

Transcriptome-wide mapping reveals reversible and dynamic N¹-methyladenosine methylome

Xiaoyu Li^{1,2,5}, Xushen Xiong^{1-3,5}, Kun Wang^{1,2}, Lixia Wang^{1,2}, Xiaoting Shu¹⁻³, Shiqing Ma¹⁻³ & Chengqi Yi^{1,2,4,*}

N¹-Methyladenosine (m¹A) is a prevalent post-transcriptional RNA modification, yet little is known about its abundance, topology and dynamics in mRNA. Here, we show that m¹A is prevalent in *Homo sapiens* mRNA, which shows an m¹A/A ratio of ~0.02%. We develop the m¹A-ID-seq technique, based on m¹A immunoprecipitation and the inherent ability of m¹A to stall reverse transcription, as a means for transcriptome-wide m¹A profiling. m¹A-ID-seq identifies 901 m¹A peaks (from 600 genes) in mRNA and noncoding RNA and reveals a prominent feature, enrichment in the 5' untranslated region of mRNA transcripts, that is distinct from the pattern for N⁶-methyladenosine, the most abundant internal mammalian mRNA modification. Moreover, m¹A in mRNA is reversible by ALKBH3, a known DNA/RNA demethylase. Lastly, we show that m¹A methylation responds dynamically to stimuli, and we identify hundreds of stress-induced m¹A sites. Collectively, our approaches allow comprehensive analysis of m¹A modification and provide tools for functional studies of potential epigenetic regulation via the reversible and dynamic m¹A methylation.

Decades of research have revealed more than 100 different types of post-transcriptional modifications to RNA molecules¹. One prominent and frequently occurring modification is N¹-methyladenosine, or m¹A. m¹A is typically found at position 58 in the tRNA T-loop, catalyzed by the essential tRNA (m¹A58) methyltransferase complex (Trmt6 and Trmt61A) for cytoplasmic tRNAs² or by Trmt61B for mitochondrial tRNAs³. m¹A58 is fully modified in initiator tRNA^{Met} and stabilizes its tertiary structure⁴, and this position is hypomodified for ~25% of other tRNA species, which affects their association with polysome and hence utility in translation⁵. m¹A is also present at position 9 of metazoan mitochondrial tRNAs^{1,6-8} and in bacterial and eukaryotic rRNAs^{1,9-11}. m¹A in tRNA can respond to environmental stress^{12,13}, and m¹A in rRNA can affect ribosome biogenesis¹¹ and mediate antibiotic resistance in bacteria¹⁴.

While m¹A in these abundant noncoding RNAs (ncRNAs) has been well studied, whether m¹A is present in mRNA is currently unknown. Progress in establishing the prevalence of m¹A has lagged behind other research on m¹A in ncRNAs, presumably as a result of the lack of methods for detecting m¹A. Because mRNA is low in abundance, an m¹A detection method must be sensitive enough to detect transcriptome-wide m¹A methylation. In addition, owing to the position of its methyl group at the Watson-Crick interface, m¹A effectively stalls reverse transcription (RT) and results in truncated RT products, although a small percentage (~10–20%) of readthrough cDNA containing an m¹A-induced misincorporated nucleotide is produced^{15–18}. However, the pattern of mismatch is sequence dependent¹⁸, and direct application of the small readthrough fraction to mRNA methylation detection would require extremely high depth of sequencing. Therefore, special caution should be taken when combining RT and high-throughput sequencing for m¹A profiling in mRNA.

Here, we investigate the abundance, topology and dynamics of m¹A in the transcriptome of human embryonic kidney (HEK293T) cells. We develop m¹A-ID-seq, which combines m¹A immunoprecipitation and the tendency of m¹A to cause truncated RT products,

and use this technique for the first transcriptome-wide characterizations of m¹A. In m¹A-ID-seq, we take advantage of the biochemical activity of the *Escherichia coli* AlkB protein—an m¹A demethylase^{19,20}—as a means to identify high-confidence m¹A peaks. We show that m¹A is a prevalent internal mRNA modification, present in 887 transcripts from 600 human genes. m¹A is highly enriched within 5' untranslated regions (UTRs) and in the vicinity of start codons, and m¹A sites are associated with novel sequence motifs. Remarkably, m¹A in mRNA can be demethylated by ALKBH3—a known DNA/RNA demethylase^{21–24}—and thousands of reversible m¹A sites are identified in ALKBH3 knockout cells. Furthermore, m¹A in mRNA responds dynamically to stress conditions, and we detected hundreds of stress-induced m¹A peaks in cells subjected to serum starvation or H₂O₂ treatment. Taking these results together, we reveal a reversible and dynamic m¹A methylation program in the human transcriptome and provide methods for functional studies of m¹A-mediated biological regulation.

