The Emerging Epitranscriptomics of Long Noncoding RNAs

Andrew Shafik[a,*]

Ulrike Schumann[a,*]

Maurits Evers[a]

Tennille Sibbritt[a,c]

Thomas Preiss[a,b]

a: EMBL–Australia Collaborating Group, Department of Genome Sciences, The John Curtin School of Medical Research (JCSMR), The Australian National University, Building 131 Garran Road, Acton (Canberra), Australian Capital Territory 2601, Australia

b: Victor Chang Cardiac Research Institute, 405 Liverpool Street, Darlinghurst (Sydney), New South Wales 2010, Australia

c: Present address: Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst (Sydney), New South Wales 2010, Australia

***:** Equal contribution

Email addresses:

AS: andrew.shafik@anu.edu.au

US: ulrike.schumann@anu.edu.au

ME: maurits.evers@anu.edu.au

TE: t.sibbritt@garvan.org.au

Corresponding author: TP (email: thomas.preiss@anu.edu.au; twitter: @tominaustralia)

Highlights

- RNA modifications are chemically diverse and implement epitranscriptomic regulation
- Technological advances in next generation sequencing have propelled the transcriptome-wide mapping of RNA modifications
- Emerging information on the role of post-transcriptional modifications in long noncoding RNA function is reviewed here

ABSTRACT

The pervasive transcription of genomes into long noncoding RNAs has been amply demonstrated in recent years and garnered much attention. Similarly, emerging 'epitranscriptomics' research has shown that chemically modified nucleosides, thought to be largely the domain of tRNAs and other infrastructural RNAs, are far more widespread and can exert unexpected influence on RNA utilization. Both areas are characterized by the often-ephemeral nature of the subject matter in that few individual examples have been fully assessed for their molecular or cellular function, and effects might often be subtle and cumulative. Here we review available information at the intersection of these two exciting areas of biology, by focusing on four RNA modifications that have been mapped transcriptome-wide: 5-methylcytidine, N6-methyladenosine, pseudouridine as well as adenosine to inosine (A-to-I) editing, and their incidence and function in long noncoding RNAs. This article is part of a Special Issue entitled: Long Noncoding RNA.

Keywords

lncRNA, 5-methylcytidine, N6-methyladenosine, pseudouridine, editing, miRNAs

Abbreviations

m^5C, 5-methylcytidine; hm^5C, 5-hydroxymethylcytidine; m^6A, N6-methyladenosine; Ψ, pseudouridine; NGS, next generation sequencing; bsRNA-seq, bisulfite RNA sequencing; Aza-IP, 5-Azacytidine RNA immunoprecipitation; iCLIP, individual-nucleotide-resolution crosslinking and immunoprecipitation; miCLIP, methylation individual-nucleotide-resolution crosslinking and immunoprecipitation

1. Introduction

The genomes of complex organisms encode an abundance of different types of noncoding RNAs (ncRNAs), which are expected to constitute an intricate layer of mostly regulatory molecules that operate on different levels of gene expression. As a class, long noncoding RNAs (lncRNAs) are nominally defined as ncRNAs longer than 200 nucleotides (nt), to distinguish them from small regulatory and infrastructural ncRNAs [1]. Of the thousands of lncRNAs that have been identified, few have as yet been assessed for their biological function. For instance, lncRNAdb, a database for functionally annotated lncRNAs currently contains only 287 eukaryotic entries, although the list is steadily growing [2]. As befits all RNAs, lncRNA function is affected by their propensity to adopt a secondary structure and assemble into ribonucleoprotein (RNP) complexes [3]. Both of these aspects will likely be affected by nucleoside modifications, if present in these molecules. Thus, there is an increasing focus on exploring the incidence and biological relevance of such epitranscriptomic marks in lncRNAs.

More than 140 post-transcriptionally modified nucleosides have been reported in RNA [4, 5]. Detection of these modifications in tRNA, rRNA and small nucleolar RNA (snoRNA) has been a major focus of the field. However, recent studies, driven by advances in next generation sequencing (NGS), have further ventured into identifying modifications in less abundant RNA types such as mRNAs and lncRNAs, spawning the new research area of 'epitranscriptomics' [6]. An intense effort is underway to globally map RNA modifications as well as to identify the cellular componentry involved, e.g. the 'writers, readers and erasers' for each modification [7-9]. This review summarizes findings with four modifications that have since been identified and mapped transcriptome-wide, including in lncRNAs: 5-methylcytidine (m^5C), N6-methyladenosine (m^6A), pseudouridine (Ψ) and we also include adenosine to inosine (A-to-I) editing. A particular emphasis is placed on detailing specific examples where effects of the modifications on lncRNA function are already known.

2. Transcriptome-Wide Detection of Post-Transcriptional RNA Modifications

Advances in NGS technology have led to the development of techniques for the detection of several post-transcriptional modifications. These approaches are described below.

2.1. m^5C Detection Methods

Three major techniques have been described to detect m^5C sites in a transcriptome-wide manner: 1) bisulfite sequencing (bsRNA-seq), 2) 5-azacytidine RNA immunoprecipitation (Aza-IP), and 3) methylation individual-nucleotide-resolution crosslinking and immunoprecipitation (miCLIP).

The gold standard technique to map m^5C sites in DNA employs sodium bisulfite treatment of the nucleic acid prior to sequencing [10]. When subjected to bisulfite treatment, m^5C remains unchanged, whereas unmodified cytidine is converted to uridine, which is read as thymidine during sequencing. This approach can also be applied to RNA [11] and its adaptation for use with NGS, termed bsRNA-seq, led to the first transcriptome-wide survey of a RNA modification [12] (Figure 1A). In bsRNA-seq, total RNA is subjected to rRNA depletion and/or poly(A)-enrichment prior to bisulfite treatment, followed by NGS library preparation, sequencing and specialized mapping methods to detect candidate m^5C sites. Sites can be validated by locus-specific RT-PCR and amplicon sequencing using either Sanger sequencing or NGS [13]. Combination of these methods with siRNA-mediated knockdown of RNA:m^5C methyltransferases (MTases), or use of genetically deleted cells, can identify specific target sites of these MTases (e.g. [12, 14]).

bsRNA-seq does come with some limitations. For instance, secondary structures may impede the bisulfite-induced cytidine deamination, resulting in incomplete conversion and therefore false positive methylation calls. Furthermore, some other cytidine modifications, including 3-methylcytidine, N4-methylcytidine, N4,2´-O-dimethylcytidine, and N4-acetylated variants are expected to be at least partially resistant to bisulfite treatment [11, 12]. Also, while 5-hydroxymethylcytosine (hm^5C) reacts with bisulfite to yield cytosine 5-methylenesulfonate, this does not induce C-to-T transitions [15], thus m^5C and hm^5C are indistinguishable through bisulfite sequencing (see also below).

The Aza-IP technique (Figure 1A) identifies RNA:m^5C MTase targets by coupling immunoprecipitation with NGS [16, 17]. Specifically, cells that overexpress an affinity-tagged RNA:m^5C MTase of interest are grown in the presence of the anti-cancer drug 5-azacytidine, which becomes incorporated into nascent RNA. This results in a covalent bond forming between 5-azacytidine, acting as a suicide inhibitor, and the m^5C MTase at the target site. The isolation and identification of RNA targets is achieved by affinity purification of the

tagged MTase, limited RNase digestion and identification of protected fragments by RNA-seq. Although this method is not strictly site-specific, a propensity for C-to-G transversions at the crosslink site in RNA-seq reads, which has also been observed for DNA, can allow the identification of the targeted cytosine [17].

A similar approach, termed miCLIP (Figure 1A), uses affinity purification of an overexpressed, affinity-tagged mutant RNA:m^5C MTase followed by sequencing of associated RNA fragments using a customized individual-nucleotide-resolution cross-linking immunoprecipitation (iCLIP) method [18]. The UV-crosslinking step is omitted here as the MTase mutant version is trapped in an irreversible, covalently bound RNA-protein intermediate [19]. During library preparation, reverse transcription terminates at the cross-link site, allowing for m^5C detection at single nucleotide resolution.

2.2. m^6A Detection Methods

Transcriptome-wide approaches to map m^6A are based on the immunoprecipitation of fragmented, poly(A)-enriched RNA using m^6A-specific antibodies prior to NGS (MeRIP-seq or m^6A-seq) [20, 21] (Figure 1B). After sequencing, the reads are mapped to the reference genome. Reads from multiple different fragments that contain the same m^6A residue will overlap in the genome alignment producing a coverage peak, the midpoint of which is considered to be the location of the m^6A site [22]. As antibodies are known to exhibit low-level nonspecific RNA binding, the immunoprecipitation step is typically performed using two different m^6A antibodies, accepting only peaks that are detected in both data sets. Further complications with MeRIP-seq arise from the fact that available m^6A antibodies also recognize N6,2´-O-dimethyladenosine (m^6Am), which is commonly found at mRNA cap structures [23]. Also, m^6A can be formed through the rearrangement of N1-methyladenosine (m^1A) in alkaline conditions [24], potentially resulting in misidentification of the adenosine modification type by MeRIP-seq. The development of a NGS approach that identifies m^6A by specific chemical modification of m^6A, as has been achieved for Ψ (see below), might avoid these complications and also provide true single nucleotide resolution.

Furthermore, two techniques have been used to detect specific m^6A sites in lncRNAs, 1) high resolution melting and 2) site-specific cleavage and radioactive-labeling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET). The former method uses a set of oligodeoxyribonucleotide pairs complementary to a particular position in total RNA. One

oligonucleotide in each pair is modified with a fluorophore at its 5′ end, whereas the other carries a quencher at its 3′ end, such that the candidate m^6A site is directly opposite the 3′ quencher [25]. A melting point analysis determines whether m^6A is present.

SCARLET utilizes sequence-specific probes that flank the m^6A candidate site [26]. These act as precise guides to target RNase H to the cleavage site, which is then radiolabeled and splint-ligated to a DNA oligonucleotide. This results in a nucleic acid species that is resistant to RNase T1/A treatment. Thin-layer chromatography is used to determine whether m^6A is present.

2.3. Pseudouridine Detection Methods

The site-specific quantitative detection of Ψ in RNA is challenging as it is an isomer of uridine and thus 'mass-silent'. Detection can nevertheless be achieved by enzymatically hydrolyzing the RNA into nucleosides that are then identified by liquid chromatography and mass spectrometry (LC/MS), reviewed in [27]. Quantitative results can be obtained by coupling this technique with radiolabeling, endonuclease digestions, gel electrophoresis or ligation-based methods. Although these methods are quite involved, specific MS-based approaches have been developed for quantitative Ψ detection in RNA [28, 29].

To allow site-specific detection, methods were developed for specific labeling of Ψ (Figure 1C). Incubation with CMC (N-cyclohexyl-N_9-(2-morpholinoethyl)-carbodiimide metho-p-toluenesulphonate) results in the formation of a N_3-CMC-Ψ adduct that is resistant to alkaline hydrolysis [30-32]. Adducts formed with U or G are reversible and can be removed by alkaline hydrolysis treatment. N_3-CMC-Ψ adducts terminate reverse transcription one nucleotide downstream of the Ψ position [30]. Here, the RNA of interest, is subjected to poly(A)-enrichment and rRNA depletion, followed by fragmentation before treatment with CMC, and finally reverse transcription. A size shift to shorter cDNA fragments compared to the size of the original RNA fragments indicates stalling of the reverse transcriptase enzyme. These shorter cDNAs are used for library preparation and NGS, allowing determination of the Ψ site at single nucleotide resolution.

Several groups have independently implemented this type of approach and termed it Ψ-seq [33], Pseudo-seq [34] or PSI-seq (in yeast; [35]). A further adaptation of this method termed CeU-seq (N_3-CMC-enriched pseudouridine sequencing), was developed by Li *et al.* [36]. Here, the RNA is incubated with a CMC derivative, which subsequently allows labeling of

the formed N_3-CMC-Ψ adduct with DBCO-$(PEG)_4$-biotin using click chemistry. Immunoprecipitation of the pseudouridylated RNA using streptavidin beads results in the enrichment of Ψ-containing RNA followed by reverse transcription and sequencing; the inclusion of the immunoprecipitation step results in increased detection sensitivity. Thus, sites obtained using this technique likely represent the most comprehensive list of Ψ-containing RNAs in different tissues and during stress conditions.

