RNA modification in gene expression control:

m6A modification of non-coding RNA and the control of mammalian gene expression.

Heather Coker [1]*, Guifeng Wei [1], Neil Brockdorff [1]

[1] Department of Biochemistry, University of Oxford, South Parks Road, Oxford, UK. OX1 3QU

*Corresponding author: **heather.coker@bioch.ox.ac.uk**.

Abstract

The biology of non-coding RNA (ncRNA) and the regulation of mammalian gene expression is a rapidly expanding field. In this review, we consider how recent advances in technology, enabling the precise mapping of modifications to RNA transcripts, has provided new opportunities to dissect post-transcriptional gene regulation. With this has come the realisation that in the absence of translation, the modification of ncRNAs may play a fundamental role in their regulation, protein interactome and subsequent downstream effector functions. We focus upon modification of RNA by N^6-methyladenosine (m6A); its readers, writers and erasers, before considering the differing role of m6A modified lncRNAs MALAT1 and Xist.

Keywords

ncRNA; Xist; MALAT1; m6A

Introduction

An extraordinary number of non protein-coding RNAs (ncRNA) exist for which there is no known function, but advancing high-throughput technology now enables us to examine novel ncRNAs in greater depth; from short ncRNAs such snRNAs and snoRNAs to highly processed long ncRNAs (lncRNA) (reviewed by [1] [2]). Already, we have glimpses of the importance of ncRNA in the context of highly specialised multicellular organisms: LncRNA, a ncRNA of 200 nucleotides or more, can act at multiple levels to regulate gene expression (reviewed by [3]): They can act as transcriptional regulators either *in cis* or *in trans* to a gene, with even the act of transcription itself able to modulate transcription of a neighbouring gene [4]. LncRNA can regulate the accessibility of RNA binding proteins, restricting their ability to interact with other transcripts or directly affect accessibility to a gene in order to influence transcription. LncRNAs can activate transcription, as in the case of the lncKdm2b acting to recruit the NURF chromatin remodeller complex to stimulate expression of the transcription factor Zfp292 [5]. In contrast, a complex cascade of events leading to chromosome wide gene repression is mediated by the lncRNA Xist (discussed in detail later in this review). Not only can lncRNAs act as post-transcriptional regulators of gene expression, but recent research has demonstrated that lncRNAs such as NEAT1 are essential components of phase separated nuclear paraspeckles, providing a link between nuclear organisation and gene regulation by a lncRNA [6-9].

Many different modifications of RNA have been identified, but with recent advances in the ability to map these modifications, attention has turned to the effects they may have on ncRNA beyond rRNA and tRNA [10, 11] [12]. This is most evident in the field of N^6-methyladenosine (m6A) modification of RNA. Development of an antibody-based m6A mapping technique, has dramatically fuelled research over the past six years into the post-transcriptional consequences of m6A modification of mouse and human RNA [13-16]). Indeed, the identification of m6A 'writer', 'reader' and 'eraser' proteins, has brought about an increasing realisation that in modifying RNA to generate what could be viewed as an 'epi-transcriptomic language', a new layer of regulation for all RNA, including ncRNA, had been identified. In this review, we give a brief overview of m6A modification of RNA and the potential for

m6A on ncRNA to regulate gene expression, considering in detail the chromatin associated ncRNAs MALAT1 and Xist.

N^6-methyladenosine and its writers, readers and erasers

The methylation of the N^6 position of adenosine (m6A) in mammalian RNA was identified several decades ago, and characterised as a highly abundant modification of mRNA [17, 18] [19] [20], [21]). It wasn't until 2012 however, that two groups mapped the transcriptomic distribution of mouse and human m6A, utilising an m6A-specific antibody approach, coupled with high throughput sequencing [13, 14]. This work demonstrated just how prevalent the m6A mark was, with Dominissini et al. identifying more than 12,000 m6A sites present on more than 7000 human genes from both a hepatocarcinoma cell line and normal brain tissue. Comparison with RNA derived from mouse liver, indicated conservation in many instances between the human and mouse 'RNA methylomes' [13]. These observations were supported when Meyer at al. identified m6A on transcripts from close to 8000 genes and more than 300 non-coding genes from mouse brain [14]. Further work has identified m6A on the transcripts of Saccharomyces cerevisiae, Arabidopsis, Drosophila and Zebrafish [22], [23], [24, 25].

m6A is thought to be placed co-transcriptionally, with the majority of m6A already evident on nascent pre-mRNA [26]. It is deposited on adenosines within the consensus motif RRACH (in which R: A or G and H: A, C or U), and whilst the distribution of m6A varies between transcripts from different genes (with multiple peaks associated with certain transcripts), there are certain trends in the distribution of m6A [13], [14]: In general, transcripts exhibit a trend of enrichment of m6A towards the stop codon and 3' UTR, hypothesized to contribute towards the control of transcript stability and translation [14]. Furthermore, Dominissini et al. also identified that m6A in exonic regions was preferentially found in longer exons of 400 nucleotides or more [13]. Mapping of m6A has been further refined with the advent of m6A-miCLIP and PA-m6A-seq in which single nucleotide resolution of m6A sites has been achieved by UV crosslinking the m6A specific antibody to RNA, before adopting an iCLIP based approach [15, 16, 27, 28]. These techniques still suffer from a loss of resolution at positions in which there are multiple, sequential m6A sites.

This has been addressed by the development of 'site-specific cleavage and radioactive-labelling followed by ligation-assisted extraction and thin-layer chromatography' (SCARLET), which has been of particular use in the elucidation of single nucleotide distribution of m6A within clustered consensus motifs, and has also importantly enabled the quantification of m6A abundance at a distinct site within the transcript population [29].

m6A writers

The specificity of m6A at RRACH consensus motifs is conferred by components of a multi-subunit m6A methyltransferase complex, which localises to the nucleus and acts on nascent RNA [26] (Figure 1). This mammalian complex includes the methyltransferase components METTL3 and METTL4, which interact to form a heterodimeric catalytic core, capable in isolation of methylating RNA *in vitro* [29] [30]. *In vivo* however, other key components of the complex are necessary for its activity: WTAP and VIRMA have been hypothesized to confer a regulatory role to the complex [29] [31] [32] [33]. WTAP has been shown to assist in the correct nuclear localisation of the methyltransferase complex and the absence of WTAP or the depletion of VIRMA has been shown to significantly deplete m6A from the majority of RNA transcripts [32, 33].