RESULTS

Quantitative analysis of m¹A in human mRNA

To determine whether m¹A is present in mRNA, we first performed quantitative MS analysis. Because m¹A is abundant in rRNA and tRNA, we adopted a previously developed mRNA purification procedure to minimize contamination by these abundant ncRNAs²⁵. m¹A is also known to undergo Dimroth rearrangement to m⁶A under alkaline conditions²⁶; therefore, we performed RNA digestion—a required step for MS quantification—at neutral pH so as to avoid underestimating m¹A abundance. Additionally, we carried out extra DNA removal steps to ensure that the purified mRNA was free of DNA contamination, which may contain m¹DA DNA damage. Furthermore, we used N⁶-isopentenyladenosine (i⁶A), which is found only in tRNA¹, as an indicator of tRNA contamination (Supplementary Results, Supplementary Fig. 1a,b). Applications of these protocols to different RNA populations revealed an m¹A/A ratio of approximately 0.02% for mRNAs isolated from various human cell lines (Supplementary Fig. 1a–c). This level of m¹A is

¹State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing, China. ²Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China. ³Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China. ⁴Synthetic and Functional Biomolecules Center, Department of Chemical Biology, College of Chemistry and Molecular Engineering, Peking University, Beijing, China. ⁵These authors contributed equally to this work. *e-mail: chengqi.yi@pku.edu.cn