2.4. A-to-I Detection Methods

A-to-I editing is the most common type RNA editing, comprising about 90% of all editing events in RNA [37, 38]. These editing events are typically identified by comparing cDNA sequences with the corresponding genomic DNA sequence (e.g. [39]). The edited inosine base pairs with cytidine in cDNA, hence editing is visible as an A-to-G sequence change (Figure 1D). Single nucleotide polymorphisms (SNPs), mapping and sequencing errors can obscure the identification of editing sites [40-42]. Sophisticated bioinformatics tools have thus been developed to maximize detection accuracy whilst minimizing the detection of false positives [43-45]. Furthermore, considerable efforts have been made to comprehensively detect and remove known SNPs from editing datasets and databases [45-47].

As the identification of true editing sites from transcriptome sequencing data is difficult, alternative methods aimed at 'marking' the inosine have been developed. Co-immunoprecipitation of ADAR (adenosine deaminase acting on RNA) enzymes with bound substrate RNA and subsequent microarray analysis of these associated RNA sequences was a first step to detect such 'marked' editing sites [48, 49]. However this method is compromised by the fact that association of ADAR with RNA is not necessarily indicative of editing. Therefore, alternative approaches that specifically target inosine-containing transcripts have been developed. These are based on the finding that glyoxal reacts with guanosine to form a stable adduct, whereas inosine glyoxal adducts are unstable. Guanosine glyoxal/borate adducts are resistant to RNase T1 digestion [50-52]. RNase T1 specifically cleaves RNA after guanosine or inosine, but is inhibited by guanosine glyoxal/borate adducts. This creates RNA fragments that carry inosine at their termini as an input for sequencing.

A similar approach to 'mark' inosine, termed Inosine Chemical Erasing (ICE) was developed by Sakurai *et al.* [53]. ICE involves the treatment of RNA with acrylonitrile, which converts the inosine to N1-cyanoethylinosine in the process of cyanoethylation, and results in the

formation of an inosine/acrylonitrile adduct that inhibits base pairing with cytidine and stalls reverse transcription (Figure 1D). Total RNA is either treated with acrylonitrile or left untreated and then reverse transcribed into cDNA. In untreated RNA, A or I at a given position is converted into T or G, respectively. In the treated sample A will be converted into T whilst the presence of inosine/acrylonitrile adducts block reverse transcription, leading to shorter cDNAs. The ICE method was extended further to combine it with NGS (ICE-seq), requiring fragmentation of poly(A)-enriched RNA before cyanoethylation and reverse transcription. ICE-seq allows transcriptome-wide detection of editing [54]. Gel purification of longer cDNA fragments effectively erases these shorter inosine/acrylonitrile adduct-containing cDNAs from the library. Subsequent sequencing and comparison of libraries identifies inosines by detecting erased reads upon cyanoethylation [54].

3. Types of Post-Transcriptional Modifications Detected in lncRNAs

3.1 Scope of Post-Transcriptional Modifications in lncRNAs

m^5C, m^6A, Ψ and A-to-I editing have each been documented in lncRNAs. Even though these modifications have been mapped transcriptome-wide, the focus of these studies was typically on mRNAs, and therefore, detection of modified sites in lncRNAs often received only scant attention. Thus, to illustrate the general scope for the presence of these four modified nucleosides in lncRNAs, we mined publicly available human datasets for candidate sites in transcripts annotated as lncRNAs. This analysis yielded a total of 9,965 m^5C sites in 1,072 transcripts, 15,357 m^6A sites in 12,348 transcripts, 162 Ψ in 150 transcripts, and 333,394 A-to-I editing sites in 8,832 transcripts (summarized in detail in Table 1). The substantial number of modifications identified in lncRNAs by this analysis highlights the potential for further discoveries in lncRNA epitranscriptomics.

3.2 5-Methylcytosine (m^5C)

tRNA Aspartic Acid Methyltransferase 1 (TRDMT1), and NOP2/Sun RNA Methyltransferase Family Members 2, 4, and 6 (NSUN2, NSUN4 and NSUN6) have each been shown to function as RNA:m^5C MTases in mammals [55-58]. Notably, TRDMT1-mediated

methylation inhibits stress-induced cleavage of modified tRNAs [59]. The NSUN family contains a further four proteins (NSUN1, NSUN3, NSUN5, and NSUN7), which are also predicted to methylate RNA based on sequence conservation of domains required for methyltransferase activity [60]. Several NSUN proteins are critically involved in disease. For example, mutations in NSUN2 cause autosomal-recessive intellectual disability syndromes [61]. NSUN5 is deleted in the Williams-Beuren syndrome [62], and a connection between a mutation in the NSUN7 gene and male sterility has been reported [63]. It remains to be discovered how the loss of RNA methylation contributes to the pathology of these diseases.

m^5C sites are prevalent in tRNAs and some sites are also known in rRNA [60]. These sites can be reliably detected by bisulfite amplicon sequencing [11], however, bsRNA-seq and other transcriptome-wide methods have together uncovered thousands of additional candidate sites in both coding and noncoding RNAs [12, 16, 19]. The role of m^5C in mRNA remains elusive. Direct effects of m^5C on RNA structure are expected to be subtle as it preserves the base pairing properties of the unmodified nucleoside. It is to be expected that the position of the m^5C site within the transcript will dictate its function, perhaps by attracting a binding protein or affecting the binding or functionality of other regulatory factors. However, no factor that specifically binds to or recognizes m^5C in RNA has yet been reported and little is known about the potential removal of m^5C from mRNA (but see section 3.2.2 below). A recent report suggests that NSUN2-mediated methylation in the 3´ UTR of the p16 mRNA confers stability [64]. Of interest, Squires *et al.* [12] noted that m^5C sites are statistically significantly enriched in mRNA 3´ UTRs, whilst Argonaute I-IV binding sites are enriched immediately upstream of m^5C sites. These observations suggest a role for m^5C methylation in the miRNA mechanism.

3.2.1. m^5C Sites in lncRNAs

Although not specifically enumerated, the recent bsRNA-seq, Aza-IP and miCLIP studies each reported numerous m^5C candidate sites in lncRNAs. Below we describe examples that have been investigated further.

3.2.1.1 Vault RNAs

Although at just under 100 nt too small in size for a *sensu strictu* lncRNA, the vault (vt)RNAs deserve a mention here. They form part of vault particles, abundant and widely conserved

cytoplasmic ribonucleoprotein complexes of unknown function [65]. Both miCLIP and Aza-IP identified m^5C sites in several vtRNAs [16, 19]. Position C70 in vtRNA1-1 was shown to be a NSUN2-dependent site that regulates its processing into small miRNA-like regulatory RNAs [66].

3.2.1.2 scaRNA2

Khoddami and Cairns [16] also detected a NSUN2-dependent m^5C site at position C316 in scaRNA2 by Aza-IP and validated it by bisulfite amplicon sequencing. scaRNA2 assembles into small Cajal body-specific ribonucleoproteins and guides 2′-O-methylation of U2 snRNA [67]. However, this site does not reside in a scaRNA2 region of known function and no follow-up work was done to assess the relevance of m^5C to scaRNA2 function.

3.2.1.3 RNase P

Three studies have independently identified and verified a NSUN2-dependent m^5C site at position C174 in RPPH1, the catalytic ncRNA component of RNase P. The site was first detected by bsRNA-seq [12] and confirmed by both miCLIP and Aza-IP [16, 19]. RNase P is involved in the processing of tRNA precursors, and may also play a role in RNA polymerase III transcription [68]. The identified m^5C site resides in the P12 domain of RPPH1 that may be important for RPPH1 function in transcription of 7SL RNA, 5S rRNA and several tRNAs [69]. Future work should address how m^5C in RPPH1 affects its function in transcription regulation or tRNA processing.

3.2.1.4 HOTAIR

Amort *et al.* [70] employed bisulfite amplicon sequencing to probe for the presence of m^5C sites in functionally relevant regions of the lncRNAs XIST and HOTAIR, both of which interact with chromatin-modifying complexes to regulate gene silencing. HOTAIR (HOX antisense intergenic RNA) is a lncRNA, ~2.2 kb in length, that is transcribed from the antisense strand of the HOXC gene cluster and contacts the histone methyltransferase polycomb-repressive complex 2 (PRC2) via its 5′ proximal region (also see article by Kaneko & Reinberg in this issue). Also, the 3′ end of HOTAIR is known to bind the histone demethylase complex, LSD1/CoREST/REST [71]. A single site (C1683) near the LSD1

binding region in HOTAIR was shown to be stoichiometrically methylated in several different human cell lines [70].

3.2.1.4 XIST

The ~17 kb long XIST transcript is derived from a regulatory locus termed X-inactivation center and functions to permanently inactivate one of the two X-chromosomes in female placental mammals [72]. XIST binds the PRC2 complex via its 5´ proximal A region [73]. Amort *et al.* [70] identified five substoichiometrically methylated, clustered m^5C sites (C701, C702, C703, C711 and C712) within repeat 8 of the A region of XIST in female human HEK293T cells. Interestingly, the corresponding sites in mouse XIST did not show any methylation. Furthermore, only about one in five human XIST molecules showed the modified cluster of sites, suggesting a regulatory rather than mandatory function of methylation. *In vitro* studies revealed that the cluster of m^5C sites in human XIST inhibited binding of purified PRC2 (Figure 2A). Overall, this evidence suggests that the presence of m^5C in XIST, and perhaps also HOTAIR, regulates epigenetic processes, by modulating their interactions with chromatin modifying protein complexes.

3.2.2. m^5C is a Dynamic Modification in RNA

In DNA, the conversion of m^5C to hm^5C by the ten-eleven translocation (TET) enzymes serves as both a step in an active DNA demethylation pathway as well as to create a stable DNA modification with potentially wide-ranging but as yet incompletely understood epigenetic functions [74, 75].

Interestingly, recent studies have heralded similar findings in epitranscriptomics research. Fu *et al.* [23] demonstrated that the TET enzyme family has the ability to catalyze the oxidative demethylation of m^5C to hm^5C in human RNA *in vitro* and *ex vivo* (in HEK293T cells). Huber *et al.* [76] employed an *in vivo* isotope-tracing approach to enable the detection of isotope-labeled m^5C and any directly formed derivatives, which are determined through LC-MS/MS techniques. This analysis resulted in the detection of hm^5C, evidently arising from the oxidation of m^5C. Furthermore, it was possible to identify hm^5C sites across all three domains of life (plants, nematodes and mammalian cells), indicating that hm^5C is a conserved RNA modification. Notably, the hm^5C modification was observed at a 40-fold increase in the poly(A)-enriched HEK293T samples compared to total RNA, indicating an enrichment of

hm^5C in mRNAs and/or certain polyadenylated lncRNAs [76], suggestive of a role in dynamic regulatory mechanisms. It will be important to map specific sites of hm^5C in RNA to gain functional insight.

3.3. N6-Methyladenosine (m^6A)

Writers, readers and erasers for m^6A in RNA have all been identified in recent years, reviewed in [7, 22]. The formation of m^6A is catalyzed by a protein complex containing the RNA:m^6A MTases, methyltransferase-like 3 and 14 (METTL3, METTL14). They further associate with a protein named KIAA1429 and the Wilm's tumour 1 associating protein (WTAP), which may direct the complex to nuclear speckles and/or target transcripts. The METTL3/14 complex modifies a large proportion of m^6A sites in the transcriptome found within the consensus sequence RRm^6ACH (R = A/G and H = A/C/U). However, adenine bases outside of this consensus context can also be targets for m^6A methylation. For example, a methylated adenine base, linked to a methylated ribose, is also often part of the mRNA cap structure, i.e. m^7Gpppm^6Am. Indeed, a proportion of m^6A sites in rRNAs, tRNAs and snRNAs, are also not located within the above mentioned motif, suggesting that other RNA:m^6A MTases remain to be discovered. Two m^6A demethylases, fat mass and obesity associated protein (FTO) and α-ketoglutarate-dependent dioxygenase alkB homolog 5 (ALKBH5), are also known, demonstrating that m^6A can be dynamically removed from RNA. Collectively, the RNA:m^6A MTases and demethylases have been shown to cause a spectrum of developmental phenotypes in model organisms, and their aberrant function is linked to developmental and neurological disorders in humans [7].