In vivo, the RNA binding protein RBM15, and its homolog RBM15b, interact with METTL3 in a WTAP dependent fashion and mass spectrophotometry has confirmed RBM15 to be a key component of the methyltransferase complex in Drosophila [31], [34], [24]. In the absence of RBM15 / RBM15b, m6A deposition is perturbed, and it has been suggested that the RRM domains of RBM15 / RBM15b preferentially bind to U-rich regions of RNA, recruiting the m6A complex to sites proximal to the m6A consensus motifs [34]. The most recent component of the m6A methylation complex to be identified is ZC3H13 [35, 36]: In both mouse ES cells and in Drosophila, this zinc finger protein was shown to be required for efficient, functional deposition of m6A and successful localisation of the complex to the nucleus. In Drospohila, ZC3H13 has been shown to act as a bridging component between Fl(2)d (WTAP) and Nito (RBM15) in the methyltransferase complex [35]. Additionally CBLL1, a E3 ubiquitin ligase, has been identified on multiple occasions as a potential member of the m6A methyltransferase complex, although its role has yet to be clearly defined

[31, 36, 37]. Current open questions include the exact physical relationship of the m6A methyltransferase core components, given the lack of structural data for an intact complex, indeed it will be interesting to observe whether there is tissue specificity to m6A methyltransferase co-factors, that may correspond to defined roles in the whole organism.

m6A erasers

Dynamic regulation of m6A was suggested by the identification of two mammalian m6A demethylase proteins; fat-mass and obesity-associated protein (FTO), and Alk B homolog 5 (ALKBH5) [38, 39] (Figure 1). The nuclear protein FTO demethylates m6A generating *in vitro* the intermediate N6-hydroxymethyladenosine (hm^6A), which it then oxidises further to N6-formyladenosine (f^6A). The detection of these intermediates in mouse and human mRNA means that it is possible *in vivo* that they themselves perform an as yet unknown function [38, 40]. The biological relevance of FTO with respect to m6A demethylation *in vivo* is however unclear: In humans, misregulation of FTO has been linked to obesity [41-43] and the cellular loss of FTO alters pre-mRNA splicing patterns, but FTO deficient mice exhibit minor effects on m6A [44] [45]. Further studies have suggested that FTO preferentially demethylates m^6Am, found adjacent to the 7-methylguanosine cap on mRNA, thereby influencing mRNA stability in an m6A independent fashion [46].

The nuclear RNA demethylase ALKBH5 is a better candidate for m6A demethylation *in vivo*. Loss of ALKBH5 has recently been demonstrated to impact upon the prevalence of m6A and the splicing and degradation of certain transcripts in male germ cells, leading to a concomitant loss of fertility in male mice [39] [47]. However, data supporting a role for ALKBH5 *in vivo* is limited to certain transcripts in a specific cellular context [48]. On a broader biological level, recent studies of the dynamics of m6A in mouse and human cells have suggested that the modification may in fact be remarkably stable, fuelling debate as to the concept of a dynamically fluctuating population of m6A modified transcripts [26, 49, 50].

m6A readers and gene expression

Given the relatively limited dynamics of m6A and that coding capacity of m6A modified mRNA is unaltered, it is unsurprising that reader proteins are necessary to

confer the downstream effects of m6A [51]. Until recently, the majority of the literature concerning m6A readers has focused on the YTH family of proteins (reviewed by [52]). These readers demonstrate m6A selectivity as a consequence of a recognition domain that has similarity to the eukaryotic YT521-B domain, from which they gain their name [53, 54] [55, 56]. There are multiple, diverse pathways through which m6A has been suggested to affect gene expression instructed by the downstream functional interactions of reader proteins. The specificity of a reader for one subset of m6A containing transcripts versus another is currently unexplained, but is regulated in part by the intercellular and tissue wide distribution of different reader proteins (Figure 1).

YTHDC1 is the only exclusively nuclear m6A reader protein (Figure 1). It has been shown to interact with a subset of m6A containing transcripts and splice factors in order to regulate gene expression and RNA export [57, 58]. Indeed, the direct interaction of m6A bound YTHDC1 with the splice factor SRSF3 has been shown to block binding of SRSF10 to m6A modified RNA, promoting exon inclusion in those select transcripts [57]. There is some debate regarding the extent to which splicing is altered in the absence of m6A, which may be a consequence of differing cellular, environment and model systems [13, 14, 59, 60] [24] [57]. There has also been a suggestion that the interaction of YTHDC1 with the lncRNA Xist, is able to recruit factors leading to silencing of gene expression on the inactive X chromosome, (discussed later in this review) [34]. YTHDC2 is present in both the nucleus and the cytoplasm (Figure 1), and exhibits a very different domain structure to YTHDC1 [52]. YTHDC2 is highly expressed in germ cells, where it is thought to play a dual role; both enhancing the efficiency with which certain m6A transcripts are translated as well as modulating their subsequent degradation. Its absence affects both male and female fertility, but have a specific role in spermatogenesis [61-63]. Most recently, YTHDC2 has been shown to interact directly with the small ribosomal subunit to promote translation of a subset of transcripts and then to degrade them via recruitment of the exoribonuclease XRN1 [64].

The cytoplasmic proteins YTHDF1, 2 and 3 together process mRNA, influence translation and coordinate mRNA degradation after translation: YTHDF1 interacts with the translation initiation factor eIF3 to promote highly efficient cap-dependent

translation, and has been shown to modulate mRNA degradation [65] [66]. YTHDF2 preferentially binds m6A in the 3' UTR, recruiting the CCR4-NOT deadenylase complex and resulting in m6A transcript degradation (Figure 1), and as a consequence, m6A modified transcripts exhibit shorter half-lives than unmodified transcripts [67-69]. RNA transcripts whose half-life is influenced by the presence of m6A include transcription factors themselves, demonstrating the complexity of regulation in an interlinked transcriptome [70]. Furthermore, under conditions of cellular stress, YTHDF2 relocates to the nucleus, binds m6A in the 5'UTR of stress-induced transcripts to promote their cap-independent translation [67, 71]. The role of YTHDF3 appears to overlap with its paralogs YTHDF1 and 2, and it has been suggested that YTHDF3 may work with these other readers in a cooperative fashion [72] [73].

In addition to the YTH family of readers, several other proteins show a preferential binding to m6A (reviewed by [74]) (Figure 1): The nuclear protein HNRNPA2B1 has been shown to influence both m6A dependent alternative splicing and also the recruitment of the miRNA processing complex DGCR8 to a subset of precursor miRNAs, containing m6A in order to facilitate their progression to mature miRNAs [75, 76]. Furthermore, it has been suggested that miRNAs are in turn themselves able to regulate m6A levels, modulating the ability of METTL3 to bind to miRNA target RNAs and place m6A [77]. miRNA target mRNAs have also been shown to frequently exhibit peaks of m6A preceding the miRNA target site [14]. IGF2 binding proteins 1, 2 and 3 contain conventional KH RNA binding domains that selectively encompass an m6A site. Approximately 80% of transcripts bound by IGF2B were found to contain m6A, with the binding motif of IGF2B shown to overlap with the m6A consensus motif [78]. In both cases, further work is needed to fully understand the role of these nonconventional m6A readers.

Whilst the modification of RNA with m6A does not lead to the inevitable loss of Watson and Crick base pairing, it has been shown to destabilise RNA structures in certain circumstances, enabling RNA binding proteins to bind previously inaccessible, regions of RNA. Heterogeneous nuclear ribonucleoprotein C and G (HNRNPC and HNRNPG) are thought to utilise this structural 'm6A switch' mechanism. Indeed, recent structural analysis of HNRNPA2B1 does not support its direct recognition of

m6A, suggesting that it too may function as an m6A-switch (discussed in further detail below) [79].