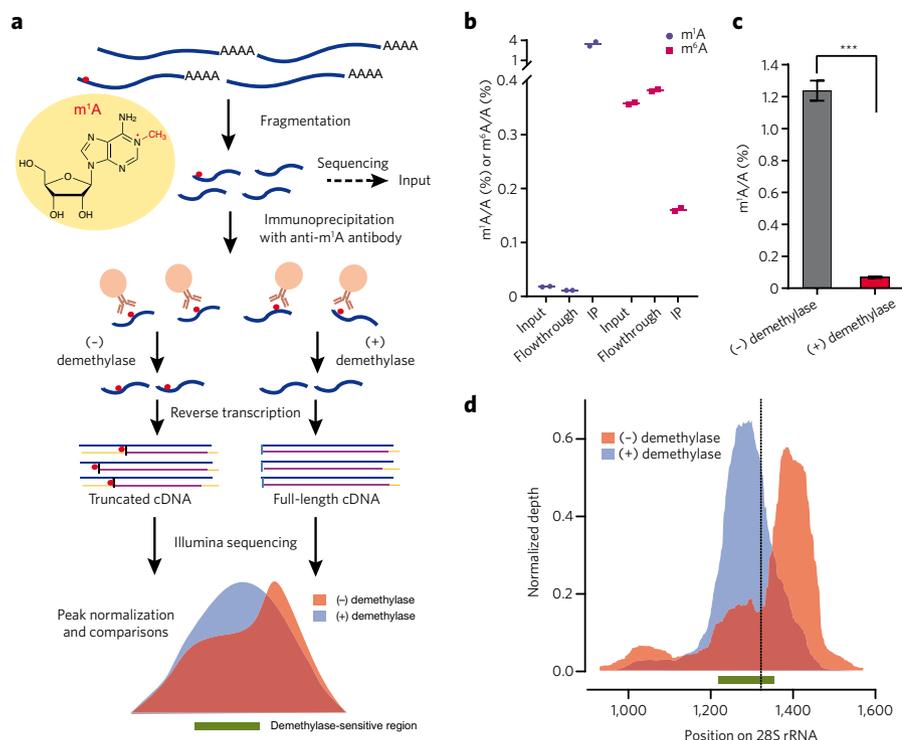


Figure 1 | m¹A-ID-seq utilizes an m¹A antibody and a biochemical demethylation reaction to enrich and identify high-confidence m¹A sites. (a) m¹A-ID-seq procedure. mRNA was fragmented (Input) and immunoprecipitated (IP). The m¹A-containing RNA was subjected to demethylase (“(+) demethylase”) or mock (“(-) demethylase”) treatment. The input, (-) demethylase and (+) demethylase samples were reverse transcribed and sequenced for high-confidence m¹A detection. Sequencing profiles of the (+) and (-) demethylase samples were compared and “demethylase-sensitive regions” (green bar) were identified. The demethylase-sensitive region harbors the m¹A site, yet base-resolution information for m¹A was difficult to obtain due to random priming during library preparation (see Online Methods). (b) The m¹A and m⁶A contents of input, flowthrough and immunoprecipitation (IP) samples in the m¹A-ID-seq were quantified by LC-MS/MS. m¹A, but not m⁶A, was highly enriched by m¹A antibody. (c) Total RNA from HEK293T cells, which has an m¹A/A ratio of ~1%, was used to optimize demethylation conditions. ~95% of m¹A in total RNA can be demethylated under optimized conditions. Values represent mean \pm s.e.m. ($n = 8$); *** $P < 0.01$ (t -test, two sided). (d) m¹A-ID-seq detects the known m¹A site in human rRNA (position 1322 in 28S rRNA, indicated by the dotted line). The ‘normalized depth’ of each nucleotide was calculated by dividing the sequencing depth of the nucleotide by the sum of the sequencing depths for all nucleotides within a peak. The green bar represents the demethylase-sensitive region.

about 5–10% the global level of m⁶A, which is the most abundant internal mRNA modification in eukaryotes and has a reported m⁶A/A ratio of about 0.2–0.5% when measured using several methods including quantitative MS, used in this study^{27,28}. Therefore, our results suggest that m¹A modification is also abundant in human mRNAs.

m¹A-ID-seq utilizes a demethylase for m¹A detection

To comprehensively study m¹A methylation in the transcriptome, we developed a method that combines m¹A-specific RNA immunoprecipitation with demethylase-assisted RNA sequencing (m¹A-ID-seq, shown in Fig. 1a). In m¹A-ID-seq, poly(A)⁺ RNA was fragmented into ~150 nucleotides (nt) and immunoprecipitated with an antibody to m¹A. To ensure the specificity of the antibody, we first performed dot-blot assays using synthetic oligonucleotides containing unmodified adenosines or site-specifically incorporated m¹A. The m¹A antibody bound selectively to the m¹A-containing oligonucleotides while showing no noticeable binding to the unmodified control (Supplementary Fig. 2a). Of particular importance, the m¹A antibody showed no cross-reactivity to the m⁶A-containing model sequence, demonstrating high specificity toward methylation

occurring only to the N¹ position of adenosines. We further tested the antibody’s ability to enrich m¹A-containing RNA: quantitative MS demonstrated high enrichment of m¹A after immunoprecipitation (Fig. 1b), and qRT-PCR results also showed that the antibody enriched the rRNA region containing m¹A1322 (a known m¹A site in human 28S rRNA using RefSeq annotation NR_003287; equivalent to m¹A1309 in ref. 18)^{1,18,29} by >500-fold (Supplementary Fig. 2b).

A unique aspect of m¹A-ID-seq is the utilization of *E. coli* AlkB, a DNA/RNA demethylase that performs repair functions *in vivo* and demethylates several base methylations including m¹A (refs. 19,20). Because m¹A can generate truncated RT products, we envisioned that AlkB-assisted removal of m¹A from the immunoprecipitated RNA would result in the accumulation of full-length cDNAs. Therefore, by comparing the sequencing profiles of the (+) and (-) demethylase samples, we would be able to detect ‘demethylase-sensitive’, and hence high-confidence, m¹A peaks (Fig. 1a). In fact, AlkB has recently been used in tRNA sequencing^{30,31}. To explore its utility in sequencing m¹A in mRNA, we performed extensive optimization of the *in vitro* reaction conditions so as to maximize the demethylation efficiency (Supplementary Fig. 3). Indeed, under optimal conditions, ~95% of total m¹A methylation can be removed, demonstrating that the AlkB demethylase is highly efficient (Fig. 1c).

m¹A-ID-seq detects known m¹A sites in rRNA

We next applied m¹A-ID-seq to characterize m¹A methylation in rRNA. To allow the identification of high-confidence m¹A peaks, we first defined a stringent peak-detection algorithm (requiring an estimated $q < 1.0 \times 10^{-10}$ and an enrichment of >3-fold; see Online Methods). Three rRNA sites are

known to be m¹A modified: A645 and A2142 in yeast 25S rRNA and A1322 in human 28S rRNA^{9,11,29}; all three sites were readily identified by the algorithm (Supplementary Fig. 4a,b).