MeRIP-seq detected modified sites in thousands of mRNAs in different cell lines and tissues from mammalian, yeast or plant sources revealing both, relatively static and dynamically changing sites [20, 21, 77-79]. MeRIP-seq peaks are situated throughout all regions of mammalian mRNAs, with some concentration near the 5′ end, likely due to the prevalence of cap-proximal m^6Am, and a pronounced enrichment around stop codons. As with m^5C, the methyl group in m^6A does not alter its Watson-Crick base pairing, however, it may represent a preferential binding surface for a 'reader' protein. Alternatively, m^6A might alter local RNA structure to allow access for RNA-binding proteins to nearby sites, as has just been shown for heterogeneous nuclear ribonucleoprotein C (HNRNPC) [80], possibly by preventing Hoogsteen base pairing between an adenosine or uracil on the adenosine base in a A:U Watson-Crick base pair [22]. Research into the molecular function of m^6A in mRNAs is fast

evolving and no simple consensus mechanism is emerging, indicating a range of context-dependent functions. Evidence has been produced for m^6A effects on mRNA splicing, stability and translation, reviewed in [22]. For instance, a propensity for m^6A to co-occur, but not overlap, with miRNA target sites, and also to stabilize mRNA was noted early on [21]. By contrast, m^6A reportedly also destabilizes mRNAs, mediated by one of the identified 'reader' proteins, the human YTH domain family 2 (YTHDF2) [81]. YTHDF1, in turn was shown to promote target mRNA translation [82]. Other recognized m^6A readers are YTHDF3, YTHDC1, and Hu-antigen R (HuR) [21, 81, 83]. It would be expected that these newly identified m^6A readers also recognize a proportion of m^6A sites in lncRNAs.

3.3.1 m^6A Sites in lncRNAs

The recent transcriptome-wide studies have identified m^6A in lncRNAs, revealing in excess of 300 m^6A containing lncRNAs, reviewed in [23]. However, recent improvements in the bioinformatics analyses of such data and reanalysis of published HEK293T MeRIP-seq data [20] resulted in the detection of 1,847 candidate m^6A sites in lncRNAs, accounting for 12.1% of the total m^6A peaks [84].

3.3.2 MALAT1

Metastasis associated lung adenocarcinoma transcript 1 (MALAT1) is an abundant, highly conserved lncRNA of ~9kb length. MALAT1 is a scaffold component of nuclear speckles with roles in regulating mRNA transcription and splicing [85]. Liu *et al.* [26] identified two highly methylated sites (A2515 and A2577) in MALAT1, both within the GGACU consensus context. Additional secondary structure prediction and structural mapping experiments of *in vitro* transcribed MALAT1 RNA segments revealed that both sites are located in hairpin stems. Follow-up work implicated the modification in destabilization of the hairpin stem, allowing access of HNRNPC to a U-rich tract situated on the hairpin arm opposite of the m^6A site at position A2577 (Figure 2B). Although the functional consequences of altered HNRNPC binding to MALAT1 itself were not further addressed, it prompted MeRIP-seq and global HNRNPC footprinting efforts that identified tens of thousands of m^6A switches in the human transcriptome. Broadly, m^6A-switch-regulated HNRNPC-binding was shown to affect the abundance and alternative splicing of target mRNAs [80].

3.3.3. Pri-miRNAs

MicroRNAs (miRNAs) are small noncoding RNAs, typically 21-25 nt in length that regulate gene expression through imperfect complementarity to target mRNAs, resulting in RNA cleavage, translation inhibition or RNA deadenylation [86, 87]. The initial step in miRNA biogenesis is the processing of the primary miRNA (pri-miRNA) by the microprocessor complex, DGCR8 and Drosha, forming a precursor miRNA (pre-miRNA) hairpin [88].

The mechanism by which the Drosha/DGCR8 complex recognizes the correct region of the pri-miRNA for cleavage is incompletely understood. Alarcón *et al.* [89] noticed an enrichment of the GGACU consensus motif at the junction between the hairpin stem and the flanking single-stranded RNA regions of pri-miRNAs from MDA-MB-231 breast cancer cells. Of note, this motif was depleted in pre-miRNA transcripts. Interestingly, METTL3 depletion reduced the binding of DGCR8 to pri-miRNAs as well as causing the accumulation of unprocessed pri-miRNAs and concomitant reduction of mature miRNAs [89], suggesting a crucial role for m^6A in miRNA biogenesis. *In vitro* processing assays were further carried out with let-7 pri-miRNA transcripts containing a m^6A or unmodified A base showing that methylated pri-let-7e was more efficiently cleaved into pre-let-7e than the unmethylated variant [89]. These experiments indicate that m^6A has a direct role in the efficient processing of pri-miRNAs into pre-miRNAs (Figure 3A). Additional analyses also suggested that DGCR8 could directly interact with methylated RNA, as METTL3 depletion resulted in a significant reduction of DGCR8-bound RNA. Thus, m^6A methylation allows DGCR8 to differentially recognize pri-miRNAs over mRNAs, which can also form short hairpin secondary structures, similar to those observed in lncRNAs [89]. As miRNA levels and METTL3 expression are both known to be altered in various cancers, this study provided a first connection between these two cancer markers.

3.4. Pseudouridine (Ψ)

Ψ is the 5´-ribosyluracil isomer of uridine and commonly occurs in tRNA, rRNA and small nuclear RNA (snRNA) [90]. mRNAs were not known to be pseudouridylated, however, Ψ-seq, or variants thereof, recently identified hundreds of novel Ψ sites in mRNA from human, mouse or yeast cells [33-36]. A mixture of conserved, inducible and tissue-specific sites were detected, suggesting plasticity and a regulatory purpose of Ψ. One report provided suggestive evidence that Ψ enhances mRNA stability [33].

In eukaryotes, uridine isomerization is mediated by either a guide RNA-dependent or -independent mechanism [90-92]. Box H/ACA snoRNAs form small nucleolar RNA-protein complexes (snoRNPs), which include the catalytic subunit Dyskerin/DKC1, and target pseudouridylation of specific rRNA regions. RNA-independent isomerization is carried out by stand-alone pseudouridine synthases from the PUS family. There are several PUS members known to target specific positions in specific RNAs [93] and recent Ψ-seq studies expanded the target repertoire of several of the Ψ synthases to include mRNAs. No reader of Ψ has as yet been described and an eraser is not expected to exist, as isomerisation of uridine to Ψ is considered to be irreversible. Ψ synthases are implicated in human disease, mutations in PUS1 cause mitochondrial myopathy and sideroblastic anemia (MLASA) [94], while DKC1 mutations cause X-linked dyskeratosis congenita [95].

Ψ can induce several changes to the physical/chemical characteristics of RNA, such as altered RNA structure, increased base stacking and greater backbone rigidity. This is possible as Ψ has a C-C linkage between the base and the sugar and also contains an extra hydrogen bond donor [96]. Additionally, a Ψ-A pair is thought to be more stable than a U-A pair and therefore Ψ improves base pairing [97]. This provides much scope for Ψ-mediated altered recognition of RNA. For example, pseudouridylation of the first position of all three stop codons using artificial Box H/ACA guide RNAs, converted them to sense codons with unexpected coding properties [98].

3.4.1. Ψ in lncRNAs

The recent Ψ-seq studies identified multiple Ψ candidate sites in lncRNAs, [33-36]. Carlile *et al.* [34] reported the presence of several Ψ sites in human lncRNAs. Two were detected in the aforementioned MALAT1 (U5160, U5590). A single site was found in each, LRRC75A Antisense RNA 1 (LRRC75A-AS1; U1537), Small Nucleolar RNA Host Gene 1 (SNHG1; U1766) and 7SK small nuclear RNA (RN7SK; U250). In yeast, the RNase MRP RNA (NME1) and RNase P (RPR1) each were determined to carry a single Ψ (U274 and U70, respectively), whereas the telomerase RNA (TLC1) contained four sites (U218, U253, U506, U1005). The study by Li *et al.* [36] confirmed the two previously identified Ψ sites (U5160 and U5590; [34]) in MALAT1, and detected one additional site, at U3374. Also, a single site in each XIST (U11249; see above) and KCNQ1 Opposite Strand/Antisense Transcript 1 (KCNQ1OT1; U64919) were identified to be pseudouridylated along with several sites in LRRC75A-AS1. Interestingly, many of the lncRNAs identified in these studies have a role in

disease [36]. As the function of pseudouridylation of these lncRNAs was not further investigated, we focus below on the lncRNA, steroid receptor RNA activator (SRA), where some progress has already been made to understand the functionality and the dynamics of pseudouridylation.

3.4.2. Steroid Receptor RNA Activator

The widely conserved and cancer-associated SRA gene locus gives rise to both coding (cSRA) and noncoding (ncSRA) transcript variants, reviewed in [99, 100]. Around 900 nt in length, ncSRA was discovered first and shown to act as a coactivator of several nuclear hormone receptors (NR) [101]. ncSRA adopts a complex secondary structure containing four domains with 25 helical regions [102] and functions as a scaffold for multicomponent NR complexes, which include the Ψ synthases PUS1 and PUS3 [103, 104]. Here, PUS1 functions by directly interacting with the DNA binding domain of NRs but also by modifying ncSRA [103]. A subsequent study showed that, similar to PUS1, PUS3 can also modify ncSRA [104]. Incubation of hSRA with PUS1 or PUS3 *in vitro* showed the modification of seven and six sites, respectively, some of which are common. Only one of these Ψ sites was shown to be modified *in vivo* and mapped to position U206, which is located in the stem-loop structure STR5 of ncSRA. In order to determine the importance of this position, a mutated form of hSRA, hSRA-U206A was analyzed. Surprisingly, hSRA-U206A became hyper-pseudouridylated, with PUS1 and PUS3 targeting ten and eight sites, respectively. This suggested that position U206 plays a central role in regulating the extent of ncSRA pseudouridylation and as such its functionality. When testing hSRA and hSRA-U206A function on estrogen and androgen receptors, as expected hSRA showed coactivator activity, whereas hSRA-U206A functioned as corepressor (Figure 2C), further indicating that position U206 is a key residue that regulates ncSRA function through regulating the extent of pseudouridylation [104].

Targeting ncSRA pseudouridylation has the therapeutic potential to reduce the level of steroid receptor signaling in hormone-sensitive cancer cells. Ghosh, Patton and Spanjaard [105] demonstrated that synthetic STR5 stem-loop RNA can inhibit PUS1-mediated pseudouridylation of SRA *in vitro*. When transfected into cultured cells it reduced the levels of estrogen receptor α and androgen receptor signaling in breast and prostate cancer cells, respectively.

3.5. A-to-I editing

RNA editing is one of the most common modifications in both coding and noncoding RNAs and is most abundant in the brain [106]. Several types of editing exist, but A-to-I RNA editing is by far the most frequent, representing more than 90% of editing events [37, 38]. A-to-I editing (referred to as editing from here on) occurs in double-stranded RNA (dsRNA) molecules. Duplex formation is promoted by the presence of inverted repetitive sequences, such as the Alu element, a member of the SINE family of retroelements. Thus, editing is very strongly associated with Alu elements, although editing also occurs in other repetitive elements such as LINE 1 retrotransposons [107]. Editing events can either be specific (occuring in shorter duplexes disrupted by mismatches and bulges) or promiscuous (in longer, more stable duplexes).

Editing is achieved by deamination of adenosine by the ADAR enzymes. Three ADAR enzymes are present in the human genome, ADAR1, which has two isoforms, ADAR1p110 and ADAR1p150, ADAR2 and ADAR3. ADAR1 and ADAR2 show similar 5′ neighbor preferences in their substrates U=A>C>G, but ADAR2 also has a 3′ neighbor sequence preference (U=G>C=A). The substrates for ADAR enzymes are dsRNA regions and both enzymes show selectivity depending on duplex length, mismatches and bulges. Editing does not occur close to duplex ends and any adenosine present within the dsRNA sequence is a potential substrate for editing. Almost half of all adenosine residues present within a duplex can be edited before destabilization of the double-stranded structure occurs, in a process termed hyperediting [108]. ADAR enzymes play an important role in neuron development and innate immunity. ADAR2 targets the mRNAs for several different ion channels and G protein-coupled receptors in the nervous system and loss of ADAR2-mediated editing of the Q/R site in the AMAP receptor subunit GluA2 results in seizures and early death in mice [109]. ADAR1 mutations are characterized by altered interferon production causing Aicardi-Goutières syndrome and dystonia, and ADAR1 mutant mice show embryonic lethality [110, 111]. Loss of ADAR1-mediated hyperediting of dsRNA regions in 3′ UTRs leads to altered RNA folding, resulting in the stimulation of the innate immune system [110]. Simultaneous deletion of MDA5, a cytoplasmic dsRNA sensor, rescues this phenotype by suppressing the innate immune system [111].