It has been suggested that m6A might be broadly used to post-transcriptionally fine tune gene expression in contrasting but similar fashion to miRNA; m6A acting as a potential counter-balance to effectively up-regulate gene expression, boosting translation when necessary, and working to buffer the miRNA system controlling gene downregulation [80]. This could potentially enable the dynamic transcriptomic response necessary to respond to environmental or developmental stimuli. Indeed, whilst it can be difficult to tease apart the primary from the widespread secondary effects on gene expression of the removal of key cellular pathways such as the depletion of METTL3, there is a general consensus that m6A is essential in a system that requires plasticity, such as the differentiation of naïve ES cells, cellular reprogramming and the observation that the *in vivo* knockout of Mettl3 is embryonic lethal [81] [82] [77]. It has also proved impossible to bring viable METTL14 null mice to term, with METTL14 activity seeming to be critical for early embryogenesis [83]. In scenarios such as these, it is likely that the effects of losing m6A on gene expression must be a combination of misregulated miRNA, splicing, stability, degradation and translation. For this reason, it is becoming increasingly obvious that perturbation of a writer, or indeed a reader can be frustratingly difficult to accurately interpret and that to progress further in our understanding we need to carefully dissect the effects of m6A on individual RNA, ideally by modulating the RNA itself rather that altering the broader cellular environment. Perturbing m6A and altering mRNA half-life or translational efficiency has a clear effect on gene expression, but understanding the differing roles of m6A on non-coding RNA in the mammalian cell is less clear. There are currently only a handful of non-coding RNA's that we understand with any certainty. The first of these that we will explore is the lncRNA MALAT1, followed by the lncRNA Xist, for which there is arguably the most comprehensive *in vivo* data.

MALAT1

Metastasis-associated lung adenocarcinoma transcript-1 (MALAT1) is a 7 - 9 kb, evolutionarily conserved lncRNA whose mutation or upregulation has been consistently associated with the development of cancer and subsequent metastasis [84] (reviewed by [85]). Processing of the MALAT1 transcript leads to both a nuclear lncRNA and a cytoplasmic tRNA like form [86]. In the nucleus, MALAT1 is thought to associate with active regions of chromatin and has been suggested to interact with and potentially regulate the activity of splice factors, with an impact on alternative splicing [87-89]. In the absence of disease however, deletion of MALAT1 does not appear to dramatically alter gene expression or the viability of mice deficient in MALAT1 [90-92] [85]. Interestingly, it has been suggested that the lack of phenotype observed with MALAT1 deficiency may be a consequence of compensation by the lncRNA NEAT1, known to facilitate the formation of phase separated nuclear paraspeckles [7, 8, 87].

MALAT1 is a highly abundant, highly m6A modified transcript, containing multiple m6A sites modified in a high proportion of transcripts from multiple cell lines [13, 14, 93]. SCARLET technology has confirmed m6A at four of the seven MALAT1 m6A consensus motifs (Figure 2) [93]. Two of these sites were identified as the first examples of an ‘m6A switch’: In contrast to the m6A specific readers discussed earlier in this review, the m6A structural switch has been hypothesized to be an alternative mode of action for m6A, in which modification of an RNA transcript increases the likelihood of single strandedness, structurally altering the binding site of an RNA binding protein (RBP) to make its interaction with the transcript more favourable [94].

The first example of an m6A structural switch factor was HNRNPC, a nuclear RBP that preferentially binds single stranded poly-U tracts and has putative roles in RNA processing and maturation [95, 96]. Photoactivatable-ribonucleoside-enhances crosslinking and immunoprecipitation (PAR-CLIP) combined with m6A-seq identified upwards of 39,000 putative m6A ‘switches’ at HNRNPC binding sites in both the introns and exons of coding and non-coding transcripts from human cells [96]. Of these sites, close to 3000 switches showed a decrease in the ability of HNRNPC to bind when m6A was globally reduced by the knockdown of either METTL3 or METTL14 [96]. In these instances it appears that modification of the

RNA with m6A favours the single strandedness of the RNA binding sites required by HNRNPC. *In vitro* studies utilised an RNA hairpin from MALAT1 containing both the m6A target A2577 and a potential poly-U HNRNPC binding site (Figure 2). In its unmethylated state, (or indeed as a consequence of an A2577U mutation) this HNRNPC binding domain was partially inaccessible. When A2577 was modified to m6A however, the stem of the RNA hairpin was destabilized, resulting in a more accessible poly-U tract and an approximate eight-fold increase in HNRNPC binding [93, 96]. This was investigated further using nuclear magnetic resonance (NMR) and Forster resonance energy transfer (FRET) to confirm that the U_5-tract was destabilized by m6A modification of A2577 even though NMR data suggested that the main structure of the hairpin was maintained [97]. Whilst no specific effects of HNRNPC binding to the MALAT1 m6A switch were identified in this study, the binding of HNRNPC genome-wide to transcripts has been shown to affect both their abundance and alternative splicing profile. Further work is needed to determine if the increased binding of HNRNPC to m6A containing MALAT1 directly affects gene expression, particularly transcripts that might directly or indirectly have an effect upon the proliferation and invasiveness of cancerous cells. This will be of particular relevance given that although many MALAT1 m6A sites are evolutionarily conserved, the A2577 site is only found in human, gorilla orangutan and baboon transcripts [93].

HNRNPG is also able to bind more easily to m6A methylated RNA. Unlike the m6A proximal binding of HNRNPC to the 2577 m6A hairpin, the binding of HNRNPG to purine-rich RNA has been shown to overlap directly with the m6A RRACH consensus motif (Figure 2) [98]. Whilst HNRNPG does not directly recognise m6A, it appears that binding accessibility is facilitated by its presence, with more than 13,000 m6A dependent HNRNPG binding sites identified, including m6A at residue A2515 in MALAT1. These targets appear to be largely distinct from those of HNRNPC and can alter the expression and alternative splicing of HNRNPG bound m6A transcripts [96, 98].

Finally, it is interesting that METTL16 has also been characterised as a MALAT1 binder. Homodimeric METTL16 is an RNA methyltransferase known to place m6A on U6 small nuclear RNAs (snRNA), and to regulate the splicing of MAT2A, the

gene that codes for, and consequentially alters, the levels of the methyltransferase substrate S-Adenosyl methionine [99]. METTL16 does not exhibit the same RRACH consensus motif as METTL3 / METTL14, but has been mapped to interact specifically with the 3' triple helix region of MALAT1, and whilst the impact of this specific interaction on gene expression is unclear, METTL16 is essential for embryonic development [100] [101, 102].

Xist mediated silencing of gene expression

Xist (X inactive-specific transcript) is perhaps the most extensively studied mammalian lncRNA. It is a 17.5 kb, capped, spliced and polyadenylated transcript that never leaves the nucleus and functions as the master regulator of X-chromosome inactivation (XCI). XCI is the process by which genes on one of the two X chromosomes of a female mammal are transcriptionally silenced to achieve dosage compensation, ensuring comparable X chromosome gene products are present in both males and females [103-109]. In other organisms, dosage compensation is achieved by the upregulation of transcripts from the single male X chromosome (reviewed in [110]). In Drosophila this process has been shown to be dependent on female-specific alternative splicing of the Sxl transcript. This key event is regulated by m6A directed recruitment of splice factors (reviewed elsewhere in this issue) [24, 59, 60]. Whilst it is tempting to speculate that m6A may also play a major role in mammalian dosage compensation, it is important to consider how different the modes of dosage compensation are in the two systems:

Xist mediates a complex cascade of events leading to heterochromatinisation and subsequent transcriptional silencing of the chosen X chromosome. This can be visualised as the 'Barr body'; the densely stained, compact inactive X chromosome, visible by light microscopy exclusively in the nucleus of female cells [111]. XCI is initiated in the inner cell mass of the developing embryo, when, during random XCI, Xist is upregulated in response to developmental signals, spreading *in cis* throughout the territory of the X chromosome that will be silenced (making mouse embryonic stem (ES) cells an ideal model system in which to study the process). Imprinted XCI occurs in a specific developmental context in mice (reviewed in [112]), but will not be considered here as the role of m6A in imprinted XCI has yet to be investigated

(although WTAP was identified in a biochemical screen as an interactor of Xist during imprinted XCI) [113].

