteins that are up/downregulated in response to a global change in the abundance of an RNA modification. As shown in a test case, predictable changes to the abundance of known m^6A reader proteins in response to global inhibition of methyltransferase activity can be successfully detected (Perez-Perri et al. 2018). It is worth noting that this method does not provide direct evidence for the binding of proteins with a particular modified RNA. Instead, it presents a preliminary investigation into the protein interactome of a particular RNA modification.

2.2.2 Co-immunoprecipitation Extraction and Identification of Epitranscriptome Writer and Eraser Proteins

The extraction and identification of epitranscriptome writer and eraser proteins cannot be readily performed by adopting the same modified RNA bait pulldown techniques available for reader proteins. These aforementioned techniques rely on the specific binding of epitranscriptome proteins to the RNA modification within an RNA bait sequence. However, writer and eraser proteins are responsible for forming and removing RNA modifications, respectively; therefore, recovery of writer and eraser proteins through modified RNA bait pulldowns is a futile endeavor. Instead, writers and erasers are commonly identified and/or confirmed by low-throughput assessment of the abundance of a particular type of modified RNA in samples where the protein in question is overexpressed and/or depleted (Jia et al. 2011; Zheng et al.

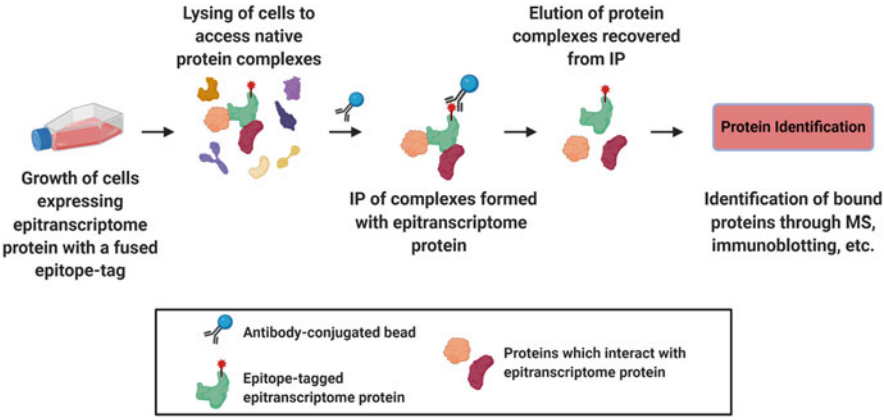


Fig. 4 Outline of the co-IP procedure used to extract and identify proteins which interact with known epitranscriptome proteins in vivo. IP of an epitope-tagged protein of interest also pulls down interacting protein partners, which are subsequently identified. IP = *i*mmuno*p*recipitation. MS = *m*ass *s*pectrometry

2013; Liu et al. 2014; Safra et al. 2017; Xu et al. 2017; Warda et al. 2017). For instance, systematic mass spectrometry measurement of the abundance of m^3C in RNA samples extracted from a series of METTL knockout/overexpression cell lines served to identify METTL8 as an m^3C writer in human mRNAs (Xu et al. 2017).

*Co-i*mmuno*p*recipitation (co-IP) represents another low-throughput strategy for mapping complete protein complexes involved in RNA modification pathways; this strategy is used to identify proteins that interact with other well-characterized epitranscriptome proteins with known activity (Liu et al. 2014; Wang et al. 2014, 2015; Huang et al. 2018; Xiang et al. 2018; Wu et al. 2019) (Fig. 4). The core assumption behind co-IP is that proteins which associate with epitranscriptome proteins in vivo are likely to have similar functionalities. Co-IP relies on the ectopic expression of an epitope-tagged protein of known epitranscriptome activity. Following lysis, protein lysates are incubated with antibody-conjugated beads designed to specifically bind to the epitope tag of the bait protein. Protein-antibody-bead complexes are subsequently extracted from the unbound protein solution through immunoprecipitation, and, after extensive washing, protein complexes are eluted and submitted for mass spectrometry identification (Liu et al. 2014; Wang et al. 2015; Huang et al. 2018). Enhanced protein complex stability has been achieved with co-IP through mild cross-linking treatment (e.g., formaldehyde) prior to immunoprecipitation (Wang et al. 2015). Further confirmation of interacting protein partners has also been accomplished through Western blotting of eluates with protein-specific antibodies (Liu et al. 2014; Wang et al. 2014, 2015; Huang et al. 2018; Xiang et al. 2018; Wu et al. 2019).

Co-IP has been used to help identify the complex formed between METTL3, METTL14, and WTAP proteins which function together to catalyze the formation of the m^6A modification (Liu et al. 2014). Similarly, the use of co-IP has led to uncovering the interaction between two readers of m^6A: YTHDF2 and Prcc2a (Wu et al. 2019). In addition to its utility as a method for uncovering novel epitranscriptome proteins, co-IP has been used to gain deeper insight into the functionality of the epitope-tagged protein. In a study of m^6A reader proteins, co-IP identified several proteins involved in translation control (e.g., translation initiation factor complex 3 (eIF3)) which associate with YTHDF1 in vivo (Wang et al. 2015). This discovery provided a justification for why the YTHDF1 protein increases the translation efficiency of targeted mRNAs (Xu et al. 2017).

2.3 *Strategies for Functional Characterization of Identified Epitranscriptome Proteins*

After identification of a specific epitranscriptome protein, studies have focused on investigating mechanistic details of the protein's role(s) within RNA modification pathways. Many experimental tools exist to guide deeper investigations into epitranscriptome protein characterization (Fig. 5). Here, we describe some of the most widely adopted techniques being applied for determining (i) the preferred RNA-binding motifs and/or binding locations of epitranscriptome proteins, (ii) the impacts of the presence and binding of epitranscriptome proteins on the fate of their targeted RNAs, and (iii) the binding kinetics and catalytic activity of epitranscriptome proteins with their target RNA substrates. In tandem, these methods provide a more wholistic functional understanding of nascent epitranscriptome proteins.

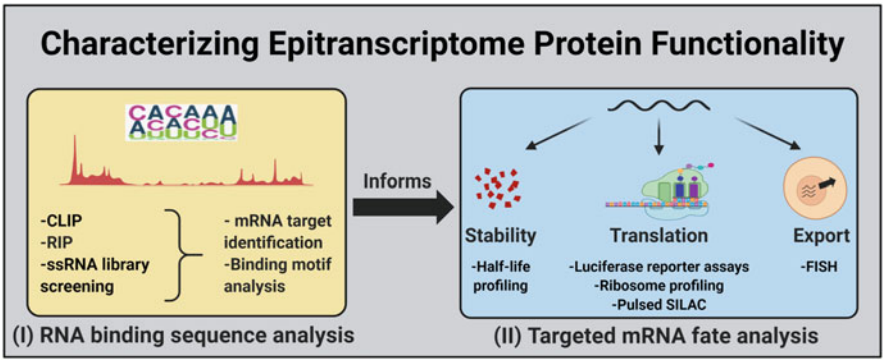


Fig. 5 Strategies for further characterization of epitranscriptome protein functionality. Methods used for identifying the binding site locations of epitranscriptome proteins (**I**) are required to inform studies into the impact of epitranscriptome proteins on the fate of their mRNA target sequences (**II**). CLIP = *c*ross-*l*inking and *i*mmuno*p*recipitation. RIP = *R*NA *i*mmuno*p*recipitation