Editing can profoundly affect cellular processes, as inosine base pairs with cytidine and is therefore interpreted as guanosine by the cellular machinery. Editing results in destabilization of the RNA molecule as I-U base pairs are less stable than A-U base pairs, resulting in

changes within the RNA structure [112]. Also, editing of coding sequences can change the genetic code and consequently alter the encoded protein sequence. Although this is rare, some mRNAs undergo editing, which is crucial for their function [113].

Many transcriptome-wide surveys reporting vast numbers of candidate editing sites are now available [39, 44, 45, 54, 114, 115]. For example, in excess of 100 million editing sites were found in human brain RNA [114]. This led to the establishment of the public editing database DARNED (database of RNA editing in humans) compiling up-to-date published editing datasets for the human genome [46]. This database was later extended to include data obtained in fly and mouse [116]. As the detection of RNA editing became easier, a second database, RADAR (rigorously annotated database of A-to-I RNA editing), was established listing human, fly and mouse data, and detailing tissue specific editing levels for each site, as well as manually curated data for each transcript [117].

3.5.1. Editing in lncRNAs

Alu elements are one of the most common repetitive sequences in the human genome and frequently found in noncoding sequences, such as pseudogenes, transposable element genes, lncRNAs or primary miRNA transcripts. Indeed, the editing databases (RADAR and DARNED) list multiple lncRNAs. Determining the extent and function of editing in lncRNAs is critical, as editing has the potential to alter interactions of the lncRNA with other nucleic acids or proteins, ultimately affecting their biological functions in diverse ways.

One of the first studies to enumerate editing in human lncRNAs was by Peng *et al.* [44], specifically listing 31 sites in MALAT1 and 41 sites in the XIST activator transcript JPX, as well as in several other lncRNAs. Subsequent studies identified further lncRNAs that had undergone editing [45, 54]. This list was extended to include a single editing site in a 7SK-like element and a further 5 clustered sites were detected in the intronic region of the coding gene MAN2A1, which contains a 7SK pseudogene in antisense orientation [54]. Although the presence of editing in lncRNAs has been established, whether these events have any consequence on their function remains unclear.

A detailed investigation into the effects of editing on a particular lncRNA has as yet not been conducted. However, many pri-miRNA transcripts are targeted for editing, and in this context more is known. The extent and biological consequences of editing in pri-miRNAs are described below.

3.5.2. pri-miRNAs

Primary miRNA transcripts form hairpins containing considerable dsRNA portions, thus making them plausible targets for editing enzymes. Primary miRNA editing can affect each step in miRNA biogenesis, e.g. changing the processing by Drosha or Dicer due to altered pri- and/or pre-miRNA structure, or the targeting properties of the mature miRNA, especially if editing occurred in the miRNA seed region. That miRNAs can indeed undergo editing was first demonstrated by detecting several sites in the pri-miR22 transcript in brain tissue from human and mice [118]. The phenomenon has since been shown to occur in a considerable number of miRNAs, although the levels of editing can be quite low (<10%), e.g. [118-122]. An estimated 16% of pri-miRNAs are edited in human brain [123]. Furthermore, both ADAR enzymes have been implicated in the regulation of pri-miRNA as well as mature miRNA levels [121, 124-127].

The interest in editing of mature miRNAs led to the development of an online tool, DREAM (detecting RNA editing associated with micro RNAs [128]). Furthermore, the miR-EdiTar database lists predicted miRNA binding sites that are edited themselves, or could become targets due to mature miRNA editing [129].

The first evidence that editing of pri- or pre-miRNA transcripts affects processing came from Nishikura and co-workers. Yang *et al.* [130] selected pri-miR142, as both mature miR142-5p and miR142-3p contained six and three highly edited sites, respectively. Cells were transfected with mimics of unmodified or edited pri-miR142 transcripts and the levels of pri-, pre- and mature miR142 monitored. Cells expressing pre-edited pri-miR142 produced virtually no pre-miR142 nor did mature miR142 accumulate, with editing sites near the Drosha cleavage site having the strongest effect. Furthermore, it could be shown that both ADAR enzymes edit the pri-miR142 transcript *in vitro*, confirming that ADAR can target the secondary structure of primary miRNA transcripts, and editing affected Drosha processing *in vitro*. Additionally, ADAR1$^{-/-}$ and ADAR2$^{-/-}$ mice, deficient in editing, showed increased levels of mature miR142, likely due to increased pri-miR142 processing efficiency, providing strong evidence that ADAR editing of pri-miR142 does occur and inhibits cleavage by Drosha *in vivo* (Figure 3B) [130]. Further examples of both, suppression or enhancement of Drosha processing by pri-miRNA editing have since been documented [123]. Some hyperedited, nuclear-retained pri-miRNAs have been found to be stable, while others are targeted for degradation by the Tudor-SN ribonuclease [125, 130, 131].

A subsequent study provided evidence that editing also alters processing of pre-miRNAs by

Dicer [112]. As in the previous study, processing intermediates of pre-edited or unedited pri-miR151 were monitored. Primary miR151 has been shown to be edited at several sites in human brain to varying degree, depending on the tissue [120]. While pre-miR151 was almost completely edited at the +3 site, it was impossible to detect any edited mature miR151-3p, indicating that editing affects the Dicer cleavage step. *In vitro* processing experiments showed that ADAR1 edited pri-miR151 and this was processed by Drosha into pre-miR151, but was not processed into mature miR151 by Dicer, although EMSA assays suggested that editing did not inhibit Dicer binding [112]. This indicates that editing specifically inhibits the cleavage of pre-miRNAs into mature miRNAs (Figure 3B).

However, some edited mature miRNAs can accumulate. Kawahara *et al.* [132] initially investigated the effect of editing on the primary transcript of the miR376 cluster. A single pri-miR376 is transcribed from this cluster, which gives rise to six individual mature miR376 variants. The primary transcript is heavily edited (+4 and +44 sites) in brain tissue, with some of the pre-miR376 transcripts being almost fully edited at the +44 position. cDNA sequencing of several mature miR376 variants showed that high levels of edited mature miR376 are present in various brain tissues, indicating that editing does not interfere with processing by either Drosha or Dicer [132]. As both editing sites (+4 and +44) are located within in the mature miR376 seed regions (5p and 3p, respectively), accumulation of edited miR376 could result in altered target selection. Computational target prediction of +4 edited mature miR376a-5p showed 82 potential targets, only two of which are in common with the 78 potential targets predicted for unedited miR376a-5p. Luciferase reporter constructs containing target sites for either the edited or unedited mature miR376-5p were co-transfected with pre-edited (+4 A-to-I or A-to-G substitutions) or unedited miR376-5p. This showed that miRNA target recognition was specific with no cross-repression occurring (i.e. the edited mature miRNA did not repress the luciferase reporter gene containing the target site for the unedited mature miRNA and vice versa). Furthermore, endogenous levels of the edited mature miR376-5p target transcripts were 2-fold lower in WT compared to ADAR2$^{-/-}$ mouse brain tissue, whereas transcript levels of the unedited mature miR376-5p target remained unchanged. Interestingly, one of the alternative targets of edited mature miR376-5p is involved in the uric acid synthesis pathway, suggesting that miR376 editing plays a role in uric acid regulation in select tissues [132]. This study provided the first experimental evidence that edited mature miRNA seed sequences, resulting from pri-miRNA transcript editing, could lead to miRNA re-direction and the subsequent recognition of an altered set of

target genes (Figure 3B).

This type of target reassignment by editing of mature miRNA has since been shown to occur for a range of additional miRNAs [131] and could play an important role in cancer [26, 133, 134]. For example, unedited mature miR376a* was shown to promote invasive growth of glioblastoma cells, whereas edited mature miR376* suppressed this phenotype [133]. It was further shown that unedited and edited mature miR376* target different genes, RAP2 and AMFR respectively. This was further confirmed by the siRNA-mediated downregulation of RAP2, which resulted in the same invasive phenotype, providing direct evidence that unedited mature miR376*-mediated RAP2 downregulation is linked to increased invasive growth of glioma cells [133]. Similarly, altered editing of miR455-5p was shown to be involved in regulating growth and invasiveness of melanoma [134]. Highly metastatic melanoma cells show reduced ADAR1 levels and an inverse correlation between ADAR1 levels and tumor growth was demonstrated. ADAR1 targets miR455 in melanoma cells and loss of ADAR1 led to increased levels of unedited mature miR455-5p. Unedited mature miR455-5p targets the CPEB1 gene, which is a known tumor suppressor gene, whereas edited mature miR455-5p is proposed to target genes with a known role in tumor growth promotion. It was shown that overexpression of unedited miR455-5p promotes tumor growth and development of lung metastasis, whereas overexpression of edited miR455-5p suppressed this phenotype [134]. These studies outline the potential far-reaching impact of editing events in human disease, as miRNA biogenesis and ADAR expression are often altered in cancers.

3.6. Outlook/Future Directions

The subject matter summarized here sits at the intersection between two newly emerging fields and thus much of the available information is fragmentary and many open questions remain. Few examples of a lncRNA can currently be named where both, its cellular function as well as the role of the epitranscriptomic marks it carries, have been documented. Nevertheless, the examples highlighted in this review show that modifications such as m^5C, m^6A and Ψ can regulate the interactions of lncRNAs with cognate binding proteins to regulate their function. A-to-I editing and m^6A modification of pri-miRNAs illustrate how a whole range of RNA processing steps and downstream functions can be influenced by an epitranscriptomic mark. Despite the small number of specific examples featured here, there are already cases of lncRNAs that carry several different modifications, e.g. MALAT1 exhibits m^6A, Ψ and editing sites. A further focus on mechanistic studies involving known

examples, together with the identification of further readers, writers and erasers, promises exciting breakthrough findings in coming years.

Efforts to map modifications in lncRNAs should also continue. Better and more sensitive methods to map the already known modifications should be joined by new methods to look for the incidence of some of the many additional species that exist in the zoo of chemical RNA modifications. Approaches and information emerging from the sister field of mRNA epitranscriptomics research will no doubt contribute to these endeavors.

Acknowledgements

This work was supported by a grant from the National Health and Medical Research Council of Australia (#1061551) to T.P.

References

[1] J.S. Mattick, J.L. Rinn, Discovery and annotation of long noncoding RNAs, Nat. Struct. Mol. Biol., 22 (2015) 5-7.

[2] X.C. Quek, D.W. Thomson, J.L. Maag, N. Bartonicek, B. Signal, M.B. Clark, B.S. Gloss, M.E. Dinger, lncRNAdb v2.0: expanding the reference database for functional long noncoding RNAs, Nucleic Acids Res., 43 (2015) D168-173.

[3] I.V. Novikova, S.P. Hennelly, C.S. Tung, K.Y. Sanbonmatsu, Rise of the RNA machines: exploring the structure of long non-coding RNAs, J. Mol. Biol., 425 (2013) 3731-3746.

[4] W.A. Cantara, P.F. Crain, J. Rozenski, J.A. McCloskey, K.A. Harris, X. Zhang, F.A. Vendeix, D. Fabris, P.F. Agris, The RNA Modification Database, RNAMDB: 2011 update, Nucleic Acids Res., 39 (2011) D195-201.

[5] M.A. Machnicka, K. Milanowska, O. Osman Oglou, E. Purta, M. Kurkowska, A. Olchowik, W. Januszewski, S. Kalinowski, S. Dunin-Horkawicz, K.M. Rother, M. Helm, J.M. Bujnicki, H. Grosjean, MODOMICS: a database of RNA modification pathways--2013 update, Nucleic Acids Res., 41 (2013) D262-267.