Xist is located within the X-inactivation centre, a complex genomic locus approximately 100 kb in size and comprised of multiple genes related to the regulation of XCI, ensuring X chromosomal counting and the coordination of Xi with embryonic development [114]. Four of these genes generate non-coding transcripts; Xist, Jpx / Enox, Tsix and Xite and of these, only Tsix and Xist are known to be modified by m6A [13, 34] [115]. Tsix is expressed in male ES cells and female ES cells prior to the onset of XCI. It is antisense to, and overlaps with, Xist, repressing Xist expression [114] [116, 117] (reviewed by [118]). Tsix exhibits a clear enrichment of m6A at the 3' end of the transcript, and it will be interesting to see what effect, if any, this has on Xist expression and XCI (Figure 3a) (unpublished data). Xist, in contrast, exhibits three major peaks of m6A enrichment in mouse ES cells (Figure 3b), and multiple potential sites in human [13] [28] [15, 34] [115]. The most prominent peak of mouse Xist m6A spans approximately 200 nucleotides of exon 1, towards the 5' end of the transcript. To date, SCARLET technology has not been applied to Xist, meaning that we do not have information as to the proportion of Xist molecules that are modified at each position.

The complex modular structure of Xist RNA specifically recruits multiple chromatin modifying complexes. This means that even as an inducible, autosomal transgene, Xist demonstrates only slightly reduced silencing efficiency compared to the endogenous locus on the X chromosome [119] [120] [115]. Repetitive elements in Xist RNA, termed the A, F, B, C, D and E-repeats, act as 'landing pads' for protein factors required for downstream silencing (Figure 3b) [121] [122]. These regions are highly conserved across species, albeit in an expanded or collapsed form (reviewed by [123]). We now understand that different regions of Xist confer different properties, with the ability to segregate for example, the regions essential for silencing, from those that affect localisation and those that have more minor roles in silencing. This has resulted in a minimal form of Xist having been characterised; capable of chromosomal silencing and containing the A, F and B-repeats, a fragment of the C-repeat region and the major peak of m6A modification, which falls just after the A-repeats (Figure 3b and c) [119][115].

The A-repeats are key to the silencing capability of Xist, and contain 7.5 copies of a 26-mer annotated repeat (8.5 copies in humans), separated by a U-rich spacer domain [122]. We now know that the silencing capability of the A-repeats results almost exclusively from their recruitment of the key silencing factor SPEN (also known as MINT and SHARP) (Figure 3c) [113, 124-126] [115]. Indeed, a systematic analysis of Xist demonstrated that its silencing capacity largely resulted from the combination of SPEN mediated silencing and polycomb mediated pathways[115]. HNRNPK has been shown to recognise the B-repeat region of Xist, recruiting PCGF3/5 of the Polycomb Repressive Complex 1 (PRC1) which places the repressive chromatin mark H2AK119ub, subsequently recruiting Polycomb Repressive Complex 2 (PRC2), placing repressive H3K27me3 [127] [119, 128-131].

RBM15, the component of the m6A methyltransferase complex suggested to bind proximal to sites of m6A, also localises to the A-repeats [34, 132]. The extensive 5' peak of m6A modification characterised in mouse Xist is found immediately after the A-repeat region and before the F and B repeats (Figure 3b and c). Interestingly, mouse Xist mutants previously defined as 'A-repeat mutants' include this m6A-enriched region, thus deleting the A-repeats would necessarily result in loss of the majority of m6A. This explains the identification of WTAP as a direct interactor of Xist in an A-repeat dependent fashion [113] [122] [125]. Experiments to determine the effect of WTAP knockdown on silencing have generated apparently conflicting results, with one group observing no effect on Xist mediated silencing [113], whilst a genetic screen identified WTAP, RBM15 and VIRMA as important factors for the efficient initiation of Xist mediated silencing [125].

Further work attempted to address this discrepancy and suggested a direct effect of Xist m6A on its ability to mediate transcriptional silencing [34]. It was suggested that RBM15 and its paralog RBM15b exhibited redundancy, such that when both proteins were depleted, there was a dramatic effect upon the efficiency of Xist mediated silencing, with more than 70 % of silencing abolished after depletion of RBM15/15B. Interestingly, iCLIP data suggested that RBM15/15B was found bound proximal to m6A sites in human HEK293 cells in which XCI was fully established, and the majority of Xist RNA were fully processed, mature RNA [34]. In addition to

biochemical data, super-resolution microscopy has also demonstrated that components of the m6A methyltransferase complex are associated with mouse Xist RNA throughout the inactive X domain, not just at the site of transcription, a curious observation given the apparent lack of dynamic change to m6A and the expectation that m6A would need to be accessible to a reader protein [26, 125]. In the study by Patil et al., Xist mediated silencing ability was assayed using a microscopy based technique in which mouse ES cells with successful knockdown of a protein were identified by immunofluorescence and then subsequently analysed using single molecule RNA fluorescence in situ hybridisation (smFISH), scoring the combined number of nuclear and cytoplasmic transcripts before and after Xist induction [34]. Using this method, METTL3 and YTHDC1 knockdown were also shown to result in up to a 95 % loss of Xist mediated silencing, which interestingly could be rescued by tethering YTHDC1 to the 3' of Xist [34]. This raises the question as to how tethering the nuclear reader protein to a distal region of Xist was able to bring about such a transcript specific effect. Furthermore, the inability to generate YTHDC1 null mouse ES cells, and the perturbation of cellular transcripts, including those measured from the X chromosome even prior to the induction of Xist mediated silencing, again highlights the limitations of either acute or chronic disruption of the m6A network to the complex transcriptional balance of the cell [34]. Indeed, evidence suggests that it may be impossible to maintain a cellular system in which m6A is completely removed, rather than depleted, given that Wtap and Mettl3 knockout cells exhibit only *reduced* levels of m6A [81] [115].

An alternative method to analyse Xist mediated silencing exploited single nucleotide polymorphisms from an interspecific mouse ES cell line to carry out allele specific analysis of chromatin-bound transcripts (Chromatin RNA-seq) from Xist containing chromosomes (transgenic or endogenous) [115]. In these experiments, the majority of isolated transcripts were nascent, meaning that analysis of Xist mediated silencing considered only the effects on transcription, rather than any changes to cytoplasmic decay rates. Furthermore, allelic analysis ensured that changes in expression from both the active and inactive X chromosome were taken into account. Using this technique, knockout of Rbm15 or Wtap had no effect upon the ability of an Xist transgene to transcriptionally silence the surrounding chromosome, whilst in the endogenous setting of the X chromosome, Rbm15 and Wtap knockouts demonstrated

a minimal effect upon XCI, far less than that observed for the Spen knockout or A-repeat deletion (in which the adjacent m6A region remained intact) [115]. A further deletion, of endogenous mouse Xist, in which seven of the nine A-repeat proximal m6A sites were removed (mutant 11G), had no effect upon Xist mediated XCI. An additional, staggered deletion, that removed only five of the m6A sites but encroached into the annotated SPEN binding domain of the A-repeats themselves, showed a small silencing defect (mutant 3A) (Figure 3b and c) [115]. Whilst a low level of m6A may have remained on Xist both in the deletion mutants and in the Wtap and Rbm15 knockouts, the effect of m6A upon Xist mediated gene expression was minimal.