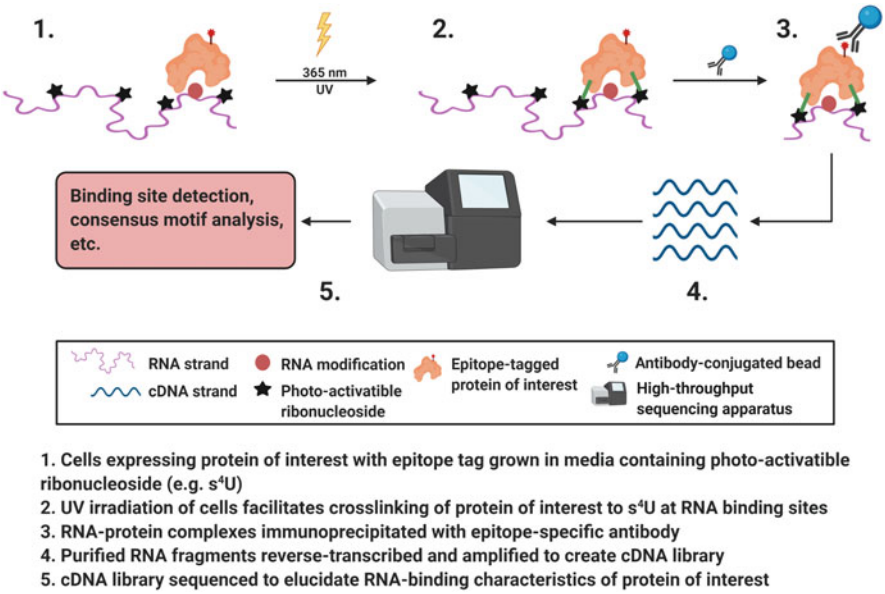


Fig. 6 Schematic depicting the PAR-CLIP technique used to identify mRNA-binding targets of epitranscriptome proteins in vivo. s^4U = 4-thiouridine

2.3.1 Sequence Analysis of Epitranscriptome Protein RNA-Binding Targets

Many of the techniques available for gaining insight into the RNA-binding landscape of epitranscriptome proteins revolve around immunoprecipitation of the complexes formed between an epitranscriptome protein and its RNA targets (Wang et al. 2015; Meyer et al. 2015; Warda et al. 2017; Huang et al. 2018; Song et al. 2019; Zheng et al. 2020). One of the most common immunoprecipitation methods employed for studying the binding profiles of proteins that interact with modified RNAs is through *p*hoto*a*ctivatable *r*ibonucleoside-enhanced *c*ross-*l*inking and *i*mmuno*p*recipitation (PAR-CLIP) (Liu et al. 2014, 2015; Zhao et al. 2014; Wang et al. 2014, 2015; Meyer et al. 2015; Edupuganti et al. 2017; Yang et al. 2017; Huang et al. 2018; Wu et al. 2019) (Fig. 6). In PAR-CLIP experiments, cells expressing an epitope-tagged protein of known epitranscriptome activity are cultured in media containing a photoactivatable ribonucleoside analogue (e.g., 4-thiouridine, s^4U). Cells harboring s^4U speckled RNA are irradiated with a low dose of 365 nm UV light to induce the specific cross-linking of RNA-binding proteins to their RNA targets at s^4U sites. Following cross-linking, lysis, and mild RNase treatment, the resulting total protein lysate is incubated with antibody-coated beads designed to bind the epitope tag on the protein of interest. After pulldown of the RNA-protein complexes, bound RNAs are eluted and processed into a cDNA library for high-

throughput sequencing and subsequent mapping of protein binding locations back onto the transcriptome. The RNA-binding sites for the protein of interest can be mapped with high precision as a result of the specific cross-links formed between the epitranscriptome proteins and the s^4U nucleotides present at/near the binding sites. Upon the elution of RNA fragments from their protein binding partners, small fragments of the protein remain on the RNAs at the location of cross-linking (s^4U sites). These protein fragments cause the reverse transcriptase enzyme to incorporate detectable T-to-C mutations during cDNA synthesis. Often the RNA-binding site location data is compared to known locations of the corresponding RNA modification, and the degree of overlap between the two data sets is assessed. A high degree of overlap (typically >10%) has served as convincing evidence of the specific interaction of the protein and RNA modification in vivo (Liu et al. 2014; Wang et al. 2014; Dai et al. 2018; Huang et al. 2018).

Due to the transcriptome-wide mapping capability of RNA-binding sites, CLIP techniques open doors for many additional investigations. For instance, bioinformatic enrichment analysis of the sequencing reads has shown the ability to uncover preferential binding motifs for the protein of study (Liu et al. 2014; Wang et al. 2014, 2015; Edupuganti et al. 2017; Huang et al. 2018; Zheng et al. 2020). Additionally, RNA-binding site location data has been used to unveil preferences for the protein binding within individual RNA transcripts (Liu et al. 2014; Wang et al. 2014, 2015; Edupuganti et al. 2017; Huang et al. 2018; Zheng et al. 2020). Such analysis was used in the study of the m^6A reader protein, YTHDF2, which led to the finding of preferential targeting of mRNAs near the stop codon, 3' UTR, and CDS regions (Wang et al. 2014). This result guided researchers to hypothesize YTHDF2's involvement in regulating target mRNA transcript stability and/or translation (Wang et al. 2014). Additionally, investigating the downstream cellular processes of targeted transcripts, through *g*ene *o*ntology (GO) term enrichment, has provided insight into the potential impact of the studied protein on overall cellular homeostasis (Edupuganti et al. 2017). GO term enrichment analysis of the mRNA targets for common m^6A reader proteins identified that mRNA metabolism and processing were among the most common impacted pathways (Edupuganti et al. 2017).

In a more low-throughput manner, CLIP techniques can serve to provide a basic litmus test for the binding activity of an epitranscriptome protein. Instead of submitting immunoprecipitated RNA fragments for high-throughput cDNA sequencing analysis, captured RNAs can be labeled radioactively (Edupuganti et al. 2017; Warda et al. 2017) or with biotin (Zhao et al. 2014; Yang et al. 2017; Wu et al. 2019), and the relative quantity of captured RNA can be detected through various visualization mechanisms (autoradiography, chemiluminescence, etc.). This method is often used to study the global impact on the RNA-binding activity of a specific epitranscriptome protein in response to the knockdown/overexpression of other proteins which interface with the same RNA modification (Zhao et al. 2014; Yang et al. 2017; Wu et al. 2019). For example, the total amount of RNA bound to the m^5C reader protein, ALYREF, was shown to be significantly decreased in response to knockdown of the m^5C writer, NSUN2 (Yang et al. 2017).

An in vitro alternative to the aforementioned CLIP techniques for analysis of epitranscriptome protein RNA-binding sites employs screening of diverse libraries of short RNA strands (ssRNAs), which possess an RNA modification of interest, against cognate reader proteins immobilized onto magnetic beads (Arguello et al. 2019). After screening, RNA sequences recovered from binding to the reader protein are similarly prepared for cDNA library sequencing. These sequencing results can be used to decipher enriched reader protein RNA-binding motifs. Such analysis was validated through using the m^6A reader, YTHDC1; the well-conserved DR(m^6A)CH RNA-binding motif was significantly enriched after performing the ssRNA screening procedure against this protein (Arguello et al. 2019). While this method lacks the ability to interrogate in vivo RNA-binding interactions, it may prove valuable for uncovering previously unknown RNA-binding motifs that are not recovered from CLIP analysis, possibly arising from the availability bias of certain RNA sequences in vivo.

2.3.2 Functional Characterization of the Impact of Epitranscriptome Proteins on Target mRNA Sequences

Investigations into the role that a particular epitranscriptome protein has in the translation efficiency, stability, and export of targeted RNAs have been made possible, in large part, due to the RNA-binding site location data obtained from the sequencing-based methodologies mentioned previously (e.g., PAR-CLIP).

The impact that epitranscriptome proteins may have on translation efficiency can be studied in both small and large scales. Small-scale translation efficiency studies have often relied on the use of a reporter protein (i.e., luciferase) inserted downstream of an mRNA sequence known to be targeted by the epitranscriptome protein. Knockdown and/or overexpression of the epitranscriptome protein is performed to assess the impact on translation of the target transcript, indirectly measured by changes in the activity level of luciferase (luminescence signal) (Wu et al. 2019; Song et al. 2019). An increase in the relative luciferase activity post-knockdown of the epitranscriptome protein suggests that the protein works to repress translation of its target mRNAs; this result was shown for HNRNPD, an m^6A reader, on *Tfeb*, one of its confirmed mRNA target sequences (Song et al. 2019). Additionally, CRISPR-dCas and CRISPR-dCas-inspired epitranscriptome fusion protein architectures have been used in conjunction with luciferase translation reporter assays to investigate the impact of specific proteins on the translation of luciferase mRNA (Rauch et al. 2018, 2019). Delivery of the YTHDF1 m^6A reader to a luciferase mRNA reporter via gRNA targeting has been shown to result in relative increases of luciferase activity (Rauch et al. 2019). This result suggests that YTHDF1 is responsible for increasing the translation efficiency of its target mRNAs.