[6] Y. Saletore, K. Meyer, J. Korlach, I.D. Vilfan, S. Jaffrey, C.E. Mason, The birth of the Epitranscriptome: deciphering the function of RNA modifications, Genome Biol., 13 (2012) 175.

[7] S. Blanco, M. Frye, Role of RNA methyltransferases in tissue renewal and pathology, Curr. Opin. Cell Biol., 31 (2014) 1-7.

[8] S.R. Jaffrey, An expanding universe of mRNA modifications, Nat. Struct. Mol. Biol., 21 (2014) 945-946.
[9] S.R. Lee, J. Lykke-Andersen, Emerging roles for ribonucleoprotein modification and remodeling in controlling RNA fate, Trends Cell Biol., 23 (2013) 504-510.
[10] S.J. Clark, J. Harrison, C.L. Paul, M. Frommer, High sensitivity mapping of methylated cytosines, Nucleic Acids Res., 22 (1994) 2990-2997.
[11] M. Schaefer, T. Pollex, K. Hanna, F. Lyko, RNA cytosine methylation analysis by bisulfite sequencing, Nucleic Acids Res., 37 (2009) e12.
[12] J.E. Squires, H.R. Patel, M. Nousch, T. Sibbritt, D.T. Humphreys, B.J. Parker, C.M. Suter, T. Preiss, Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA, Nucleic Acids Res., 40 (2012) 5023-5033.
[13] T. Sibbritt, A. Shafik, S.J. Clark, T. Preiss, Nucleotide-level profiling of m5C RNA methylation, Methods Mol. Biol., in press (2015).
[14] F. Tuorto, R. Liebers, T. Musch, M. Schaefer, S. Hofmann, S. Kellner, M. Frye, M. Helm, G. Stoecklin, F. Lyko, RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis, Nat. Struct. Mol. Biol., 19 (2012) 900-905.
[15] Y. Huang, W.A. Pastor, Y. Shen, M. Tahiliani, D.R. Liu, A. Rao, The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing, PLoS One, 5 (2010) e8888.
[16] V. Khoddami, B.R. Cairns, Identification of direct targets and modified bases of RNA cytosine methyltransferases, Nat. Biotechnol., 31 (2013) 458-464.
[17] V. Khoddami, B.R. Cairns, Transcriptome-wide target profiling of RNA cytosine methyltransferases using the mechanism-based enrichment procedure Aza-IP, Nat. Protoc., 9 (2014) 337-361.
[18] I. Huppertz, J. Attig, A. D'Ambrogio, L.E. Easton, C.R. Sibley, Y. Sugimoto, M. Tajnik, J. Konig, J. Ule, iCLIP: protein-RNA interactions at nucleotide resolution, Methods, 65 (2014) 274-287.
[19] S. Hussain, J. Aleksic, S. Blanco, S. Dietmann, M. Frye, Characterizing 5-methylcytosine in the mammalian epitranscriptome, Genome Biol., 14 (2013) 215.
[20] K.D. Meyer, Y. Saletore, P. Zumbo, O. Elemento, C.E. Mason, S.R. Jaffrey, Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons, Cell, 149 (2012) 1635-1646.
[21] D. Dominissini, S. Moshitch-Moshkovitz, S. Schwartz, M. Salmon-Divon, L. Ungar, S. Osenberg, K. Cesarkas, J. Jacob-Hirsch, N. Amariglio, M. Kupiec, R. Sorek, G. Rechavi,

Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq, Nature, 485 (2012) 201-206.
[22] K.D. Meyer, S.R. Jaffrey, The dynamic epitranscriptome: N6-methyladenosine and gene expression control, Nat. Rev. Mol. Cell Biol., 15 (2014) 313-326.
[23] Y. Fu, D. Dominissini, G. Rechavi, C. He, Gene expression regulation mediated through reversible m(6)A RNA methylation, Nat Rev Genet, 15 (2014) 293-306.
[24] J.B. Macon, R. Wolfenden, 1-Methyladenosine. Dimroth rearrangement and reversible reduction, Biochemistry, 7 (1968) 3453-3458.
[25] A.Y. Golovina, M.M. Dzama, K.S. Petriukov, T.S. Zatsepin, P.V. Sergiev, A.A. Bogdanov, O.A. Dontsova, Method for site-specific detection of m6A nucleoside presence in RNA based on high-resolution melting (HRM) analysis, Nucleic Acids Res., 42 (2014) e27.
[26] N. Liu, M. Parisien, Q. Dai, G. Zheng, C. He, T. Pan, Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA, RNA, 19 (2013) 1848-1856.
[27] A. Durairaj, P.A. Limbach, Mass spectrometry of the fifth nucleoside: a review of the identification of pseudouridine in nucleic acids, Anal. Chim. Acta, 623 (2008) 117-125.
[28] B. Addepalli, P.A. Limbach, Mass spectrometry-based quantification of pseudouridine in RNA, J. Am. Soc. Mass Spectrom., 22 (2011) 1363-1372.
[29] M. Taucher, B. Ganisl, K. Breuker, Identification, localization, and relative quantitation of pseudouridine in RNA by tandem mass spectrometry of hydrolysis products, Int J Mass Spectrom, 304 (2011) 91-97.
[30] A. Bakin, J. Ofengand, 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 (1993) 9754-9762.
[31] N.W. Ho, P.T. Gilham, The reversible chemical modification of uracil, thymine, and guanine nucleotides and the modification of the action of ribonuclease on ribonucleic acid, Biochemistry, 6 (1967) 3632-3639.
[32] J. Ofengand, M. Del Campo, Y. Kaya, Mapping pseudouridines in RNA molecules, Methods, 25 (2001) 365-373.
[33] S. Schwartz, D.A. Bernstein, M.R. Mumbach, M. Jovanovic, R.H. Herbst, B.X. Leon-Ricardo, J.M. Engreitz, M. Guttman, R. Satija, E.S. Lander, G. Fink, A. Regev, Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA, Cell, 159 (2014) 148-162.

[34] T.M. Carlile, M.F. Rojas-Duran, B. Zinshteyn, H. Shin, K.M. Bartoli, W.V. Gilbert, Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells, Nature, 515 (2014) 143-146.
[35] A.F. Lovejoy, D.P. Riordan, P.O. Brown, Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in S. cerevisiae, PLoS One, 9 (2014) e110799.
[36] X. Li, P. Zhu, S. Ma, J. Song, J. Bai, F. Sun, C. Yi, Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome, Nat. Chem. Biol., 11 (2015) 592-597.
[37] A. Mallela, K. Nishikura, A-to-I editing of protein coding and noncoding RNAs, Crit. Rev. Biochem. Mol. Biol., 47 (2012) 493-501.
[38] G. Nigita, D. Veneziano, A. Ferro, A-to-I RNA Editing: Current Knowledge Sources and Computational Approaches with Special Emphasis on Non-Coding RNA Molecules, Front Bioeng Biotechnol, 3 (2015) 37.
[39] J.H. Bahn, J.H. Lee, G. Li, C. Greer, G. Peng, X. Xiao, Accurate identification of A-to-I RNA editing in human by transcriptome sequencing, Genome Res., 22 (2012) 142-150.
[40] C.L. Kleinman, J. Majewski, Comment on "Widespread RNA and DNA sequence differences in the human transcriptome", Science, 335 (2012) 1302; author reply 1302.
[41] W. Lin, R. Piskol, M.H. Tan, J.B. Li, Comment on "Widespread RNA and DNA sequence differences in the human transcriptome", Science, 335 (2012) 1302; author reply 1302.
[42] J.K. Pickrell, Y. Gilad, J.K. Pritchard, Comment on "Widespread RNA and DNA sequence differences in the human transcriptome", Science, 335 (2012) 1302; author reply 1302.
[43] M.J. de Hoon, R.J. Taft, T. Hashimoto, M. Kanamori-Katayama, H. Kawaji, M. Kawano, M. Kishima, T. Lassmann, G.J. Faulkner, J.S. Mattick, C.O. Daub, P. Carninci, J. Kawai, H. Suzuki, Y. Hayashizaki, Cross-mapping and the identification of editing sites in mature microRNAs in high-throughput sequencing libraries, Genome Res., 20 (2010) 257-264.
[44] Z. Peng, Y. Cheng, B.C. Tan, L. Kang, Z. Tian, Y. Zhu, W. Zhang, Y. Liang, X. Hu, X. Tan, J. Guo, Z. Dong, Y. Liang, L. Bao, J. Wang, Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome, Nat. Biotechnol., 30 (2012) 253-260.

[45] G. Ramaswami, R. Zhang, R. Piskol, L.P. Keegan, P. Deng, M.A. O'Connell, J.B. Li, Identifying RNA editing sites using RNA sequencing data alone, Nat Methods, 10 (2013) 128-132.
[46] A. Kiran, P.V. Baranov, DARNED: a DAtabase of RNa EDiting in humans, Bioinformatics, 26 (2010) 1772-1776.
[47] R. Piskol, G. Ramaswami, J.B. Li, Reliable identification of genomic variants from RNA-seq data, Am. J. Hum. Genet., 93 (2013) 641-651.
[48] J. Ohlson, M. Enstero, B.M. Sjoberg, M. Ohman, A method to find tissue-specific novel sites of selective adenosine deamination, Nucleic Acids Res., 33 (2005) e167.
[49] J. Ohlson, M. Ohman, A method for finding sites of selective adenosine deamination, Methods Enzymol., 424 (2007) 289-300.
[50] P.B. Cattenoz, R.J. Taft, E. Westhof, J.S. Mattick, Transcriptome-wide identification of A > I RNA editing sites by inosine specific cleavage, RNA, 19 (2013) 257-270.
[51] D.P. Morse, B.L. Bass, Detection of inosine in messenger RNA by inosine-specific cleavage, Biochemistry, 36 (1997) 8429-8434.
[52] C.N. Tseng, H.W. Chang, J. Stocker, H.C. Wang, C.C. Lu, C.H. Wu, J.G. Yang, C.L. Cho, H.W. Huang, A method to identify RNA A-to-I editing targets using I-specific cleavage and exon array analysis, Mol. Cell. Probes, 27 (2013) 38-45.
[53] M. Sakurai, T. Yano, H. Kawabata, H. Ueda, T. Suzuki, Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome, Nat. Chem. Biol., 6 (2010) 733-740.
[54] M. Sakurai, H. Ueda, T. Yano, S. Okada, H. Terajima, T. Mitsuyama, A. Toyoda, A. Fujiyama, H. Kawabata, T. Suzuki, A biochemical landscape of A-to-I RNA editing in the human brain transcriptome, Genome Res., 24 (2014) 522-534.
[55] M.G. Goll, F. Kirpekar, K.A. Maggert, J.A. Yoder, C.L. Hsieh, X. Zhang, K.G. Golic, S.E. Jacobsen, T.H. Bestor, Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2, Science, 311 (2006) 395-398.
[56] M. Frye, F.M. Watt, The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors, Curr. Biol., 16 (2006) 971-981.
[57] M.D. Metodiev, H. Spahr, P. Loguercio Polosa, C. Meharg, C. Becker, J. Altmueller, B. Habermann, N.G. Larsson, B. Ruzzenente, NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly, PLoS Genet, 10 (2014) e1004110.