It is likely that the differences observed in these studies are a consequence of the differing methods used to assay transcriptional silencing, and highlight the importance of careful analysis, in particular when perturbing cellular systems with the potential for extensive secondary effects. It may be that acute as opposed to chronic perturbation of the system contributes to the differing results observed here, as it would seem, does the chromosomal context in which Xist mediated silencing takes place. What can be concluded, is that Xist mediated gene silencing is not heavily dependent upon m6A, although there is the potential that under certain circumstances there may be a more nuanced effect. It will be interesting in the future to probe more deeply the m6A distribution across the population of Xist molecules and to understand whether the half-life or stability of Xist is altered by m6A modification. It is also interesting to contemplate why certain components of the m6A machinery appear to be prolific and stable binders of Xist.

Conclusion

For more than forty years we have known that like DNA, RNA is chemically modified, with evidence of RNA modifications identified from viruses to Arabidopsis, mouse and man [17-21, 133]. Characterisation of highly abundant modified tRNA and rRNA first informed us of the plethora of structural and functional roles for modified RNA (reviewed by [134, 135]. The advent of new nucleotide mapping techniques has facilitated the selective enrichment of modifications on less abundant forms of RNA and enabled single nucleotide resolution of m6A on coding and non-coding RNA transcripts.

Perturbation of the m6A methyltransferase complex and identification of the m6A reader proteins has provided evidence that in mRNA and non-coding RNA, modification is likely to play a role in the stability and function of transcripts. Indeed, with the dynamics of m6A seemingly fairly limited, understanding how different m6A reader proteins recognise and select their target transcripts to modulate the specific downstream effects of m6A is a key open question. Clear specificity for m6A, conferred by the YTH domain, appears easier to dissect than more amorphous m6A 'switch' targets, in which m6A alters RNA structure such that it facilitates binding of certain RNA binding proteins. Even so, there is currently no understanding of how YTHDC1, as the predominant nuclear m6A reader, is able to distinguish between different sets of co-transcriptionally modified RNA transcripts. When YTHDC1 is tethered to Xist for example, in a non sequence-specific manner, what regulates its ability to mediate downstream transcriptional silencing in that context, when in another context, YTHDC1 might be expected to recruit spliceosome related proteins?

The effects of m6A upon gene expression are complex and, at a cellular level, difficult to dissect. Removing the m6A methyltransferase complex is challenging, with knockout often resulting in cell lethality. For those proteins in which knockout can be achieved, there is seemingly a hypomorphic effect, with reduced levels of m6A remaining across transcripts, confounding the interpretation of results [81] [115]. The effects of acute knockdown as opposed to chronic depletion of m6A may also differ, with emerging degron systems a future valuable tool with which to address this. m6A affects the stability, half-life and translational efficiency of modified transcripts, although again, dissecting primary from secondary effects is challenging. On a broader scale however, it is clear that m6A is needed for sufficient transcriptomic plasticity to enable dramatic cellular change [77, 81, 82].

The careful dissection of the contribution of m6A to individual RNA transcripts may lead to a greater understanding of the subtleties of the system. Understanding the pattern of m6A modification on individual transcripts may help tease out sub-populations that may, perhaps, have differing roles. m6A modification of RNA transcripts that can alter their useful lifespan, whether by attenuating half life or

promoting translational efficiency are important in the cellular context, but it is tempting to speculate that a deeper understanding of the effects of RNA modification upon ncRNA would demonstrate the most profound effects, with modified ncRNA being the final functional molecule, rather than a vehicle to check and balance translation.

In this review, we have focused upon the lncRNA MALAT1 and Xist, the ncRNAs for which there is currently the greatest body of literature concerning the effect of m6A. These two systems are dramatically different, with MALAT1 having a profound effect upon disease state, but with far less evidence for its role in healthy cells. As with the earliest RNA studies however, the abundance of MALAT1 has enabled researchers to map its m6A modifications with relative ease. This has led to the identification of the m6A 'switch', along with some of the most in depth biochemical analysis of the structural changes that take place as a consequence of m6A, albeit *in vitro*. To fully understand how m6A influences the *in vivo* behaviour of an individual molecule of MALAT1, and its direct downstream consequences on gene expression will however take considerable work, necessitating mutation of MALAT1 in normal and disease models.

Researchers working to understand the role of m6A on the lncRNA Xist have the advantage that multiple, sophisticated, *in vivo* systems are established in which mutation of Xist and the direct observation of its effect upon gene expression are possible. Furthermore, the restricted role of Xist, both spatially and functionally, means that such experiments do not have detrimental effects upon the cell as a whole. Limiting the impact of mutation in this way, avoiding constitutive knockout of components of the methyltransferase system, will be necessary for us to elegantly dissect the m6A system further. Indeed, deletion of m6A domains in Xist and analysis of those genes immediately affected downstream, suggests the role of m6A on Xist to be far less dramatic than when cellular wide systems are disrupted, and whole cell transcript levels used to analyse gene expression.

Biology does not idly conserve redundant systems through evolution. m6A is a conserved modification of an ancient molecule, whose functions still elude researchers. Technological advances in the last six years have provided us with the

resources to map m6A, and with that the ability to target specific transcripts and dissect their function. We must use these tools to pick apart the puzzle of m6A, but to fully appreciate the roles of different modified RNAs and their reader proteins will take time and carefully designed scientific strategies.

Acknowledgements

We would like to thank colleagues in the Brockdorff lab for discussion and comments. Work in the Brockdorff lab is funded by the Wellcome Trust (103768).

Figure 1. m6A writers, readers and erasers. m6A is placed on RNA at RRACH motifs (where R: A or G and H: A, C or U) by the m6A methyltransferase complex. A cartoon representation of the mammalian complex depicts the catalytic core of METTL3 and METTL14 along with putative sites of interaction with the regulatory components WTAP, VIRMA, RBM15/15B, ZC3H13 and CBLL1. m6A can be removed by ALKBH5 and FTO. Readers of m6A show specificity for different RNA targets within the nucleus (YTHDC1, HNRNPA2B1, HNRNPC, HNRNPG), cytoplasm (YTHDF1, YTHDF3, IGF2BP's) or both (YTHDC2). YTHDF2 is a predominantly cytoplasmic protein, which is able to translocate to the nucleus under conditions of heat shock stress in order to aid cap-independent translation [71].

Figure 2. MALAT1 is highly modified with m6A that can act as a structural 'switch'. m6A-miCLIP mapping of the human lncRNA MALAT1 identified multiple peaks of m6A [15]. Of these, SCARLET technology confirmed four m6A motifs to be consistently modified in multiple cell lines [93]. Residues A2515 and A2577 reside within stem loop structures in MALAT1, which when modified by m6A, facilitate binding of HNRNPG and HNRNPC respectively, either encompassing or proximal to the RRACH consensus motif [96, 98].