Other methods, such as ribosomal profiling, have been used for analyzing the impact of epitranscriptome proteins on target mRNA translation in a global context. In this strategy, global mRNA translation is first inhibited through the addition of cycloheximide. mRNA immediately isolated from ribosomal fractions is then

submitted for RNA-seq to gain a snapshot of the mRNAs actively being transcribed at this specific time point. The resultant sequencing data is compared between samples obtained from epitranscriptome protein knockdown/overexpression cell lines and from control samples to determine the change in the landscape of ribosome-bound mRNA target sequences. A decrease in the amount of known mRNA-binding targets recovered from ribosomal fractions post-knockdown of a given epitranscriptome protein indicates an involvement of the protein in increasing translation efficiency of its target mRNAs; this result was used to showcase the role of YTHDF1 in increasing translational efficiency (Wang et al. 2015; Dai et al. 2018).

Alternatively, a pulsed-SILAC approach has been employed to study the impact of epitranscriptome proteins on target mRNA translation rates in a global context. Specifically, this isotopic labeling approach has been used to investigate the impact of the m^6A reader, FMR1, on translation of its known mRNA targets. FMR1 overexpression and control cell cultures were transferred to different variations of isotopically labeled growth media, and protein samples were harvested across several time points post-media transfer. Mass spectrometry analysis was subsequently used to compare the turnover rate of the proteins encoded from mRNA targets—indicated by a shift in the mass spectrometry peaks from a normal to an isotopically unique amino acid signature. The average rate of protein turnover was faster in control samples than in FMR1 overexpression samples, which suggested FMR1's role as an inhibitor to translation of its mRNA targets (Edupuganti et al. 2017).

A different mechanism for investigating the impact of epitranscriptome proteins on RNA fate is through mRNA lifetime profiling. In this approach, cell cultures are treated with the potent transcription inhibitor actinomycin D to halt transcriptional production of nascent RNAs. mRNA samples are harvested over several time points (typically spanning ~6 hours) post-actinomycin D treatment and submitted to RNA-seq for transcriptome-wide analysis of mRNA half-life (Wang et al. 2015; Edupuganti et al. 2017; Huang et al. 2018). Monitoring the abundance of specific mRNAs over time permits the calculation of average mRNA half-life (Wang et al. 2015; Edupuganti et al. 2017; Huang et al. 2018; Wu et al. 2019; Liu et al. 2019). Comparing the average half-life of known mRNA targets between control and protein overexpression/knockdown cell cultures provides insight into the impact of the epitranscriptome protein on RNA stability. Such analyses have led to the discovery of the role of the YTHDF2 m^6A reader in driving its mRNA targets toward degradation (Wang et al. 2014, 2015) and of the role of the IGF2BP3 m^6A reader in stabilization of its target mRNAs (Huang et al. 2018). An alternative, low-throughput method has been used to examine RNA half-lives wherein RNA samples harvested post-actinomycin D treatment are subjected to RT-qPCR for the temporal monitoring of specific mRNA transcript abundances (Edupuganti et al. 2017; Wu et al. 2019; Liu et al. 2019; Zheng et al. 2020).

Furthermore, *f*luorescence *i*n *s*itu *h*ybridization (FISH) has been an effective method for analyzing the role of epitranscriptome proteins on mRNA localization and nuclear export. FISH utilizes fluorescently labeled RNA probes possessing base-

pairing complementarity to facilitate annealing to specific target mRNAs (Wang et al. 2014; Yang et al. 2017) or general poly(A) RNAs (Zheng et al. 2013). Upon knockdown and/or overexpression of the epitranscriptome protein of interest, the change in the level of mRNA nuclear export can be quantified via fluorescence microscopy. For example, FISH targeting poly(A) mRNAs in ALKBH5 knockdown HeLa cell cultures resulted in increased cytoplasmic accumulation, suggesting this protein's role in nuclear retention of mRNAs (Zheng et al. 2013). For a more comprehensive analysis, FISH may be coupled with immunostaining of epitranscriptome proteins of interest to observe the colocalization of these proteins with their specific mRNA targets (Yang et al. 2017).

Combining methods that assess the impact of epitranscriptome proteins on target mRNA translation efficiency, stability, and localization have provided a more global understanding of the intracellular role of these proteins, as demonstrated for the well-characterized YTHDF1 and YTHDF2 m^6A reader proteins (Wang et al. 2014).

2.3.3 Characterization of Modified RNA-Protein Binding Kinetics and Catalysis

In addition to the techniques outlined above, EMSA and *m*icro*s*cale *t*hermophoresis (MST) measurements have been widely used to assess in vitro binding interactions between epitranscriptome proteins and specific regions within modified and unmodified RNA substrates. In doing so, these techniques have provided direct evidence for the binding preference of readers toward many common RNA modifications, including m^6A (Wang et al. 2014; Huang et al. 2018; Arguello et al. 2019), m^5C (Yang et al. 2017; Dai et al. 2020), m^1A (Dai et al. 2018; Seo and Kleiner 2020), and 8-oxoG (Hayakawa et al. 2002; Gonzalez-Rivera et al. 2020c). Importantly, these methods are incapable of truly capturing the kinetic activity of writer and eraser proteins, due to the catalytic nature of their function. Instead, studies involving the RNA modification/removal characteristics of writers and erasers have relied on HPLC, mass spectrometry, or radioactivity-based in vitro assays. By taking the HPLC/mass spectrometry spectra of digested nucleotides recovered before and after incubation of the protein with modified or unmodified synthetic RNA oligomers, the relative change in the level of the peak corresponding to the RNA modification of interest is used as an indicator of protein activity (Aas et al. 2003; Jia et al. 2011; Zheng et al. 2013; Fu et al. 2013; Ma et al. 2019; Rau et al. 2019). Alternatively, incubating proteins with radiolabeled RNA oligomers followed by RNA precipitation and scintillation counting of the supernatant has also proven to be effective for characterizing eraser protein functionality (Aas et al. 2003; Ougland et al. 2004). Through meticulous monitoring of time and substrate concentrations, this radiolabeling strategy was shown to achieve catalytic parameter quantification for both writers and erasers of m^6A (Li et al. 2016a). Techniques for quantification of epitranscriptome protein catalytic properties have recently shown high-throughput screening potential (Shen et al. 2019; Wiedmer et al. 2019). Such a screening potential has profound implications in medicine, as small molecule inhibitors of

these proteins are desirable for the treatment of diseases arising from aberrant RNA modification abundance.

3 Advances on Computational Approaches for Accelerated Study of Modified RNA-Protein Interactions

The wet-lab methods described above provide valuable information about interactions of modified RNA-protein complexes. However, these experimental methods may incur high-cost and laborious experiments and may exhibit practical limitations related to the detection limit, reactivity, and scalability (Lin and Miles 2019). To this end, computational methods may offer sustainable avenues to simultaneously screen a large number of candidates or conditions to both guide and enrich wet-lab experimentation. Despite recent advances in experimental strategies to investigate modified RNA-dependent protein interactions, we have recognized a lag of momentum to construct computational strategies that can guide and/or predict potential binding partners in modified RNA-protein complexes. Herein, we identify predictive strategies that mainly fall into two categories (Fig. 7a). The first strategy is based on the known or predicted physicochemical, structural, and functional properties of RNA-protein complexes by applying artificial intelligence algorithms. The second strategy is based on experimentally determined RNA-protein structures and molecular dynamic simulations. This section aims to showcase the current state of computational strategies that characterize RNA-protein interactions in the context of RNA modifications to stimulate innovation in this subfield within epitranscriptomics.