[58] S. Haag, A.S. Warda, J. Kretschmer, M.A. Gunnigmann, C. Hobartner, M.T. Bohnsack, NSUN6 is a human RNA methyltransferase that catalyzes formation of m5C72 in specific tRNAs, RNA, 9 (2015) 1532-43.
[59] M. Schaefer, T. Pollex, K. Hanna, F. Tuorto, M. Meusburger, M. Helm, F. Lyko, RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage, Genes Dev., 24 (2010) 1590-1595.
[60] Y. Motorin, F. Lyko, M. Helm, 5-methylcytosine in RNA: detection, enzymatic formation and biological functions, Nucleic Acids Res., 38 (2010) 1415-1430.
[61] L. Abbasi-Moheb, S. Mertel, M. Gonsior, L. Nouri-Vahid, K. Kahrizi, S. Cirak, D. Wieczorek, M.M. Motazacker, S. Esmaeeli-Nieh, K. Cremer, R. Weissmann, A. Tzschach, M. Garshasbi, S.S. Abedini, H. Najmabadi, H.H. Ropers, S.J. Sigrist, A.W. Kuss, Mutations in NSUN2 cause autosomal-recessive intellectual disability, Am. J. Hum. Genet., 90 (2012) 847-855.
[62] A. Doll, K.H. Grzeschik, Characterization of two novel genes, WBSCR20 and WBSCR22, deleted in Williams-Beuren syndrome, Cytogenet. Cell Genet., 95 (2001) 20-27.
[63] T. Harris, B. Marquez, S. Suarez, J. Schimenti, Sperm motility defects and infertility in male mice with a mutation in Nsun7, a member of the Sun domain-containing family of putative RNA methyltransferases, Biol. Reprod., 77 (2007) 376-382.
[64] X. Zhang, Z. Liu, J. Yi, H. Tang, J. Xing, M. Yu, T. Tong, Y. Shang, M. Gorospe, W. Wang, The tRNA methyltransferase NSun2 stabilizes p16INK(4) mRNA by methylating the 3'-untranslated region of p16, Nat Commun, 3 (2012) 712.
[65] H. Persson, A. Kvist, J. Vallon-Christersson, P. Medstrand, A. Borg, C. Rovira, The non-coding RNA of the multidrug resistance-linked vault particle encodes multiple regulatory small RNAs, Nat. Cell Biol., 11 (2009) 1268-1271.
[66] S. Hussain, A.A. Sajini, S. Blanco, S. Dietmann, P. Lombard, Y. Sugimoto, M. Paramor, J.G. Gleeson, D.T. Odom, J. Ule, M. Frye, NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs, Cell Reports, 4 (2013) 255-261.
[67] K.T. Tycowski, A. Aab, J.A. Steitz, Guide RNAs with 5' caps and novel box C/D snoRNA-like domains for modification of snRNAs in metazoa, Curr. Biol., 14 (2004) 1985-1995.

[68] R. Reiner, Y. Ben-Asouli, I. Krilovetzky, N. Jarrous, A role for the catalytic ribonucleoprotein RNase P in RNA polymerase III transcription, Genes Dev., 20 (2006) 1621-1635.
[69] N. Jarrous, V. Gopalan, Archaeal/eukaryal RNase P: subunits, functions and RNA diversification, Nucleic Acids Res., 38 (2010) 7885-7894.
[70] T. Amort, M.F. Souliere, A. Wille, X.Y. Jia, H. Fiegl, H. Worle, R. Micura, A. Lusser, Long non-coding RNAs as targets for cytosine methylation, RNA Biol., 10 (2013) 1003-1008.
[71] A. Bhan, S.S. Mandal, LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer, Biochim. Biophys. Acta, 1856 (2015) 151-164.
[72] A. Wutz, Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation, Nat Rev Genet, 12 (2011) 542-553.
[73] J. Zhao, B.K. Sun, J.A. Erwin, J.J. Song, J.T. Lee, Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome, Science, 322 (2008) 750-756.
[74] M.R. Branco, G. Ficz, W. Reik, Uncovering the role of 5-hydroxymethylcytosine in the epigenome, Nat Rev Genet, 13 (2012) 7-13.
[75] A. Breiling, F. Lyko, Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond, Epigenetics Chromatin, 8 (2015) 24.
[76] S.M. Huber, P. van Delft, L. Mendil, M. Bachman, K. Smollett, F. Werner, E.A. Miska, S. Balasubramanian, Formation and abundance of 5-hydroxymethylcytosine in RNA, Chembiochem, 16 (2015) 752-755.
[77] Y. Li, X. Wang, C. Li, S. Hu, J. Yu, S. Song, Transcriptome-wide N(6)-methyladenosine profiling of rice callus and leaf reveals the presence of tissue-specific competitors involved in selective mRNA modification, RNA Biol., 11 (2014) 1180-1188.
[78] S. Schwartz, S.D. Agarwala, M.R. Mumbach, M. Jovanovic, P. Mertins, A. Shishkin, Y. Tabach, T.S. Mikkelsen, R. Satija, G. Ruvkun, S.A. Carr, E.S. Lander, G.R. Fink, A. Regev, High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis, Cell, 155 (2013) 1409-1421.
[79] S. Schwartz, M.R. Mumbach, M. Jovanovic, T. Wang, K. Maciag, G.G. Bushkin, P. Mertins, D. Ter-Ovanesyan, N. Habib, D. Cacchiarelli, N.E. Sanjana, E. Freinkman, M.E. Pacold, R. Satija, T.S. Mikkelsen, N. Hacohen, F. Zhang, S.A. Carr, E.S. Lander, A. Regev, Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites, Cell Reports, 8 (2014) 284-296.

[80] N. Liu, Q. Dai, G. Zheng, C. He, M. Parisien, T. Pan, N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions, Nature, 518 (2015) 560-564.
[81] X. Wang, Z. Lu, A. Gomez, G.C. Hon, Y. Yue, D. Han, Y. Fu, M. Parisien, Q. Dai, G. Jia, B. Ren, T. Pan, C. He, N6-methyladenosine-dependent regulation of messenger RNA stability, Nature, 505 (2014) 117-120.
[82] X. Wang, B.S. Zhao, I.A. Roundtree, Z. Lu, D. Han, H. Ma, X. Weng, K. Chen, H. Shi, C. He, N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency, Cell, 161 (2015) 1388-1399.
[83] C. Xu, X. Wang, K. Liu, I.A. Roundtree, W. Tempel, Y. Li, Z. Lu, C. He, J. Min, Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain, Nat. Chem. Biol., 10 (2014) 927-929.
[84] X. Cui, J. Meng, M.K. Rao, Y. Chen, Y. Huang, HEPeak: an HMM-based exome peak-finding package for RNA epigenome sequencing data, BMC Genomics, 16 Suppl 4 (2015) S2.
[85] T. Gutschner, M. Hammerle, M. Eissmann, J. Hsu, Y. Kim, G. Hung, A. Revenko, G. Arun, M. Stentrup, M. Gross, M. Zornig, A.R. MacLeod, D.L. Spector, S. Diederichs, The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells, Cancer Res., 73 (2013) 1180-1189.
[86] T.H. Beilharz, D.T. Humphreys, J.L. Clancy, R. Thermann, D.I. Martin, M.W. Hentze, T. Preiss, microRNA-mediated messenger RNA deadenylation contributes to translational repression in mammalian cells, PLoS One, 4 (2009) e6783.
[87] E. Huntzinger, E. Izaurralde, Gene silencing by microRNAs: contributions of translational repression and mRNA decay, Nat Rev Genet, 12 (2011) 99-110.
[88] M. Ha, V.N. Kim, Regulation of microRNA biogenesis, Nat. Rev. Mol. Cell Biol., 15 (2014) 509-524.
[89] C.R. Alarcón, H. Lee, H. Goodarzi, N. Halberg, S.F. Tavazoie, N6-methyladenosine marks primary microRNAs for processing, Nature, 519 (2015) 482-485.
[90] J. Ge, Y.T. Yu, RNA pseudouridylation: new insights into an old modification, Trends Biochem. Sci., 38 (2013) 210-218.
[91] H. Adachi, Y.T. Yu, Insight into the mechanisms and functions of spliceosomal snRNA pseudouridylation, World J. Biol. Chem., 5 (2014) 398-408.

[92] G. Wu, A.T. Yu, A. Kantartzis, Y.T. Yu, Functions and mechanisms of spliceosomal small nuclear RNA pseudouridylation, Wiley Interdiscip Rev RNA, 2 (2011) 571-581.
[93] T. Hamma, A.R. Ferre-D'Amare, Pseudouridine synthases, Chem. Biol., 13 (2006) 1125-1135.
[94] Y. Bykhovskaya, K. Casas, E. Mengesha, A. Inbal, N. Fischel-Ghodsian, Missense mutation in pseudouridine synthase 1 (PUS1) causes mitochondrial myopathy and sideroblastic anemia (MLASA), Am. J. Hum. Genet., 74 (2004) 1303-1308.
[95] N.S. Heiss, S.W. Knight, T.J. Vulliamy, S.M. Klauck, S. Wiemann, P.J. Mason, A. Poustka, I. Dokal, X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions, Nat. Genet., 19 (1998) 32-38.
[96] D.R. Davis, Stabilization of RNA stacking by pseudouridine, Nucleic Acids Res., 23 (1995) 5020-5026.
[97] E. Kierzek, M. Malgowska, J. Lisowiec, D.H. Turner, Z. Gdaniec, R. Kierzek, The contribution of pseudouridine to stabilities and structure of RNAs, Nucleic Acids Res., 42 (2014) 3492-3501.
[98] J. Karijolich, Y.T. Yu, Converting nonsense codons into sense codons by targeted pseudouridylation, Nature, 474 (2011) 395-398.
[99] S.M. Colley, P.J. Leedman, Steroid Receptor RNA Activator - A nuclear receptor coregulator with multiple partners: Insights and challenges, Biochimie, 93 (2011) 1966-1972.
[100] C. Cooper, D. Vincett, Y. Yan, M.K. Hamedani, Y. Myal, E. Leygue, Steroid Receptor RNA Activator bi-faceted genetic system: Heads or Tails?, Biochimie, 93 (2011) 1973-1980.
[101] R.B. Lanz, N.J. McKenna, S.A. Onate, U. Albrecht, J. Wong, S.Y. Tsai, M.J. Tsai, B.W. O'Malley, A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex, Cell, 97 (1999) 17-27.
[102] I.V. Novikova, S.P. Hennelly, K.Y. Sanbonmatsu, Structural architecture of the human long non-coding RNA, steroid receptor RNA activator, Nucleic Acids Res., 40 (2012) 5034-5051.
[103] X. Zhao, J.R. Patton, S.L. Davis, B. Florence, S.J. Ames, R.A. Spanjaard, Regulation of nuclear receptor activity by a pseudouridine synthase through posttranscriptional modification of steroid receptor RNA activator, Mol. Cell, 15 (2004) 549-558.

[104] X. Zhao, J.R. Patton, S.K. Ghosh, N. Fischel-Ghodsian, L. Shen, R.A. Spanjaard, Pus3p- and Pus1p-dependent pseudouridylation of steroid receptor RNA activator controls a functional switch that regulates nuclear receptor signaling, Mol. Endocrinol., 21 (2007) 686-699.
[105] S.K. Ghosh, J.R. Patton, R.A. Spanjaard, A small RNA derived from RNA coactivator SRA blocks steroid receptor signaling via inhibition of Pus1p-mediated pseudouridylation of SRA: evidence of a novel RNA binding domain in the N-terminus of steroid receptors, Biochemistry, 51 (2012) 8163-8172.
[106] M.S. Paul, B.L. Bass, Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA, EMBO J., 17 (1998) 1120-1127.
[107] A. Athanasiadis, A. Rich, S. Maas, Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome, PLoS Biol., 2 (2004) e391.
[108] R.W. Wagner, J.E. Smith, B.S. Cooperman, K. Nishikura, 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. U. S. A., 86 (1989) 2647-2651.
[109] M. Higuchi, S. Maas, F.N. Single, J. Hartner, A. Rozov, N. Burnashev, D. Feldmeyer, R. Sprengel, P.H. Seeburg, Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2, Nature, 406 (2000) 78-81.
[110] N.M. Mannion, S.M. Greenwood, R. Young, S. Cox, J. Brindle, D. Read, C. Nellaker, C. Vesely, C.P. Ponting, P.J. McLaughlin, M.F. Jantsch, J. Dorin, I.R. Adams, A.D. Scadden, M. Ohman, L.P. Keegan, M.A. O'Connell, The RNA-editing enzyme ADAR1 controls innate immune responses to RNA, Cell Reports, 9 (2014) 1482-1494.
[111] B.J. Liddicoat, R. Piskol, A.M. Chalk, G. Ramaswami, M. Higuchi, J.C. Hartner, J.B. Li, P.H. Seeburg, C.R. Walkley, RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself, Science, 349 (2015) 1115-1120.
[112] Y. Kawahara, B. Zinshteyn, T.P. Chendrimada, R. Shiekhattar, K. Nishikura, RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex, EMBO Rep, 8 (2007) 763-769.
[113] Y. Pinto, H.Y. Cohen, E.Y. Levanon, Mammalian conserved ADAR targets comprise only a small fragment of the human editosome, Genome Biol., 15 (2014) R5.
[114] L. Bazak, A. Haviv, M. Barak, J. Jacob-Hirsch, P. Deng, R. Zhang, F.J. Isaacs, G. Rechavi, J.B. Li, E. Eisenberg, E.Y. Levanon, A-to-I RNA editing occurs at over a hundred