We also defined a ‘demethylase sensitivity’ (DS) score as a second metric with which to evaluate the m¹A peaks. The sequencing profiles in the (+) and (-) demethylase samples were normalized and compared; DS scores were then calculated for 10-nt-wide sliding windows that span the entire peaks at an increment of 5 nt (Online Methods). Indeed, a ‘demethylase-sensitive region’, formed by several adjacent sliding windows that show positive DS scores, was identified around A1322 in 28S rRNA (the green bar in Fig. 1d).

To further prove that this inherent property of m¹A during RT could be used for m¹A mapping, we screened different reverse transcriptases (RTases) and RT conditions. We first showed that such demethylase-sensitive regions of high-throughput sequencing profiles were found for both SuperScript III and AMV RTase (Supplementary Fig. 5a); moreover, we were able to maximize RT stops at sites of m¹A for AMV RTase through optimization of the RT conditions (Supplementary Fig. 5b), thereby increasing the sensitivity of m¹A-ID-seq. Interestingly, we noticed that DS scores can be used to improve the resolution of m¹A detection: by scoring

windows within the matched peaks, we effectively narrowed down the m¹A site from a peak spanning ~350 nt to a 130-nt region harboring A1322 (Fig. 1d). Altogether, these results demonstrated that m¹A-ID-seq readily detects the known m¹A sites in rRNA.

m¹A-ID-seq detects 901 m¹A peaks in human transcriptome

Having validated the ability of m¹A-ID-seq to detect m¹A in rRNA, we next analyzed m¹A methylation in mRNA. While requiring strict peak-calling criteria (Online Methods), we identified 901 high-confidence m¹A peaks from the transcriptome of HEK293T cells (Supplementary Data Set 1). These m¹A peaks originate from 887 transcripts, which are encoded by 600 human genes; hence, the majority (~96.9%) of the transcripts contain only one m¹A peak. Both messenger RNAs (841) and noncoding transcripts (46) are m¹A modified (Fig. 2a–c and Supplementary Fig. 6); for instance, the disease-related long noncoding RNA MALAT1 contains m¹A methylation (Supplementary Data Set 1).

To validate mRNA m¹A peaks identified by m¹A-ID-seq, we designed an ‘orthogonal’ method that is independent of the m¹A antibody and the AlkB demethylase (Supplementary Figs. 7 and 8). In this method, we first used Dimroth rearrangement to convert endogenous m¹A to m⁶A (ref. 26), and then performed immunoprecipitation experiments using a widely approved m⁶A antibody^{32,33}. Eight m¹A-containing transcripts were selected for examination, and all eight could be validated using this approach.

To predict molecular functions that m¹A may potentially participate in, we used the DAVID database to identify the Gene Ontology

(GO) terms that are enriched for m¹A-containing transcripts. The results showed that transcripts with m¹A are involved in various processes and functions, with enrichment in transcription-factor binding ($P = 3.93 \times 10^{-8}$), RNA binding ($P = 6.88 \times 10^{-6}$), enzyme binding ($P = 6.99 \times 10^{-6}$), mRNA processing ($P = 3.49 \times 10^{-5}$) and others (Supplementary Fig. 9). Collectively, these data demonstrated that m¹A-containing RNAs are involved in a variety of processes and functions of human cells.