Figure 3. The m6A profiles of the lncRNA Xist and antisense Tsix in mouse ES cells. A. The Tsix transcript is transcribed across the transcriptional start site of Xist and is modified with m6A (unpublished data). B. Xist is comprised of A, F, B, C, D and E-repeat regions, with SPEN binding to the A-repeats critical for downstream transcriptional silencing [136]. The most prominent of the three m6A peaks spans approximately 200bp immediately after the A-repeats region and has been profiled using different m6A mapping techniques [13, 15, 28, 115]. C. An enlarged scale view of the 5' region of Xist illustrates the position of the 7.5 A repeat regions, the SPEN binding domain and the broad peak of m6A enrichment immediately after the A-repeats. Xist deletion mutant 11G removes seven of the nine m6A sites but has no effect upon silencing, whilst deletion mutant 3A, which encroaches further into the SPEN binding domain has a small, but significant effect[115].

References

1. Cech, T.R. and J.A. Steitz, *The noncoding RNA revolution-trashing old rules to forge new ones.* Cell, 2014. **157**(1): p. 77-94.
2. Esteller, M., *Non-coding RNAs in human disease.* Nat Rev Genet, 2011. **12**(12): p. 861-74.
3. Kopp, F. and J.T. Mendell, *Functional Classification and Experimental Dissection of Long Noncoding RNAs.* Cell, 2018. **172**(3): p. 393-407.
4. Furlan, G., et al., *The Ftx Noncoding Locus Controls X Chromosome Inactivation Independently of Its RNA Products.* Mol Cell, 2018. **70**(3): p. 462-472.e8.
5. Liu, B., et al., *Long noncoding RNA lncKdm2b is required for ILC3 maintenance by initiation of Zfp292 expression.* Nat Immunol, 2017. **18**(5): p. 499-508.
6. Romero-Barrios, N., et al., *Splicing regulation by long noncoding RNAs.* Nucleic Acids Research, 2018. **46**(5): p. 2169-2184.
7. Yamazaki, T., et al., *Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation.* Molecular Cell, 2018. **70**(6): p. 1038-1053.e7.
8. Lin, Y., et al., *Structural analyses of NEAT1 lncRNAs suggest long-range RNA interactions that may contribute to paraspeckle architecture.* Nucleic Acids Research, 2018. **46**(7): p. 3742-3752.
9. Sunwoo, H., et al., *Men ε/β nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles.* Genome Research, 2009. **19**(3): p. 347-359.
10. Gilbert, W.V., T.A. Bell, and C. Schaening, *Messenger RNA modifications: Form, distribution, and function.* Science, 2016. **352**(6292): p. 1408-12.
11. Roundtree, I.A., et al., *Dynamic RNA Modifications in Gene Expression Regulation.* Cell, 2017. **169**(7): p. 1187-1200.
12. Esteller, M. and P.P. Pandolfi, *The Epitranscriptome of Noncoding RNAs in Cancer.* Cancer Discov, 2017. **7**(4): p. 359-368.
13. Dominissini, D., et al., *Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq.* Nature, 2012. **485**(7397): p. 201-6.
14. Meyer, K.D., et al., *Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons.* Cell, 2012. **149**(7): p. 1635-46.
15. Linder, B., et al., *Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome.* Nat Methods, 2015. **12**(8): p. 767-72.
16. Grozhik, A.V., et al., *Mapping m(6)A at Individual-Nucleotide Resolution Using Crosslinking and Immunoprecipitation (miCLIP).* Methods Mol Biol, 2017. **1562**: p. 55-78.
17. Dubin, D.T. and R.H. Taylor, *The methylation state of poly A-containing messenger RNA from cultured hamster cells.* Nucleic Acids Res, 1975. **2**(10): p. 1653-68.
18. Perry, R.P. and K. Scherrer, *The methylated constituents of globin mRNA.* FEBS Lett, 1975. **57**(1): p. 73-8.
19. Wei, C.M., A. Gershowitz, and B. Moss, *Methylated nucleotides block 5' terminus of HeLa cell messenger RNA.* Cell, 1975. **4**(4): p. 379-86.

20. Desrosiers, R., K. Friderici, and F. Rottman, *Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells.* Proc Natl Acad Sci U S A, 1974. **71**(10): p. 3971-5.
21. Horowitz, S., et al., *Mapping of N6-methyladenosine residues in bovine prolactin mRNA.* Proc Natl Acad Sci U S A, 1984. **81**(18): p. 5667-71.
22. Schwartz, S., et al., *High-resolution mapping reveals a conserved, widespread, dynamic meiotically regulated mRNA methylation program.* Cell, 2013. **155**(6): p. 1409-1421.
23. Luo, G.Z., et al., *Unique features of the m6A methylome in Arabidopsis thaliana.* Nat Commun, 2014. **5**: p. 5630.
24. Lence, T., et al., *m(6)A modulates neuronal functions and sex determination in Drosophila.* Nature, 2016. **540**(7632): p. 242-247.
25. Zhao, B.S., et al., *m(6)A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition.* Nature, 2017. **542**(7642): p. 475-478.
26. Ke, S., et al., *m(6)A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover.* Genes Dev, 2017. **31**(10): p. 990-1006.
27. Chen, K., et al., *High-Resolution N(6)-Methyladenosine (m(6)A) Map Using Photo-Crosslinking-Assisted m(6)A Sequencing().* Angewandte Chemie (International ed. in English), 2015. **54**(5): p. 1587-1590.
28. Ke, S., et al., *A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation.* Genes Dev, 2015. **29**(19): p. 2037-53.
29. Liu, J., et al., *A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation.* Nat Chem Biol, 2014. **10**(2): p. 93-5.
30. Wang, P., K.A. Doxtader, and Y. Nam, *Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases.* Mol Cell, 2016. **63**(2): p. 306-317.
31. Horiuchi, K., et al., *Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle.* J Biol Chem, 2013. **288**(46): p. 33292-302.
32. Ping, X.-L., et al., *Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase.* Cell Research, 2014. **24**: p. 177.
33. Schwartz, S., et al., *Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites.* Cell Rep, 2014. **8**(1): p. 284-96.
34. Patil, D.P., et al., *m(6)A RNA methylation promotes XIST-mediated transcriptional repression.* Nature, 2016. **537**(7620): p. 369-373.
35. Knuckles, P., et al., *Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m(6)A machinery component Wtap/Fl(2)d.* Genes Dev, 2018. **32**(5-6): p. 415-429.
36. Wen, J., et al., *Zc3h13 Regulates Nuclear RNA m6A Methylation and Mouse Embryonic Stem Cell Self-Renewal.* Molecular Cell, 2018. **69**(6): p. 1028-1038.e6.
37. Kamil, R., et al., *Identification of factors required for m6A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI.* New Phytologist, 2017. **215**(1): p. 157-172.