million genomic sites, located in a majority of human genes, Genome Res., 24 (2014) 365-376.
[115] H.T. Porath, S. Carmi, E.Y. Levanon, A genome-wide map of hyper-edited RNA reveals numerous new sites, Nat Commun, 5 (2014) 4726.
[116] A.M. Kiran, J.J. O'Mahony, K. Sanjeev, P.V. Baranov, Darned in 2013: inclusion of model organisms and linking with Wikipedia, Nucleic Acids Res., 41 (2013) D258-261.
[117] G. Ramaswami, J.B. Li, RADAR: a rigorously annotated database of A-to-I RNA editing, Nucleic Acids Res., 42 (2014) D109-113.
[118] D.J. Luciano, H. Mirsky, N.J. Vendetti, S. Maas, RNA editing of a miRNA precursor, RNA, 10 (2004) 1174-1177.
[119] S. Alon, E. Mor, F. Vigneault, G.M. Church, F. Locatelli, F. Galeano, A. Gallo, N. Shomron, E. Eisenberg, Systematic identification of edited microRNAs in the human brain, Genome Res., 22 (2012) 1533-1540.
[120] M.J. Blow, R.J. Grocock, S. van Dongen, A.J. Enright, E. Dicks, P.A. Futreal, R. Wooster, M.R. Stratton, RNA editing of human microRNAs, Genome Biol., 7 (2006) R27.
[121] T. Chen, J.F. Xiang, S. Zhu, S. Chen, Q.F. Yin, X.O. Zhang, J. Zhang, H. Feng, R. Dong, X.J. Li, L. Yang, L.L. Chen, ADAR1 is required for differentiation and neural induction by regulating microRNA processing in a catalytically independent manner, Cell Res., 25 (2015) 459-476.
[122] Y. Ekdahl, H.S. Farahani, M. Behm, J. Lagergren, M. Ohman, A-to-I editing of microRNAs in the mammalian brain increases during development, Genome Res., 22 (2012) 1477-1487.
[123] Y. Kawahara, M. Megraw, E. Kreider, H. Iizasa, L. Valente, A.G. Hatzigeorgiou, K. Nishikura, Frequency and fate of microRNA editing in human brain, Nucleic Acids Res., 36 (2008) 5270-5280.
[124] G. Chawla, N.S. Sokol, ADAR mediates differential expression of polycistronic microRNAs, Nucleic Acids Res., 42 (2014) 5245-5255.
[125] J. Garcia-Lopez, D. Hourcade Jde, J. Del Mazo, Reprogramming of microRNAs by adenosine-to-inosine editing and the selective elimination of edited microRNA precursors in mouse oocytes and preimplantation embryos, Nucleic Acids Res., 41 (2013) 5483-5493.

[126] S. Tomaselli, F. Galeano, S. Alon, S. Raho, S. Galardi, V.A. Polito, C. Presutti, S. Vincenti, E. Eisenberg, F. Locatelli, A. Gallo, Modulation of microRNA editing, expression and processing by ADAR2 deaminase in glioblastoma, Genome Biol., 16 (2015) 5.
[127] C. Vesely, S. Tauber, F.J. Sedlazeck, M. Tajaddod, A. von Haeseler, M.F. Jantsch, ADAR2 induces reproducible changes in sequence and abundance of mature microRNAs in the mouse brain, Nucleic Acids Res., 42 (2014) 12155-12168.
[128] S. Alon, M. Erew, E. Eisenberg, DREAM: a webserver for the identification of editing sites in mature miRNAs using deep sequencing data, Bioinformatics, (2015).
[129] A. Lagana, A. Paone, D. Veneziano, L. Cascione, P. Gasparini, S. Carasi, F. Russo, G. Nigita, V. Macca, R. Giugno, A. Pulvirenti, D. Shasha, A. Ferro, C.M. Croce, miR-EdiTar: a database of predicted A-to-I edited miRNA target sites, Bioinformatics, 28 (2012) 3166-3168.
[130] W. Yang, T.P. Chendrimada, Q. Wang, M. Higuchi, P.H. Seeburg, R. Shiekhattar, K. Nishikura, Modulation of microRNA processing and expression through RNA editing by ADAR deaminases, Nat. Struct. Mol. Biol., 13 (2006) 13-21.
[131] H. Kume, K. Hino, J. Galipon, K. Ui-Tei, A-to-I editing in the miRNA seed region regulates target mRNA selection and silencing efficiency, Nucleic Acids Res., 42 (2014) 10050-10060.
[132] Y. Kawahara, B. Zinshteyn, P. Sethupathy, H. Iizasa, A.G. Hatzigeorgiou, K. Nishikura, Redirection of silencing targets by adenosine-to-inosine editing of miRNAs, Science, 315 (2007) 1137-1140.
[133] Y. Choudhury, F.C. Tay, D.H. Lam, E. Sandanaraj, C. Tang, B.T. Ang, S. Wang, Attenuated adenosine-to-inosine editing of microRNA-376a* promotes invasiveness of glioblastoma cells, J. Clin. Invest., 122 (2012) 4059-4076.
[134] E. Shoshan, A.K. Mobley, R.R. Braeuer, T. Kamiya, L. Huang, M.E. Vasquez, A. Salameh, H.J. Lee, S.J. Kim, C. Ivan, G. Velazquez-Torres, K.M. Nip, K. Zhu, D. Brooks, S.J. Jones, I. Birol, M. Mosqueda, Y.Y. Wen, A.K. Eterovic, A.K. Sood, P. Hwu, J.E. Gershenwald, A.G. Robertson, G.A. Calin, G. Markel, I.J. Fidler, M. Bar-Eli, Reduced adenosine-to-inosine miR-455-5p editing promotes melanoma growth and metastasis, Nat. Cell Biol., 17 (2015) 311-321.

Figure legends:

Figure 1: Next generation sequencing (NGS) techniques to detect post-transcriptional RNA modifications. A) Detection of m^5C is achieved by either bsRNA-seq, Aza-IP or the miCLIP technique. bsRNA-seq: Sodium-bisulfite ($NaHSO_3$) treatment converts unmethylated cytosine into uracil whereas methylated cytosines are resistant to conversion. Aza-IP: The cytidine analog 5-Azacytidine (5-Aza-C) is incorporated into RNA. Parallel overexpression of V5-tagged MTases allows the immunoprecipitation of covalently bound 5-Aza-C containing RNA. miCLIP: Overexpression of a myc-tagged NSUN2 mutant (C271A) protein results in the irreversible binding of its RNA targets, allowing identification of methylated RNA fragments through immunoprecipitation. Reverse transcription stalls at the binding site, allowing site-specific identification of the modified cytosine. **B)** MeRIP-seq: m^6A sites are detected by immunoprecipitation using an anti-m^6A antibody. Specific sites are estimated by peak analysis, where the adenine at the center is considered to be the m^6A site. **C)** Ψ-seq: CMC (or CMCT) treatment of RNA results in a stable Ψ-adduct, which inhibits reverse transcription allowing precise detection of the Ψ site by comparison to an untreated library. CeU-seq: Using a CMC derivative allows for the subsequent addition of a biotin label to the Ψ-adduct by click chemistry. This enables enrichment of Ψ-containing transcripts through streptavidin immunoprecipitation, resulting in increased sensitivity. **D)** Sequence comparison between cDNA and DNA sequences from the same sample allows the detection of A-to-G mutations as likely editing sites. ICE-seq: RNA is treated with acrylonitrile to form inosine-acrylonitrile adducts, which results in stalling of the reverse transcriptase. Sequence comparison with an untreated library allows the detection of inosine sites by identifying reads 'erased' from the treated library.

$NaHSO_3$ - sodium bisulfite, 5-Aza-C - 5-azacytidine, C* - 5-azacytidine, m^5C* - 5-methyl-5-azacytidine, MTase-V5 - V5 epitope-tagged RNA methyltransferase, IP - immunoprecipitation, C271A-myc - myc-tagged NSUN2 carrying a C271A mutation, CMC - N-cyclohexyl-N_9-(2-morpholinoethyl)-carbodiimide metho-p-toluenesulphonate, CMCT - 1-cyclohexyl-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate, Ψ - pseudouridine, I* - inosine-acrylonitrile adduct

Figure 2: The proposed functions of post-transcriptional modifications in long noncoding RNAs. **A)** The P8 helix (black) of the lncRNA XIST contains five cytosines that are located at the surface of the RNA. Methylation inhibits binding of XIST to the PRC2 complex. **B)** In MALAT1 an adenine is located in a stem-loop structure opposite a uridine stretch. Formation of m^6A is predicted to result in a conformational change, allowing the enhanced binding of HNRNPC to the uridine region of MALAT1. **C)** Pseudouridylation of the appropriate sites in the Steroid Receptor RNA Activator (SRA) is essential for binding to the N-terminal domain of nuclear receptors and transcription activation. Hyper-pseudouridylation of SRA triggers a switch in function from coactivator to corepressor, suppressing transcription activation.

NR - nuclear receptor, Ψ - pseudouridine, NTD - N-terminal domain, DBD - DNA-binding domain

Figure 3: The proposed function of post-transcriptional RNA modifications in miRNA biogenesis. A) m^6A methylation at the base of the pri-miRNA stem-loop is proposed to increase Drosha processing. This was confirmed by METTL3 depletion which results in decreased pri-miRNA processing and consequently reduced accumulation of mature miRNA. **B)** A-to-I editing effects on miRNA biogenesis are specific to each pri-miRNA transcript. Editing can inhibit either Drosha or Dicer processing, but may also enhance Drosha processing of the pri-miRNA. Editing sites present within the mature miRNA can lead to the targeting of a distinct alternative set of transcripts.