m¹A is enriched within 5' UTR and near start codons

We next investigated the distribution of m¹A peaks within regions of the transcriptome. 674 of 901 (~75%) peaks exhibit demethylase-sensitive regions in both replicates, and 766 of 901 (~85%) peaks exhibit demethylase-sensitive regions in at least one replicate (Supplementary Data Set 1), demonstrating the high specificity and robustness of m¹A-ID-seq. An additional benefit of the demethylase-sensitive set is that the resolution of m¹A detection was improved from ~105 nt to ~55 nt (Supplementary Fig. 10); thus, the 766 demethylase-sensitive regions were used for further topology analysis. Four nonoverlapping transcript segments were defined for m¹A peak assignments: 5' UTR, near start codon (50-nt window centered on the start codon), coding sequence (CDS) and 3' UTR. m¹A peaks were markedly most common within two segments: 5' UTR (59.1% of m¹A peaks) and near start codon (14.9% of m¹A peaks) (Fig. 2a). To determine whether m¹A methylation is preferentially enriched in these transcript segments, we normalized the m¹A peaks using the relative fraction each segment occupied in the transcriptome. The 5' UTR and near start codon segments stood out as most enriched in m¹A peaks, representing 9.2- and 10.3-fold enrichment, respectively, over the distribution expected by chance ($P < 2.2 \times 10^{-16}$ and $P = 7.2 \times 10^{-10}$, Fig. 2c and Supplementary Fig. 6b). This is in sharp contrast to m⁶A, which is most strongly enriched in the 3' UTR segment^{32–35}. Hence, the distribution patterns of these adenosine methylations are distinct from each other.

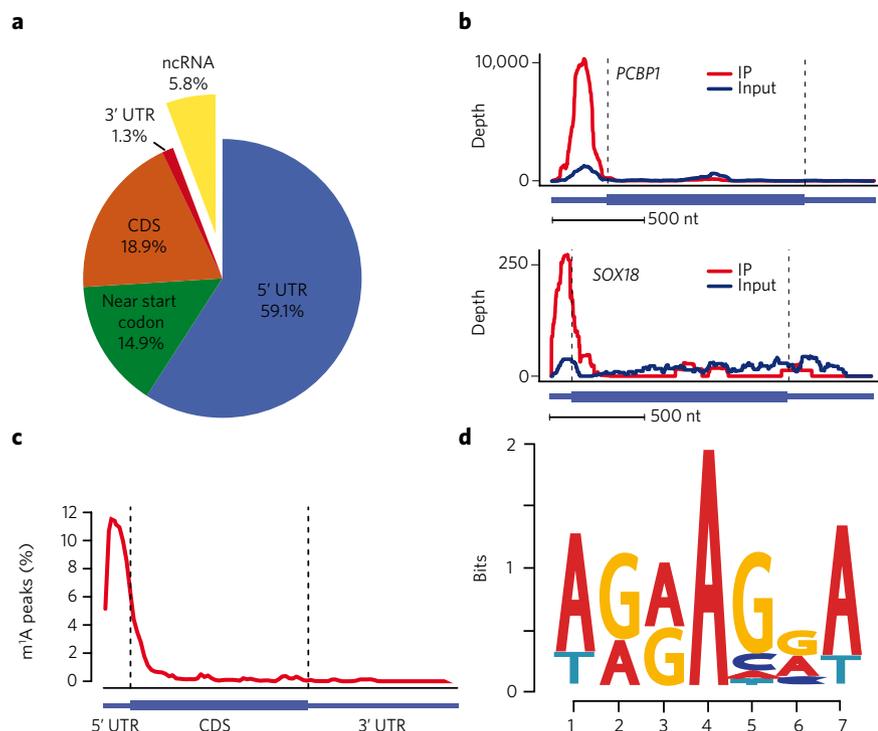


Figure 2 | m¹A-ID-seq reveals the topology of m¹A in the human transcriptome. (a) The pie chart shows the percentage of m¹A peaks in each nonoverlapping segment. The region ‘Near start codon’ was defined as 25 nt upstream and 25 nt downstream from the start codon. (b) Representative views of typical m¹A peaks on mRNA (*PCBP1* and *SOX18*). Transcript architecture is shown below, with thin and thick bars representing UTRs and CDSs, respectively. (c) Distribution of m¹A peaks across mRNA segments. Each segment was normalized according to its average length in RefSeq annotation. The total normalized transcript was binned into regions of 1% total length, and the percentage of m¹A peaks belonging to each bin was calculated. (d) One representative sequence motif enriched with m¹A (E value = 1.2×10^{-67}). Because the start codon does not match the sequences of the enriched motifs, it is less likely that the start codon itself is modified by m¹A.