38. Jia, G., et al., *N(6)-Methyladenosine in Nuclear RNA is a Major Substrate of the Obesity-Associated FTO.* Nature chemical biology, 2011. **7**(12): p. 885-887.
39. Zheng, G., et al., *ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility.* Molecular cell, 2013. **49**(1): p. 18-29.
40. Fu, Y., et al., *FTO-mediated formation of N6-hydroxymethyladenosine and N6-formyladenosine in mammalian RNA.* Nature Communications, 2013. **4**: p. 1798.
41. Dina, C., et al., *Variation in FTO contributes to childhood obesity and severe adult obesity.* Nature Genetics, 2007. **39**: p. 724.
42. Scuteri, A., et al., *Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits.* PLoS Genet, 2007. **3**(7): p. e115.
43. Frayling, T.M., et al., *A Common Variant in the FTO Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity.* Science, 2007. **316**(5826): p. 889.
44. Hess, M.E., et al., *The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry.* Nat Neurosci, 2013. **16**(8): p. 1042-8.
45. Bartosovic, M., et al., *N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3'-end processing.* Nucleic Acids Res, 2017. **45**(19): p. 11356-11370.
46. Mauer, J., et al., *Reversible methylation of m(6)Am in the 5' cap controls mRNA stability.* Nature, 2017. **541**(7637): p. 371-375.
47. Tang, C., et al., *ALKBH5-dependent m6A demethylation controls splicing and stability of long 3′-UTR mRNAs in male germ cells.* Proceedings of the National Academy of Sciences, 2018. **115**(2): p. E325-E333.
48. Zhang, S., et al., *m6A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program.* Cancer Cell, 2017. **31**(4): p. 591-606.e6.
49. Darnell, R.B., S. Ke, and J.E. Darnell, Jr., *Pre-mRNA processing includes N(6) methylation of adenosine residues that are retained in mRNA exons and the fallacy of "RNA epigenetics".* Rna, 2018. **24**(3): p. 262-267.
50. Zhao, B.S., et al., *Our views of dynamic N(6)-methyladenosine RNA methylation.* Rna, 2018. **24**(3): p. 268-272.
51. Edupuganti, R.R., et al., *N(6)-methyladenosine (m(6)A) recruits and repels proteins to regulate mRNA homeostasis.* Nat Struct Mol Biol, 2017. **24**(10): p. 870-878.
52. Patil, D.P., B.F. Pickering, and S.R. Jaffrey, *Reading m(6)A in the Transcriptome: m(6)A-Binding Proteins.* Trends Cell Biol, 2018. **28**(2): p. 113-127.
53. Imai, Y., et al., *Cloning of a gene, YT521, for a novel RNA splicing-related protein induced by hypoxia/reoxygenation.* Brain Res Mol Brain Res, 1998. **53**(1-2): p. 33-40.
54. Hartmann, A.M., et al., *The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn).* Mol Biol Cell, 1999. **10**(11): p. 3909-26.
55. Stoilov, P., I. Rafalska, and S. Stamm, *YTH: a new domain in nuclear proteins.* Trends Biochem Sci, 2002. **27**(10): p. 495-7.

56. Zhang, Z., et al., *The YTH domain is a novel RNA binding domain.* J Biol Chem, 2010. **285**(19): p. 14701-10.
57. Xiao, W., et al., *Nuclear m6A Reader YTHDC1 Regulates mRNA Splicing.* Molecular Cell, 2016. **61**(4): p. 507-519.
58. Roundtree, I.A., et al., *YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs.* Elife, 2017. **6**.
59. Haussmann, I.U., et al., *m(6)A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination.* Nature, 2016. **540**(7632): p. 301-304.
60. Kan, L., et al., *The m(6)A pathway facilitates sex determination in Drosophila.* Nat Commun, 2017. **8**: p. 15737.
61. Hsu, P.J., et al., *Ythdc2 is an N(6)-methyladenosine binding protein that regulates mammalian spermatogenesis.* Cell Res, 2017. **27**(9): p. 1115-1127.
62. Bailey, A.S., et al., *The conserved RNA helicase YTHDC2 regulates the transition from proliferation to differentiation in the germline.* Elife, 2017. **6**.
63. Wojtas, M.N., et al., *Regulation of m(6)A Transcripts by the 3'-->5' RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline.* Mol Cell, 2017. **68**(2): p. 374-387.e12.
64. Kretschmer, J., et al., *The m6A reader protein YTHDC2 interacts with the small ribosomal subunit and the 5'-3' exoribonuclease XRN1.* Rna, 2018.
65. Wang, X., et al., *N6-methyladenosine Modulates Messenger RNA Translation Efficiency.* Cell, 2015. **161**(6): p. 1388-1399.
66. Meyer, K.D., et al., *5' UTR m(6)A Promotes Cap-Independent Translation.* Cell, 2015. **163**(4): p. 999-1010.
67. Wang, X., et al., *N6-methyladenosine-dependent regulation of messenger RNA stability.* Nature, 2014. **505**(7481): p. 117-20.
68. Du, H., et al., *YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex.* Nature Communications, 2016. **7**: p. 12626.
69. Zhao, B.S., I.A. Roundtree, and C. He, *Post-transcriptional gene regulation by mRNA modifications.* Nat Rev Mol Cell Biol, 2017. **18**(1): p. 31-42.
70. Wang, X. and C. He, *Dynamic RNA Modifications in Posttranscriptional Regulation.* Molecular Cell, 2014. **56**(1): p. 5-12.
71. Zhou, J., et al., *Dynamic m(6)A mRNA methylation directs translational control of heat shock response.* Nature, 2015. **526**(7574): p. 591-4.
72. Shi, H., et al., *YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA.* Cell Res, 2017. **27**(3): p. 315-328.
73. Li, A., et al., *Cytoplasmic m(6)A reader YTHDF3 promotes mRNA translation.* Cell Res, 2017. **27**(3): p. 444-447.
74. Zhou, K.I. and T. Pan, *An additional class of m6A readers.* Nature Cell Biology, 2018. **20**(3): p. 230-232.
75. Alarcon, C.R., et al., *HNRNPA2B1 Is a Mediator of m(6)A-Dependent Nuclear RNA Processing Events.* Cell, 2015. **162**(6): p. 1299-308.
76. Alarcon, C.R., et al., *N6-methyladenosine marks primary microRNAs for processing.* Nature, 2015. **519**(7544): p. 482-5.

77. Chen, T., et al., *m⁶A RNA Methylation Is Regulated by MicroRNAs and Promotes Reprogramming to Pluripotency.* Cell Stem Cell, 2015. **16**(3): p. 289-301.
78. Huang, H., et al., *Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation.* Nat Cell Biol, 2018. **20**(3): p. 285-295.
79. Wu, B., et al., *Molecular basis for the specific and multivariant recognitions of RNA substrates by human hnRNP A2/B1.* Nat Commun, 2018. **9**(1): p. 420.
80. Roignant, J.Y. and M. Soller, *m(6)A in mRNA: An Ancient Mechanism for Fine-Tuning Gene Expression.* Trends Genet, 2017. **33**(6): p. 380-390.
81. Batista, P.J., et al., *m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells.* Cell Stem Cell, 2014. **15**(6): p. 707-19.
82. Geula, S., et al., *Stem cells. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation.* Science, 2015. **347**(6225): p. 1002-6.
83. Wang, Y., et al., *N6-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications.* Nature Neuroscience, 2018. **21**(2): p. 195-206.
84. Ji, P., et al., *MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer.* Oncogene, 2003. **22**(39): p. 8031-41.
85. Zhang, X., M.H. Hamblin, and K.J. Yin, *The long noncoding RNA Malat1: Its physiological and pathophysiological functions.* RNA Biol, 2017. **14**(12): p. 1705-1714.
86. Wilusz, J.E., S.M. Freier, and D.L. Spector, *3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA.* Cell, 2008. **135**(5): p. 919-32.
87. West, J.A., et al., *The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites.* Mol Cell, 2014. **55**(5): p. 791-802.
88. Bernard, D., et al., *A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression.* Embo j, 2010. **29**(18): p. 3082-93.
89. Tripathi, V., et al., *The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation.* Mol Cell, 2010. **39**(6): p. 925-38.
90. Zhang, B., et al., *The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult.* Cell Rep, 2012. **2**(1): p. 111-23.
91. Nakagawa, S., et al., *Malat1 is not an essential component of nuclear speckles in mice.* Rna, 2012. **18**(8): p. 1487-99.
92. Eissmann, M., et al., *Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development.* RNA Biol, 2012. **9**(8): p. 1076-87.
93. Liu, N., et al., *Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA.* Rna, 2013. **19**(12): p. 1848-56.
94. Spitale, R.C., et al., *Structural imprints in vivo decode RNA regulatory mechanisms.* Nature, 2015. **519**(7544): p. 486-490.