3.1 Artificial Intelligence Progress of Prediction of RNA-Protein Interactions

Artificial intelligence strategies have been widely applied in identifying RNA-binding sites and predicting RNA-protein interaction pairs. Conventional machine learning approaches primarily include supervised models such as *s*upport *v*ector *m*achine (SVM), *h*idden *M*arkov *m*odel (HMM), and *r*andom *f*orest (RF) that have been widely used to predict RNA-binding residues of specific proteins—an area extensively reviewed in Pan et al. (2019) and Sagar and Xue (2019). These conventional machine learning methods identify patterns from existing databases, largely built from single-base resolution techniques (i.e., CLIP-seq data sets or analogous techniques), and, based on these trained models, infer hidden patterns to predict novel binding sites (Fig. 7b). Recent applications of conventional learning methods include a stacking assemble learning framework called RPI-SE for effectively predicting *n*on*c*oding *RNA*s (ncRNA) interactions with proteins that generally

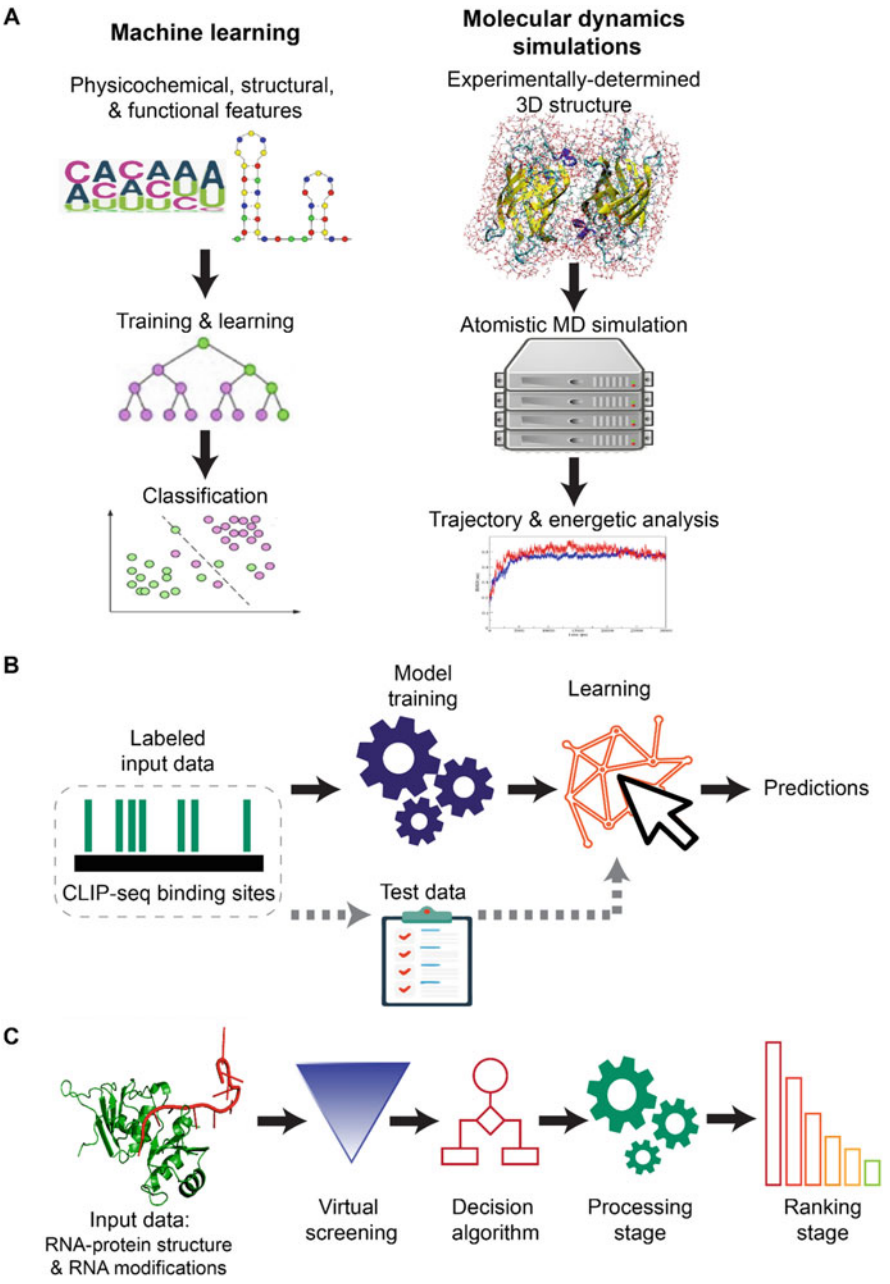


Fig. 7 (**a**) Artificial intelligence approaches and molecular dynamic simulations are two major computational strategies for predicting RNA-protein interactions. (**b**) General schematic of supervised learning algorithms. (**c**) Framework for virtual screening of RNA modifications applied to discovery of protein readers described in Orr et al. (2018)

lack large-scale levels of annotation (thereby preventing the utilization of more advanced learning algorithms). For instance, by integrating three individual learning models (*g*radient *b*oosting *d*ecision *t*ree (GBDT), extremely randomized trees (ExtraTree), and SVM), a study has shown high accuracy prediction of ncRNA-protein interactions entirely based on sequence information compared to previous methods (Yi et al. 2020). However, many predictive approaches combine sequence information with functional and/or structural properties to improve the estimation of the interactions. For example, a recent study presented a tool called *s*equence-*s*tructure *h*idden *M*arkov *m*odel (ssHMM) that considers the structural context of RNAs (Heller et al. 2017) to learn both sequence and structure preferences.

Other widely applied features that are integrated into learning models to enhance accuracy of RNA-protein prediction models include hydrogen bonding, van der Waals' contributions, mRNA half-life, isoelectric point, protein localization, GO annotations, codon bias, and solvent accessibility (Pancaldi and Bähler 2011). Importantly, the performance of machine learning models is greatly dependent on the selection and efficient extraction of these features (Bellucci et al. 2011; Pancaldi and Bähler 2011). As such, methods adapting annotated features may involve highly time-consuming classifications or may be limited by lack of experimental information (Pan et al. 2019). To aid with this critical task, toolboxes to quickly extract features that enable predictions are becoming increasingly common (Torkamanian-Afshar et al. 2020). Likewise, more advanced methods employ deep learning approaches to automatically learn hidden patterns from more complicated features. For example, Deep RBP binding preference (DeepRiPe) (Ghanbari and Ohler 2020) simultaneously extracts features from the different types of RNA sequences (i.e., 3' UTR, CDS, 5′, or UTR) for several RNA-protein binding complexes that might share similar modes of binding. In summary, these learning methods offer large-scale predictions of RNA partners (within RNA-protein binding complexes), but it is currently unclear how available experimental data of interacting modified RNA-protein complexes is sufficient for precisely building conventional learning or deep learning models of protein readers, erasers, or writers within these complexes.

3.2 *Current Applications of Artificial Intelligence in Epitranscriptomics*

A common application of learning methods in epitranscriptomics is the prediction of modified sites (e.g., m^6A, m^5C, etc.) to boost the genome-wide profiling of modified nucleotides, ultimately enriching the ability to annotate, visualize, and interpret the functional roles of RNA modifications. Prediction models based on conventional machine learning and deep learning show reliable performance on identification of many types of RNA modifications, including m^6A (Chen et al. 2015, 2018, 2019; Zou et al. 2019), m^5C (Feng et al. 2016; Li et al. 2018; Zhang et al. 2018), m^1A

(Feng et al. 2017; Li et al. 2018), and Ψ (Panwar and Raghava 2014; Li et al. 2015; Liu et al. 2020b). Application of these methods in epitranscriptomics has been recently reviewed (Liu et al. 2020a). The physicochemical properties employed in these model include (i) sequence features such as functional groups, ring structures, and hydrogen bonds of the nucleotides (Chen et al. 2017), (ii) genomic features such as transcript regions, (iii) the relative position of the modified site (i.e., nucleotide distances toward the splicing junction) (Chen et al. 2019), and (iv) RNA secondary structure properties such as enthalpy, entropy, and free energy (Chen et al. 2015). As previously discussed, a major challenge is the designation of the features given that this requires knowledge expertise to extract correct feature descriptors; this challenge has been addressed by employing deep learning models (Wei et al. 2019). Computational models used in these previously described studies have already been reviewed (Chen et al. 2017).