Table 1: Summary of epitranscriptomic marks within lncRNAs

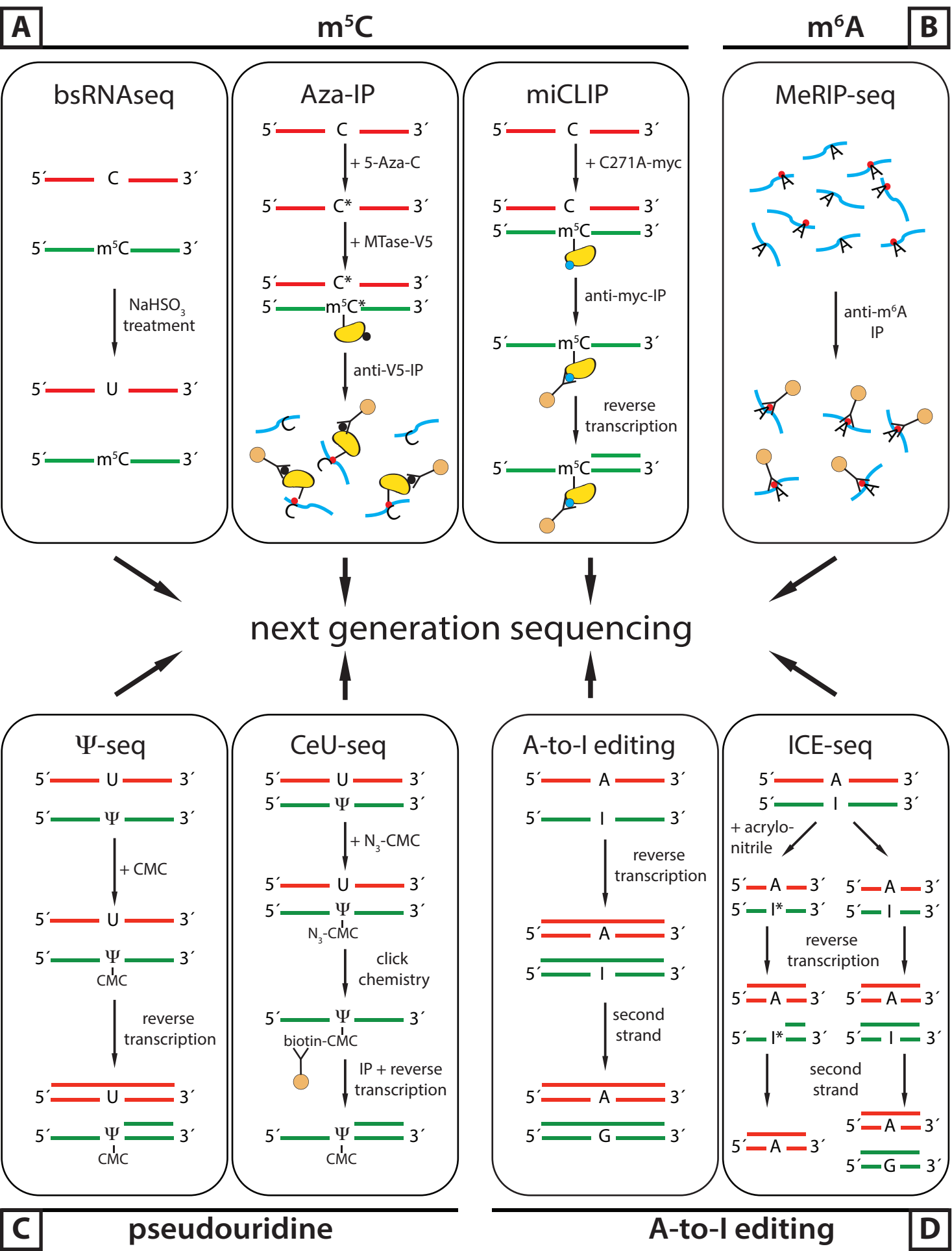
Modification type	Study/database	Total # of modifications[$]	# of modifications in lncRNAs[^]		# of individual lncRNA transcripts[^]	
			strand-specific[$]	***strand-unspecific***[$]	***strand-specific***[$]	***strand-unspecific***[$]
m^5C	*Total unique sites*	*31,175*	*3,134*	*9,965*	*755*	*1,072*
	Khoddami & Cairns (2013)	20,553	1,580	7,726	38	60
	Hussain *et al.* (2013)	1,084	110	121	41	45
	Squires *et al.* (2012)	10,490	1,544	2,234	711	1,010
m^6A	*Total unique sites*	*24,449*	*6,880*	*15,357*	*6,167*	*12,348*
	Dominissini *et al.* (2012)	25,776	7,397	16,111	6,165	12,347
	Meyer *et al.* (2012)	4,341	57	352	20	144
Ψ	*Total unique sites*	*1,805*	*69*	*162*	*62*	*150*
	Li *et al.* (2015)	1,489	58	140	54	132
	Schwartz *et al.* (2014)	396	15	36	11	30
	Carlile *et al.* (2014)	8	4	7	3	5
A-to-I editing	*Total unique sites*	*2,603,542*	*222,481*	*333,394*	*6,626*	*8,832*
	DARNED	259,654	23,574	32,879	1,833	2,608
	RADAR	2,576,289	218,793	327,817	6,376	8,546
	Sakurai *et al.* (2014)	20,482	3,050	4,926	400	635
	Peng *et al.* (2012)	22,686	4,425	6,241	846	1,310

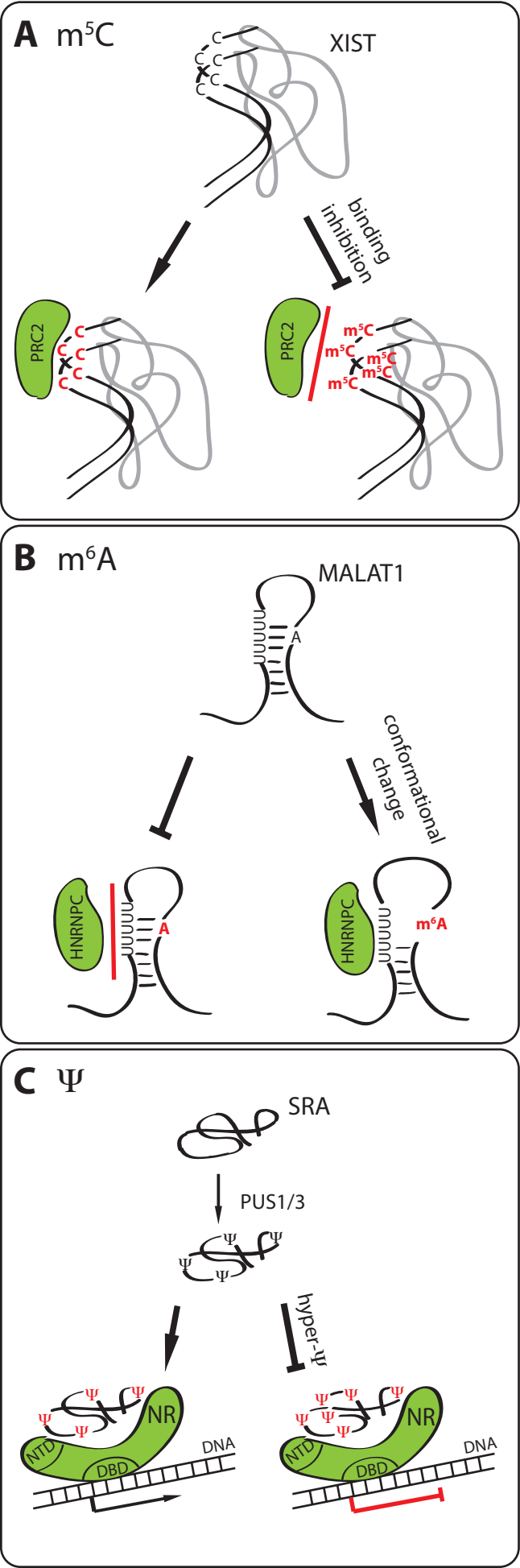
[$]Genomic coordinates of all reported modifications were transformed to coordinates within the current genome assembly version GRCh38 (ensembl.org). Strand-specific and -unspecific overlap of reported modifications with annotated lncRNAs was determined and counted, both on the level of individual modifications and on a per-gene level. Numbers may be different from those reported in the original publication, due to changes (additions, deletions etc.) across different genome assembly versions.

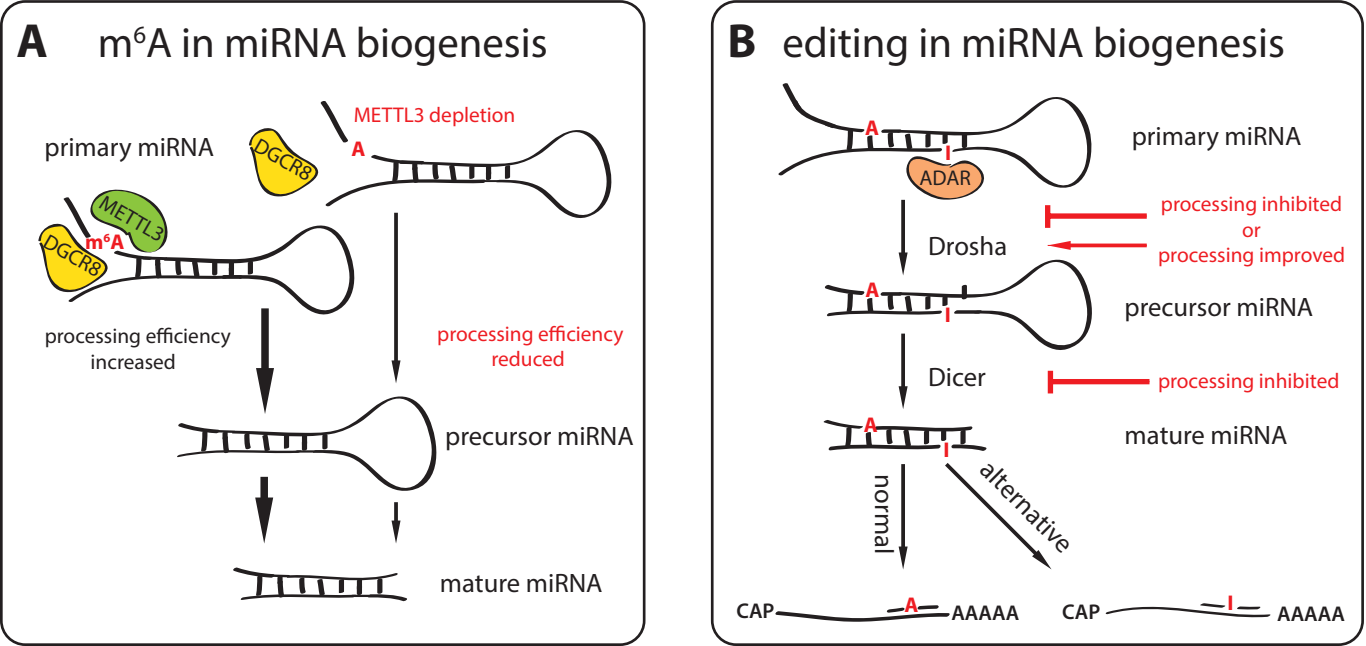
^ Reported modification sites in lncRNAs were determined from the current Ensembl human gene annotation version GRCh38.80 with lncRNAs and pseudogenes. lncRNAs are classified according to http://vega.sanger.ac.uk/info/about/gene_and_transcript_types.html.

Highlights

- RNA modifications are chemically diverse and implement epitranscriptomic regulation
- Technological advances in next generation sequencing have propelled the transcriptome-wide mapping of RNA modifications
- Emerging information on the role of post-transcriptional modifications in long noncoding RNA function is reviewed here


A
m5C
m6A
B
bsRNAseq
5´ C 3´
5´ m5C 3´
NaHSO3 treatment
5´ U 3´
5´ m5C 3´
Aza-IP
5´ C 3´
+ 5-Aza-C
5´ C* 3´
+ MTase-V5
5´ C* 3´
5´ m5C* 3´
anti-V5-IP
miCLIP
5´ C 3´
+ C271A-myc
5´ C 3´
5´ m5C 3´
anti-myc-IP
5´ m5C 3´
reverse transcription
5´ m5C 3´
MeRIP-seq
anti-m6A IP
next generation sequencing
Ψ-seq
5´ U 3´
5´ Ψ 3´
+ CMC
5´ U 3´
5´ Ψ 3´
CMC
reverse transcription
5´ U 3´
5´ Ψ 3´
CMC
CeU-seq
5´ U 3´
5´ Ψ 3´
+ N3-CMC
5´ U 3´
5´ Ψ 3´
N3-CMC
click chemistry
5´ Ψ 3´
biotin-CMC
IP + reverse transcription
5´ Ψ 3´
CMC
A-to-I editing
5´ A 3´
5´ I 3´
reverse transcription
5´ A 3´
5´ I 3´
second strand
5´ A 3´
5´ G 3´
ICE-seq
5´ A 3´
5´ I 3´
+ acrylo-nitrile
5´ A 3´
5´ I* 3´
5´ A 3´
5´ I 3´
reverse transcription
5´ A 3´
5´ I* 3´
5´ A 3´
5´ I 3´
second strand
5´ A 3´
5´ A 3´
5´ G 3´
C
pseudouridine
A-to-I editing
D


A
m5C
XIST
C
C
C
C
binding
inhibition
PRC2
C
C
C
C
PRC2
m5C
m5C
m5C
m5C
m5C
B
m6A
MALAT1
A
conformational
change
HNRNPC
A
HNRNPC
m6A
C
Ψ
SRA
PUS1/3
Ψ
Ψ
Ψ
Ψ
hyper-Ψ
Ψ
Ψ
Ψ
Ψ
Ψ
Ψ
NR
NTD
DBD
DNA
Ψ
Ψ
Ψ
Ψ
Ψ
Ψ
NR
NTD
DBD
DNA


A
m6A in miRNA biogenesis
primary miRNA
DGCR8
METTL3 depletion
A
METTL3
m6A
DGCR8
processing efficiency increased
processing efficiency reduced
precursor miRNA
mature miRNA
B
editing in miRNA biogenesis
A
I
ADAR
primary miRNA
Drosha
processing inhibited or processing improved
precursor miRNA
Dicer
processing inhibited
mature miRNA
normal
alternative
CAP
A
AAAAA
CAP
I
AAAAA