95. Cienikova, Z., et al., *Structural and mechanistic insights into poly(uridine) tract recognition by the hnRNP C RNA recognition motif.* J Am Chem Soc, 2014. **136**(41): p. 14536-44.
96. Liu, N., et al., *N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions.* Nature, 2015. **518**(7540): p. 560-4.
97. Zhou, K.I., et al., *N6-methyladenosine modification in a long non-coding RNA hairpin predisposes its conformation to protein binding.* Journal of molecular biology, 2016. **428**(5 Pt A): p. 822-833.
98. Liu, N., et al., *N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein.* Nucleic Acids Res, 2017. **45**(10): p. 6051-6063.
99. Pendleton, K.E., et al., *The U6 snRNA m6A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention.* Cell, 2017. **169**(5): p. 824-835.e14.
100. Mendel, M., et al., *Methylation of Structured RNA by the m6A Writer METTL16 Is Essential for Mouse Embryonic Development.* Molecular Cell, 2018. **71**(6): p. 986-1000.e11.
101. Brown, J.A., et al., *Methyltransferase-like protein 16 binds the 3'-terminal triple helix of MALAT1 long noncoding RNA.* Proc Natl Acad Sci U S A, 2016. **113**(49): p. 14013-14018.
102. Ruszkowska, A., et al., *Structural insights into the RNA methyltransferase domain of METTL16.* Sci Rep, 2018. **8**(1): p. 5311.
103. Lyon, M.F., *Sex chromatin and gene action in the mammalian X-chromosome.* Am J Hum Genet, 1962. **14**: p. 135-48.
104. Lyon, M.F., *Gene action in the X-chromosome of the mouse (Mus musculus L.).* Nature, 1961. **190**: p. 372-3.
105. Brockdorff, N., et al., *The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus.* Cell, 1992. **71**(3): p. 515-26.
106. Brown, C.J., et al., *The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus.* Cell, 1992. **71**(3): p. 527-42.
107. Clemson, C.M., et al., *XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure.* J Cell Biol, 1996. **132**(3): p. 259-75.
108. Penny, G.D., et al., *Requirement for Xist in X chromosome inactivation.* Nature, 1996. **379**(6561): p. 131-7.
109. Hong, Y.-K., et al., *A new structure for the murine Xist gene and its relationship to chromosome choice/counting during X-chromosome inactivation.* Proceedings of the National Academy of Sciences, 1999. **96**(12): p. 6829.
110. Disteche, C.M., *Dosage Compensation of the Sex Chromosomes.* Annual review of genetics, 2012. **46**: p. 537-560.
111. Barr, M.L. and E.G. Bertram, *A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis.* Nature, 1949. **163**(4148): p. 676.
112. Sado, T. and A.C. Ferguson-Smith, *Imprinted X inactivation and reprogramming in the preimplantation mouse embryo.* Hum Mol Genet, 2005. **14 Spec No 1**: p. R59-64.

113. Chu, C., et al., *Systematic discovery of Xist RNA binding proteins.* Cell, 2015. **161**(2): p. 404-16.
114. Lee, J.T., N. Lu, and Y. Han, *Genetic analysis of the mouse X inactivation center defines an 80-kb multifunction domain.* Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3836-41.
115. Nesterova, T.B., et al., *Systematic Allelic Analysis Defines the Interplay of Key Pathways in X Chromosome Inactivation.* bioRxiv, 2018.
116. Stavropoulos, N., N. Lu, and J.T. Lee, *A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10232-7.
117. Sado, T., et al., *Regulation of imprinted X-chromosome inactivation in mice by Tsix.* Development, 2001. **128**(8): p. 1275-86.
118. Rougeulle, C. and P. Avner, *The role of antisense transcription in the regulation of X-inactivation.* Curr Top Dev Biol, 2004. **63**: p. 61-89.
119. Pintacuda, G., et al., *hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing.* Mol Cell, 2017. **68**(5): p. 955-969.e10.
120. Loda, A., et al., *Genetic and epigenetic features direct differential efficiency of Xist-mediated silencing at X-chromosomal and autosomal locations.* Nature Communications, 2017. **8**(1): p. 690.
121. Brockdorff, N., *X-chromosome inactivation: closing in on proteins that bind Xist RNA.* Trends Genet, 2002. **18**(7): p. 352-8.
122. Wutz, A., T.P. Rasmussen, and R. Jaenisch, *Chromosomal silencing and localization are mediated by different domains of Xist RNA.* Nat Genet, 2002. **30**(2): p. 167-74.
123. Pintacuda, G., A.N. Young, and A. Cerase, *Function by Structure: Spotlights on Xist Long Non-coding RNA.* Front Mol Biosci, 2017. **4**: p. 90.
124. Monfort, A., et al., *Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells.* Cell Rep, 2015. **12**(4): p. 554-61.
125. Moindrot, B., et al., *A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing.* Cell Rep, 2015. **12**(4): p. 562-72.
126. McHugh, C.A., et al., *The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3.* Nature, 2015. **521**(7551): p. 232-6.
127. Tavares, L., et al., *RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3.* Cell, 2012. **148**(4): p. 664-78.
128. Almeida, M., et al., *PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation.* Science, 2017. **356**(6342): p. 1081-1084.
129. Cooper, S., et al., *Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment.* Cell Rep, 2014. **7**(5): p. 1456-1470.
130. Cooper, S., et al., *Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2.* Nat Commun, 2016. **7**: p. 13661.
131. Blackledge, N.P., et al., *Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation.* Cell, 2014. **157**(6): p. 1445-59.

132. Cirillo, D., et al., *Quantitative predictions of protein interactions with long noncoding RNAs.* Nat Methods, 2016. **14**(1): p. 5-6.
133. McIntyre, W., et al., *Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection.* Nucleic Acids Research, 2018. **46**(11): p. 5776-5791.
134. Pan, T., *Modifications and functional genomics of human transfer RNA.* Cell Res, 2018. **28**(4): p. 395-404.
135. Sloan, K.E., et al., *Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function.* RNA Biol, 2017. **14**(9): p. 1138-1152.
136. Chen, C.K., et al., *Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing.* Science, 2016. **354**(6311): p. 468-472.