In the context of applications of learning algorithms to predict modified RNA-protein interactions, a recent study compared various traditional machine learning models (e.g., logistic regression (LR), SVM, and RF) to distinguish the RNA substrates of six m^6A readers, the five members of the YTH family protein (YTHDC1–2, YTHDF1–3) and Eukaryotic Translation Initiation Factor 3 Subunit A (EIF3A) (Zhen et al. 2020). Taking advantage of reported cross-linking immunoprecipitation information (i.e., PAR-CLIP and iCLIP data sets) and RNA methylation profiling (miCLIP, m^6A-CLIP, MAZTER-seq, m^6A-REF-seq, and PA-m^6A-seq data sets) to train the model, this study provided effective computational prediction for the six m^6A readers using 58 genomic and sequence features of the RNA targets. Most importantly, this computational analysis unveiled clues on target specificity and biological functions of the different m^6A readers. However, in the absence of immunoprecipitation and sequencing profiling data, it is of value to develop predictive models based on the structurally characterized RNA-protein complexes available in the RCSB Protein Data Bank (PDB) (Burley et al. 2019).

3.3 Advances in Molecular Dynamics (MD) Simulations in Studying Modified RNA-Protein Interactions

Interactions between modified RNAs and RBPs can also be predicted from experimentally determined RNA-protein structures and atomistic molecular dynamics simulations (Tuszynska et al. 2014). Biophysical models of macromolecular structures have matured to provide mechanistic description of RNA-protein interactions (Vuković et al. 2016). For that reason, MD simulations are an alternative to gain both atomic-level insights into the basic biochemical principles of protein recognition of modified RNA bases as well as to design novel molecules in silico (Frohlich et al. 2016). For example, in a series of publications, the molecular binding mechanism of a short, 15 amino acid peptide to the anticodon stem loop of hypermodified $tRNA^{Lys3}$, which is associated with viral RNA replication (Li et al. 1996), was

explored as well as its computational design. By performing atomistic simulations and energetic calculations of the binding energy, these studies identified how key hydrophobic residues in the peptide recognize the anticodon stem loop of $tRNA^{Lys3}$ due to specific chemical interactions, including van der Waals energies that contribute to specificity (Xiao et al. 2014), while electrostatic and polar solvation energies disfavored the binding (Xiao et al. 2015). Moreover, the simulations described that recognition is mediated by direct interactions with the modified bases of 5-methoxycarbonylmethyl-2-thiouridine at position 34 ($mcm^5s^2U_{34}$) and 2-methylthio-N6-threonylcarbamoyladenosine at position 37 ($ms^2t^6A_{37}$) (Xiao et al. 2016). Based on these principles, peptides that resulted in a tenfold increase in the binding affinity to the starting sequence were further designed in silico (Spears et al. 2014).

Recently, our laboratory has developed a virtual screening framework to predict interactions between modified RNAs and any RNA-binding protein (RBP), regardless of their size, based on MD simulations (Fig. 7c) (Orr et al. 2018). The framework provides a sustainable and flexible approach to test hypothesis-driven interactions of a specific RBP with a defined sequence that can contain one or more modified nucleotides from a large library of more than 100 possible types of RNA modifications previously identified (Boccaletto et al. 2018). To study all the potential interactions efficiently, we implemented the screening system based on tree structures. The four canonical ribonucleotides (e.g., A, C, G, and U) define the seeds for four hierarchical trees that integrate the structural and physicochemical information of the corresponding chemical modifications propagating from each canonical molecule. The process begins by a rapid virtual screening stage that calculates the interaction free energy from 5 ns implicit-solvent MD simulations and the energy of the residues 10 Å of any atom of the RNA in a hierarchical way. Two energetic features of the interaction are evaluated from these calculations: (i) the average interaction energy between the RNA-protein complex and (ii) the total Gibbs free energy of the modified nucleoside. Then, a decision algorithm filters out the modifications with unfavorable interaction energy compared to the predecessor modification and/or the canonical ribonucleotide to prohibit further search of additional modifications within the same hierarchical branch. This strategy efficiently decreases the number of screenings by assuming that more complex modifications stemming from already unfavorable modifications of lower complexity will also be energetically unfavorable. Subsequently, the modifications passing the decision algorithm enter a processing stage where 30 ns explicit-solvent MD simulations are conducted to determine association free energies. Finally, the resulting association free energies are ranked to establish their favorability to interact with the studied protein. This model was trained with 9-nucleotide long RNA fragments containing one modification (out of 46 total modifications selected) at two different positions of the RNA strand and was tested in the context of polynucleotide phosphorylase (PNPase) protein. PNPase forms a homotrimer structure, where each monomer subunit is composed of 711 amino acids. Importantly, when tested experimentally, results from these simulations produced highly accurate predictions of the interaction affinities between modified RNAs and PNPase, with a regression correlation

coefficient (R^2) of 0.944, demonstrating a remarkable prediction ability from a computationally demanding modified RNA-protein model.

One application of this framework is the accurate prediction of PNPase variants with stronger binding affinity for chemically oxidized RNA containing 8-oxoG RNA modifications (Gonzalez-Rivera et al. 2020c). In this study, the virtual screening framework was modified to interrogate PNPase protein variants that exhibited higher specificity of a defined chemical modification, relative to wild-type PNPase. 126 rationally selected mutations in the RNA-binding site were screened using the same processive steps of a rapid virtual screening, followed by a decision stage, and lastly a processing and ranking stage. It is worth noting that mutations were selected based on the bioinformatic analysis of 782 homolog PNPase sequences from different organisms. Importantly, this analysis resulted in five variants that showed strong discrimination of 8-oxoG in silico. Experimental characterization demonstrated a fold change increase in 8-oxoG binding affinity between 1.2 and 1.5 and/or selectivity between 1.5 and 1.9 in the PNPase variants. Furthermore, the improvement in 8-oxoG binding was accompanied by increased cellular tolerance to oxidative stress when the variants were endogenously expressed in K12 Δpnp *E. coli* cells, demonstrating the biological relevance of the in silico designed mutants. Collective results from this study showed important flexibility of the virtual screening framework to identify RNA-binding proteins and their interactions in a high-throughput manner. Ultimately this approach leads to valuable insights into biochemical mechanisms of RNA-protein interactions as well as potential applications for engineering these interfaces.

Overall, computational methods can readily assist to functionally interrogate protein readers, writers, and erasers by examining key functional regions (i.e., binding and catalytic sites) and by examining their capacity to specifically interact with modified RNAs. Recent studies speculate that proteins involved in the recognition of the epitranscriptome have different degrees of specificity toward multiple RNA modifications (Lao and Barron 2019); the opportunity to conveniently design and chemically alter any RNA ligand computationally provides directions to interrogate this unexplored aspect of the biochemistry of RNA-protein interactions. Furthermore, computational methods can prompt the investigation of lowly abundant chemical modifications (that are difficult to probe in vivo) to warrant further examination in their cellular environment.

4 Future Applications in the Field of Epitranscriptome Protein Discovery and Characterization

As discussed above, a variety of experimental and computational techniques have contributed toward identification and further characterization of the intracellular roles of epitranscriptome proteins. Experimental and computational advances in the identification and characterization of epitranscriptome proteins and of their

interactions with modified RNAs have introduced a variety of tools to target RNA modifications as potential disease diagnostics and therapeutics (Sherman and Contreras 2018). However, it is clear that greater implications of RNA modifications, particularly as it relates to therapeutic targets, will continue to emerge with advances in this field. It is clear that epitranscriptome editing for therapeutic intervention will be most effective with advanced biological tools that can interact with RNA modifications in a targeted, transcript-specific manner.

4.1 Targeted Delivery of Epitranscriptome Regulation to Combat RNA Modification-Related Diseases

Recently, tools developed for studying modified RNA-protein interactions have led to a new approach in precision medicine—one which uses epitranscriptome fusion protein architectures for targeting RNAs harboring disease-relevant modifications. For instance, RNA methylations have been observed to be upregulated in cases of solid tumor pathogenesis, lung cancers, and acute myeloid leukemia cells (Wu et al. 2018). Specifically, m^6A has been found to be present on oncogene mRNA and can therefore act to promote or inhibit tumor development. Failures along the m^6A regulatory network can induce additional health-related complications by interfering with normal development and cellular function (Zhou and Yang 2020).