To examine whether m¹A peaks are associated with potential sequence elements, we performed an unbiased motif search using the MEME software suite. In order to ensure the reliability of the discovered motifs, we adopted two different randomization methods to define the control sequence (Online Methods). Several motifs were identified by MEME (Fig. 2d and Supplementary Fig. 11); one prominent example is a GA-rich consensus (Fig. 2d). Such motifs again differ greatly from the motif for m⁶A, RRACU (where R stands for purine)^{32,33}. We also examined several motifs known to be present in 5' UTR (for instance, iron-responsive element and upstream AUG); however, we did not find significant enrichment for these regulatory sequences. In addition, the sequences of m¹A peaks have higher GC content than do those of non-m¹A controls (Mann-Whitney U -test, $P < 2.2 \times 10^{-16}$, Supplementary Fig. 12a); methylated sites also appear to be more structured than randomly selected counterparts from other genes (Mann-Whitney U -test, $P < 2.2 \times 10^{-16}$, Supplementary Fig. 12b).

m¹A is reversible by the ALKBH3 demethylase

Reversible methylation has recently emerged as a prominent mechanism of epigenetic regulations in mammalian cells. To explore whether the m¹A modification is reversible, we focused on ALKBH3, which is a human homolog of *E. coli* AlkB and has been shown to repair methylation damage to RNA *in vitro*^{22,23}. We first showed that recombinant ALKBH3 protein has robust demethylation activity against a synthetic RNA substrate containing a site-specifically incorporated m¹A modification (Fig. 3a and Supplementary Fig. 13a,b); in fact, this activity on RNA is slightly higher than that of ALKBH3 in repairing single-stranded m¹dA-containing DNA oligos (Supplementary Fig. 13b).

To investigate whether m¹A methylation in mRNA is a substrate of ALKBH3 *in vivo*, we first measured the m¹A/A ratio in cells lacking or overexpressing ALKBH3 (an ALKBH3^{-/-} HEK293T cell line was generated with the CRISPR-Cas9 system). Indeed, knockout and overexpression of ALKBH3 in HEK293T cells led to an increased or decreased m¹A/A ratio in mRNA, respectively (Supplementary Figs. 14 and 18). We also showed that manipulation of ALKBH3 protein levels could change the mRNA m¹A/A ratio in different cell lines, although endogenous levels of ALKBH3 appeared to be diverse (Supplementary Figs. 14, 15 and 18). To further investigate whether ALKBH3 could work on m¹A in mRNA, we characterized the m¹A profile of the ALKBH3^{-/-} HEK293T cells. Application of m¹A-ID-seq to ALKBH3^{-/-} cells revealed 1,989 high-confidence m¹A peaks, twice the number identified in the wild-type (WT) cells (Fig. 3b,c and Supplementary Data Set 1). While the majority (~77%) of the peaks in the WT cells were also identified in ALKBH3^{-/-} cells, we observed 774 m¹A peaks that were specific to the ALKBH3^{-/-} samples (Fig. 3b). To examine the substrate specificity of ALKBH3, we analyzed the distribution pattern and sequence motifs for the 774 ALKBH3^{-/-}-specific peaks. These peaks were enriched in 5' UTRs and also contained sequence motifs similar to those found in the WT cells (Fig. 3d,e and Supplementary Fig. 16). Whereas the m¹A pattern and motifs in the WT cells are the combined results of methylation and demethylation processes, those of the 774 ALKBH3^{-/-}-specific peaks reflect the preference of the putative methyltransferase(s). Additionally, these observations also suggested that ALKBH3 has minimal sequence preference and act globally in the transcriptome. Taken together, these results showed that m¹A methylation is reversible and that m¹A in mRNA is an *in vivo* substrate of ALKBH3.

m¹A is dynamically regulated by stress conditions

Because emerging evidence shows that RNA methylation can change in response to stimuli³², we then sought to determine whether m¹A methylation in mRNA can be dynamically regulated. We characterized the m¹A patterns of cells under serum starvation conditions and H₂O₂ treatment and compared these to the m¹A profiles of untreated controls (Fig. 4a–c and Supplementary Fig. 17a,b).