Ongoing studies are adapting novel RNA targeting tools to focus on epitranscriptome editing for therapeutic applications. Advancements in CRISPR-Cas-based technologies have led to the development of several novel techniques for targeting RNA regulatory processes and gene expression. Novel fusion architectures which link CRISPR-dCas systems with epitranscriptome protein effectors are capable of interfacing with RNA modifications in a diverse manner (Liu et al. 2019). Sequence-specific installation or depletion of m^6A by means of CRISPR-dCas9-epitranscriptome protein fusion systems provides a versatile toolkit for the targeting of m^6A-associated cellular processes involved in chronic diseases such as obesity and cancer (Rau et al. 2019; Liu et al. 2019). A promising CRISPR-dCas13-directed *R*NA *e*diting for *p*rogrammable *A* to *I* *r*eplacement (REPAIR) system for genetic disease treatment and targeted epitranscriptome modulation has also been developed. An advantage of these systems is that RNA editing is transient in nature, which allows for temporal control and reversible editing outcomes. Therefore, REPAIR-mediated edits could treat temporal intracellular fluctuations (e.g., inflammation) implicated in disease (Cox et al. 2017).

The development of alternatives to traditional CRISPR-Cas-based technologies is also of interest given that bacterial-based CRISPR-Cas systems may induce unwanted immune responses in host cells for therapeutic applications. CIRTS, a CRISPR-inspired tool built from human protein-derived parts, has been demonstrated to effectively direct epitranscriptome effector proteins to specific RNA sequences in vivo. The nature of such a tool, being derived from human protein

origin as opposed to bacterial (CRISPR-Cas), holds promise as an epitranscriptome modulation therapy that subverts any immunological consequences linked to introducing foreign elements into human cells. To aid application of this technology, viral delivery was successfully used to deliver the CIRTS targeting machinery in HEK293T cells (Rauch et al. 2019). Viral delivery methods, such as the *a*denovirus-*a*ssociated *v*irus (AAV) delivery vehicle, are known to have a low risk of genome insertion and elicit low immune response stimulation from host cells (Vasileva and Jessberger 2005). The latter result provides a potential pathway for the development of clinically relevant viral delivery of CIRTS to regulate the epitranscriptome at the transcript-specific level (Rauch et al. 2019). Developments in site-specific epitranscriptome regulation tools offer a promising outlook for future therapeutic intervention strategies.

4.2 High-Throughput Screening Strategies to Uncover Small Molecule Inhibitors of Epitranscriptome Proteins

Another exciting recent discovery has been that RNA modification-dependent protein-RNA interaction sites are druggable—the activity of these sites can be both inhibited and enhanced by small molecules that interact with high binding affinity or ligand efficiency (Buker et al. 2020). Crystallographic analysis has been implemented as a method for exploration of these small molecule ligands to the YTHDC1 m^6A reader domain. In this study, 30 small protein fragments were analyzed for their affinity to the m^6A binding pocket of YTHDC1; MD simulations were used to model the binding of these ligands to the protein binding pocket. Evidence from this study suggests that small molecules can disrupt protein-RNA interactions in an m^6A-specific manner (Bedi et al. 2020).

New research tools can assist in the development and discovery of epitranscriptome proteins as novel therapeutic targets. *S*elf-*a*ssembled *m*onolayer *d*esorption *i*onization (SAMDI) is a new high-throughput technique which has been adapted for screening targets of epitranscriptome writers such as METTL3/METTL14, demonstrating its ability to characterize inhibitors to these proteins and their associated IC_{50} values (Buker et al. 2020). Additionally, antibody-based assays have been developed to further investigate and identify novel inhibitors of RNA demethylases, as was shown for the m^6A eraser, ALKBH5 (Shen et al. 2019). Moreover, *h*omogeneous *t*ime-*r*esolved *f*luorescence (HTRF) was recently developed as an alternative epitranscriptome protein inhibitor screening strategy. In this method, a *f*luorescence *r*esonance *e*nergy *t*ransfer (FRET)-labeled antibody is bound to a purification tag on an established epitranscriptome reader protein, and, upon interaction with a methylated RNA carrying a complementary FRET acceptor, the two fluorophores produce a detectable emission spectra output (Wiedmer et al. 2019). Finding small molecule inhibitors to these epitranscriptome proteins is an

important research thrust moving forward as a mechanism for combating abnormal RNA modification abundance in vivo.

5 Conclusions

The identification and characterization of epitranscriptome proteins, both experimentally and computationally, has generated an increased understanding of how these proteins regulate the RNA modification landscape. The utilization of unique pulldown strategies has been successful in identifying several readers/writers/erasers of the epitranscriptome. Subsequent characterization techniques have led to a better understanding of how these proteins are involved in broader cellular processes. Computational efforts to characterize epitranscriptome proteins have also presented themselves to be cost- and time-effective alternatives and complements to traditional wet-lab techniques. Machine learning and algorithmic models have enabled the prediction of RBP binding locations and RNA modification sites. Additionally, molecular dynamic simulations have provided a flexible virtual platform amenable for high-throughput screening of modified RNA-protein binding interactions. All of the aforementioned methodologies have helped to illuminate the broad involvements of epitranscriptome proteins on the landscape of RNA modifications and provide a platform for future discoveries, including the development of epitranscriptome protein-based tools designed to tackle abnormal RNA modification abundance.

Acknowledgments This work was supported by the National Institutes of Health under award number R21ES032124. J.C.G-R was supported by the Administrative Department of Science, Technology, and Innovation (COLCIENCIAS) of Colombia and Fulbright (Grant 479). The authors would like to thank all members of the Contreras lab at the University of Texas at Austin for discussions, suggestions, and support. The authors would especially like to thank lab members Sean Engels, Lucas Miller, and Mark Sherman for their significant contributions toward editing this work. The authors would like to acknowledge BioRender.com for assistance with figure creation.

References

Aas PA, Otterlei M, Falnes P et al (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature 421:859–863

Anreiter I, Mir Q, Simpson JT et al (2020) New twists in detecting mRNA modification dynamics. Trends Biotechnol S0167-7799(20):30166–30169

Arguello AE, Deliberto AN, Kleiner RE (2017) RNA chemical proteomics reveals the N^6-Methyladenosine (m^6A)-regulated protein-RNA Interactome. J Am Chem Soc 139:33

Arguello AE, Leach RW, Kleiner RE (2019) In vitro selection with a site-specifically modified RNA library reveals the binding preferences of N^6-Methyladenosine reader proteins. Biochemistry 58:3386–3395

Baldridge KC, Zavala J, Surratt J et al (2015) Cellular RNA is chemically modified by exposure to air pollution mixtures. Inhal Toxicol 27:74–82

Bedi RK, Huang D, Wiedmer L et al (2020) Selectively disrupting m^{6}A-dependent protein-RNA interactions with fragments. ACS Chem Biol 15:618–625

Bellucci M, Agostini F, Masin M et al (2011) Predicting protein associations with long noncoding RNAs. Nat Methods 8:444–445

Boccaletto P, MacHnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307

Buker SM, Gurard-Levin ZA, Wheeler BD et al (2020) A mass spectrometric assay of METTL3/METTL14 methyltransferase activity. SLAS Discovery 25:361–371

Burley SK, Berman HM, Bhikadiya C et al (2019) RCSB protein data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. Nucleic Acids Res 47:464–474

Chen W, Ding H, Zhou X et al (2018) iRNA(m^{6}A)-PseDNC: identifying N^6-methyladenosine sites using pseudo dinucleotide composition. Anal Biochem 562:59–65

Chen W, Feng P, Ding H et al (2015) iRNA-methyl: identifying N^6-methyladenosine sites using pseudo nucleotide composition. Anal Biochem 490:26–33

Chen X, Sun YZ, Liu H et al (2017) RNA methylation and diseases: experimental results, databases, web servers and computational models. Brief Bioinform 20:896–917

Chen K, Wei Z, Zhang Q et al (2019) WHISTLE: a high-accuracy map of the human N^6-methyladenosine (m^{6}A) epitranscriptome predicted using a machine learning approach. Nucleic Acids Res 47:e41

Contreras LM, Gonzalez-Rivera JC, Baldridge KC et al (2020) Understanding the functional impact of VOC-ozone mixtures on the chemistry of RNA in epithelial lung cells. Res Rep Health Eff Inst 201:3–43

Cox DBT, Gootenberg JS, Abudayyeh OO et al (2017) RNA editing with CRISPR-Cas13. Science 358:1019–1027

Dai X, Gonzalez G, Li L et al (2020) YTHDF2 binds to 5-Methylcytosine in RNA and modulates the maturation of ribosomal RNA. Anal Chem 92:1346–1354

Dai X, Wang T, Gonzales G et al (2018) Identification of YTH domain-containing proteins as the readers for N1-Methyladenosine in RNA. Anal Chem 90:6380–6384

David R, Burgess A, Parker B et al (2017) Transcriptome-wide mapping of RNA 5-methylcytosine in arabidopsis mRNAs and noncoding RNAs. Plant Cell 29:445–460

Dominissini D, Moshitch-Moshkovitz S, Schwartz S et al (2012) Topology of the human and mouse m^{6}A RNA methylomes revealed by m^{6}A-seq. Nature 485:201–206

Edupuganti RR, Geiger S, H Lindeboom RG et al (2017) N^6-methyladenosine (m^{6}A) recruits and repels proteins to regulate mRNA homeostasis. Nat Struct Mol Biol 24:870–878

Feng P, Ding H, Chen W et al (2016) Identifying RNA 5-methylcytosine sites via pseudo nucleotide compositions. Mol BioSyst 12:3307–3311

Feng P, Ding H, Yang H et al (2017) iRNA-PseColl: identifying the occurrence sites of different RNA modifications by incorporating collective effects of nucleotides into PseKNC. Mol Ther Nucleic Acids 7:155–163

Frohlich KM, Sarachan KL, Todd GC et al (2016) Post-transcriptional modifications of RNA: impact on RNA function and human health. In: Jurga S, Erdmann (Deceased) V, Barciszewski J (eds) Modified nucleic acids in biology and medicine. RNA technologies. Springer, Cham

Fu Y, Jia G, Pang X et al (2013) FTO-mediated formation of N^6-hydroxymethyladenosine and N^6-formyladenosine in mammalian RNA. Nat Commun 4:1–8

Ghanbari M, Ohler U (2020) Deep neural networks for interpreting RNA-binding protein target preferences. Genome Res 30:214–226

Gilbert WV, Bell TA, Schaening C (2016) Messenger RNA modifications: form, distribution, and function. Science 352:1408–1412

Gonzalez-Rivera JC, Baldridge KC, Wang DS et al (2020a) Post-transcriptional air pollution oxidation to the cholesterol biosynthesis pathway promotes pulmonary stress phenotypes. Commun Biol 3:1–16

Gonzalez-Rivera JC, Orr AA, Engels SM et al (2020c) Computational evolution of an RNA-binding protein towards enhanced oxidized-RNA binding. Comput Struct Biotechnol J 18:137–152

Gonzalez-Rivera JC, Sherman MW, Wang DS et al (2020b) RNA oxidation in chromatin modification and DNA-damage response following exposure to formaldehyde. Sci Rep 10:16545

Hayakawa H, Fujikane A, Ito R et al (2010) Human proteins that specifically bind to 8-oxoguanine-containing RNA and their responses to oxidative stress. Biochem Biophys Res Commun 403:220–224

Hayakawa H, Sekiguchi M (2006) Human polynucleotide phosphorylase protein in response to oxidative stress. Biochemistry 45:6749–6755

Hayakawa H, Uchiumi T, Fukuda T et al (2002) Binding capacity of human YB-1 protein for RNA containing 8-oxoguanine. Biochemistry 41:12739–12744

Heller D, Krestel R, Ohler U et al (2017) SSHMM: extracting intuitive sequence-structure motifs from high-throughput RNA-binding protein data. Nucleic Acids Res 45:11004–11018

Helm M, Motorin Y (2017) Detecting RNA modifications in the epitranscriptome: predict and validate. Nat Rev Genet 18:275–291

Huang H, Weng H, Sun W et al (2018) Recognition of RNA N^6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol 20:285–295

Ishii T, Hayakawa H, Igawa T et al (2018) Specific binding of PCBP1 to heavily oxidized RNA to induce cell death. Proc Natl Acad Sci U S A 115:6715–6720

Jia G, Fu Y, Zhao X et al (2011) N^6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887

Lao N, Barron N (2019) Cross-talk between m^6A and m^1A regulators, YTHDF2 and ALKBH3 fine-tunes mRNA expression. bioRxiv 589747

Li J, Huang Y, Yang X et al (2018) RNAm5Cfinder: a web-server for predicting RNA 5-methylcytosine (m5C) sites based on random Forest. Sci Rep 8:1–5

Li F, Kennedy S, Hajian T et al (2016a) A radioactivity-based assay for screening human m^6A-RNA methyltransferase, METTL3-METTL14 complex, and demethylase ALKBH5. J Biomol Screen 21:290–297

Li Z, Shalom A, Huang Y et al (1996) Multiple forms of tRNA(Lys3) in HIV-1. Biochem Biophys Res Commun 227:530–540

Li X, Xiong X, Yi C (2016b) Epitranscriptome sequencing technologies: decoding RNA modifications. Nat Methods 14:23–31

Li YH, Zhang G, Cui Q (2015) PPUS: a web server to predict PUS-specific pseudouridine sites. Bioinformatics 31:3362–3364

Lin Z, Hsu PJ, Xing X et al (2017) Mettl3−/Mettl14-mediated mRNA N^6-methyladenosine modulates murine spermatogenesis. Cell Res 27:1216–1230

Lin C, Miles WO (2019) Beyond CLIP: advances and opportunities to measure RBP-RNA and RNA-RNA interactions. Nucleic Acids Res 47:5490–5501

Liu K, Chen W, Lin H (2020b) XG-PseU: an eXtreme gradient boosting based method for identifying pseudouridine sites. Mol Gen Genomics 295:13–21

Liu N, Dai Q, Zheng G et al (2015) N^6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature 518:560–564

Liu L, Song B, Ma J et al (2020a) Bioinformatics approaches for deciphering the epitranscriptome: recent progress and emerging topics. Comput Struct Biotechnol J 18:1587–1604

Liu J, Yue Y, Han D et al (2014) A METTL3-METTL14 complex mediates mammalian nuclear RNA N^6-adenosine methylation. Nat Chem Biol 10:93–95

Liu XM, Zhou J, Mao Y et al (2019) Programmable RNA N^6-methyladenosine editing by CRISPR-Cas9 conjugates. Nat Chem Biol 15:865–871

Ma C-J, Ding J-H, Ye T-T et al (2019) AlkB homologue 1 Demethylates N^3-Methylcytidine in mRNA of mammals. ACS Chem Biol 14:1418–1425

Meyer KD, Patil DP, Zhou J et al (2015) 5′ UTR m^6A promotes cap-independent translation. Cell 163:999–1010

Orr AA, Gonzalez-Rivera JC, Wilson M et al (2018) A high-throughput and rapid computational method for screening of RNA post-transcriptional modifications that can be recognized by target proteins. Methods 143:34–47

Ougland R, Zhang CM, Liiv A et al (2004) AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. Mol Cell 16:107–116

Pan X, Yang Y, Xia C et al (2019) Recent methodology progress of deep learning for RNA–protein interaction prediction. Wiley Interdiscip Rev RNA 10:e1544

Pancaldi V, Bähler J (2011) In silico characterization and prediction of global protein-mRNA interactions in yeast. Nucleic Acids Res 39:5826–5836

Panwar B, Raghava GPS (2014) Prediction of uridine modifications in tRNA sequences. BMC Bioinformatics 15:326

Patil DP, Pickering BF, Jaffrey SR (2018) Reading m^6A in the transcriptome: m^6A-binding. Proteins 28:113–127

Perez-Perri JI, Rogell B, Schwarzl T et al (2018) Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. Nat Commun 9:1–13

Rau K, Rösner L, Rentmeister A (2019) Sequence-specific m^6A demethylation in RNA by FTO fused to RCas9. RNA 25:1311–1323

Rauch S, He C, Dickinson BC (2018) Targeted m^6A reader proteins to study Epitranscriptomic regulation of single RNAs. J Am Chem Soc 140:12

Rauch S, He E, Srienc M et al (2019) Programmable RNA-guided RNA effector proteins built from human parts. Cell 178 e12:122–134

Reichle VF, Petrov DP, Weber V et al (2019) NAIL-MS reveals the repair of 2-methylthiocytidine by AlkB in E. coli. Nat Commun 10:1–11

Reichle VF, Weber V, Kellner S (2018) NAIL-MS in E. coli determines the source and fate of methylation in tRNA. Chembiochem 19:2575–2583

Safra M, Nir R, Farouq D et al (2017) TRUB1 is the predominant pseudouridine synthase acting on mammalian mRNA via a predictable and conserved code. Genome Res 27:393–406

Sagar A, Xue B (2019) Recent advances in machine learning based prediction of RNA-protein interactions. Protein Pept Lett 26:601–619

Sajini AA, Choudhury NR, Wagner RE et al (2019) Loss of 5-methylcytosine alters the biogenesis of vault-derived small RNAs to coordinate epidermal differentiation. Nat Commun 10:2550

Seo KW, Kleiner RE (2020) YTHDF2 recognition of N^1-Methyladenosine (m^1A)-modified RNA is associated with transcript destabilization. ACS Chem Biol 15:132–139

Shen DD, Suo FZ, Song QM et al (2019) Development of formaldehyde dehydrogenase-coupled assay and antibody-based assays for ALKBH5 activity evaluation. J Pharm Biomed Anal 162:9–15

Shen Z, Wu W, Hazen SL (2000) Activated leukocytes Oxidatively damage DNA, RNA, and the nucleotide Pool through halide-dependent formation of hydroxyl radical. Biochemistry 39:5474–5482

Sherman M, Contreras L (2018) Computational approaches in design of nucleic acid-based therapeutics. Curr Opin Biotechnol 53:232–239

Shinoda K, Suda A, Otonari K et al (2020) Programmable RNA methylation and demethylation using PUF RNA binding proteins. Chem Commun 56:1365–1368

Shu X, Cao J, Cheng M et al (2020) A metabolic labeling method detects m6A transcriptome-wide at single base resolution. Nat Chem Biol 16:887–895

Sindelar M, Despic V, Mauer J et al (2019) FTO controls reversible m^6Am RNA methylation during snRNA biogenesis. Nat Chem Biol 15:340–347

Song H, Feng X, Zhang H et al (2019) METTL3 and ALKBH5 oppositely regulate m^6A modification of TFEB mRNA, which dictates the fate of hypoxia/reoxygenation-treated cardiomyocytes. Autophagy 15:1419–1437

Spears JL, Xiao X, Hall CK et al (2014) Amino acid signature enables proteins to recognize modified tRNA. Biochemistry 53:1125–1133

Torkamanian-Afshar M, Lanjanian H, Nematzadeh S et al (2020) RPINBASE: an online toolbox to extract features for predicting RNA-protein interactions. Genomics 112:2623–2632

Tuszynska I, Matelska D, Magnus M et al (2014) Computational modeling of protein-RNA complex structures. Methods 65:310–319

Vasileva A, Jessberger R (2005) Precise hit: adeno-associated virus in gene targeting. Nat Rev Microbiol 3:837–847

Vuković L, Chipot C, Makino DL et al (2016) Molecular mechanism of Processive 3′ to 5′ RNA translocation in the active subunit of the RNA exosome complex. J Am Chem Soc 138:4069–4078

Wang X, Lu Z, Gomez A et al (2014) m^6A-dependent regulation of messenger RNA stability. Nature 505:1–20

Wang X, Zhao BS, Roundtree IA et al (2015) N^6-methyladenosine modulates messenger RNA translation efficiency. Cell 161:1388–1399

Warda AS, Kretschmer J, Hackert P et al (2017) Human METTL16 is a N^6-methyladenosine (m^6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep 18:2004–2014

Wei L, Su R, Wang B et al (2019) Integration of deep feature representations and handcrafted features to improve the prediction of N^6-methyladenosine sites. Neurocomputing 324:3–9

Wiedmer L, Eberle SA, Bedi RK et al (2019) A reader-based assay for m^6A writers and erasers. Anal Chem 91:3078–3084

Wilson C, Chen PJ, Miao Z et al (2020) Programmable m^6A modification of cellular RNAs with a Cas13-directed methyltransferase. Nat Biotechnol:1–10

Wu J, Li Z (2008) Human polynucleotide phosphorylase reduces oxidative RNA damage and protects HeLa cell against oxidative stress. Biochem Biophys Res Commun 372:288–292

Wu R, Li A, Sun B et al (2019) A novel m^6A reader Prrc2a controls oligodendroglial specification and myelination. Cell Res 29:23–41

Wu X, Sang L, Gong Y (2018) N6-methyladenine RNA modification and cancers. Am J Cancer Res 8:1957–1966

Xiang JF, Yang Q, Liu CX et al (2018) N^6-Methyladenosines modulate A-to-I RNA editing. Mol Cell 69:126–135.e6

Xiao X, Agris PF, Hall CK (2015) Molecular recognition mechanism of peptide chain bound to the tRNA(Lys3) anticodon loop in silico. J Biomol Struct Dyn 33:14–27

Xiao X, Hall CK, Agris PF (2014) The design of a peptide sequence to inhibit HIV replication: a search algorithm combining Monte Carlo and self-consistent mean field techniques. J Biomol Struct Dyn 32:1523–1536

Xiao X, Zhao B, Agris PF, Hall CK (2016) Simulation study of the ability of a computationally-designed peptide to recognize target $tRNA^{Lys3}$ and other decoy tRNAs. Protein Sci 25:2243–2255

Xu L, Liu X, Sheng N et al (2017) Three distinct 3-methylcytidine (m^3C) methyltransferases modify tRNA and mRNA in mice and humans. J Biol Chem 292:14695–14703

Yang Y, Hsu PJ, Chen YS et al (2018) Dynamic transcriptomic m^6A decoration: writers, erasers, readers and functions in RNA metabolism. Cell Res 28:616–624

Yang X, Yang Y, Sun BF et al (2017) 5-methylcytosine promotes mRNA export-NSUN2 as the methyltransferase and ALYREF as an m^5C reader. Cell Res 27:606–625

Yi HC, You ZH, Wang MN et al (2020) RPI-SE: a stacking ensemble learning framework for ncRNA-protein interactions prediction using sequence information. BMC Bioinformatics 21:60

Zaccara S, Ries RJ, Jaffrey SR (2019) Reading, writing and erasing mRNA methylation. Nat Rev Mol Cell Biol 20:608–624

Zhang M, Xu Y, Li L et al (2018) Accurate RNA 5-methylcytosine site prediction based on heuristic physical-chemical properties reduction and classifier ensemble. Anal Biochem 550:41–48

Zhao YY, Mao MW, Zhang WJ et al (2018) Expanding RNA binding specificity and affinity of engineered PUF domains. Nucleic Acids Res 46:4771–4782

Zhao BS, Roundtree IA, He C (2016) Post-transcriptional gene regulation by mRNA modifications. Nat Rev Mol Cell Biol 18:31–42

Zhao X, Yang Y, Sun BF et al (2014) FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res 24:1403–1419

Zhen D, Wu Y, Zhang Y et al (2020) m^6A reader: Epitranscriptome target prediction and functional characterization of N^6-Methyladenosine (m^6A) readers. Front Cell Dev Biol 8:741

Zheng G, Dahl JA, Niu Y et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29

Zheng Q, Gan H, Yang F et al (2020) Cytoplasmic m^1A reader YTHDF3 inhibits trophoblast invasion by downregulation of m^1A-methylated IGF1R. Cell Discov 6:12

Zhou LL, Yang CG (2020) Targeting Epitranscriptomic proteins for therapeutic intervention. Biochemistry 59:125–127

Zou Q, Xing P, Wei L et al (2019) Gene2vec: gene subsequence embedding for prediction of mammalian N^6-methyladenosine sites from mRNA. RNA 25:205–218