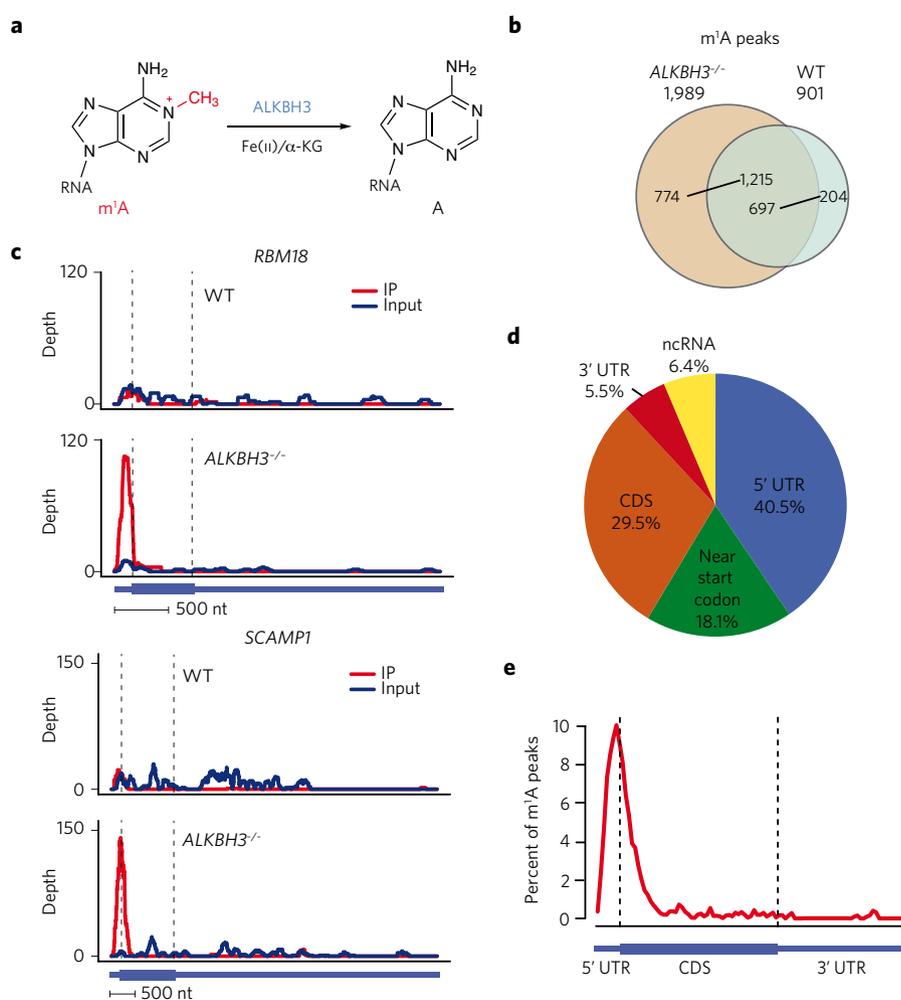


Figure 3 | m¹A in RNA is reversible by ALKBH3. (a) Oxidative demethylation of m¹A mediated by ALKBH3. (b) The majority of m¹A peaks identified in wild-type (WT) cells fall within those of ALKBH3^{-/-} cells. 1,215 peaks are shared by both ALKBH3^{-/-} replicates and at least one WT replicate, and 697 peaks are shared by both WT replicates and at least one ALKBH3^{-/-} replicate. 774 peaks are specific to ALKBH3^{-/-} cells. (c) Representative views of specific m¹A peaks in ALKBH3^{-/-} cells. (d) Pie chart showing the percentage of ALKBH3^{-/-}-specific m¹A peaks in each nonoverlapping mRNA segment. (e) Distribution of ALKBH3^{-/-}-specific m¹A peaks across mRNA segments.

476 and 208 inducible m¹A peaks were identified for serum-starvation conditions and H₂O₂ treatment, respectively (Fig. 4a–c and Supplementary Data Set 1); the majority (459 and 191, respectively) of these peaks appeared to be stress specific, while a small number (17) were common to both stimuli (Fig. 4a–c). Interestingly, these inducible m¹A peaks also exhibited enrichment in the 5' UTR and near start codon regions, recapitulating the distribution pattern in the WT cells ('constitutive peaks') (Supplementary Fig. 17c,d). Of note, we ruled out differences in mRNA expression as an explanation for the inducible m¹A methylation (Supplementary Fig. 17e,f). In all, these data showed that m¹A methylation in mRNA is dynamically regulated by different stress conditions.

DISCUSSION

m¹A is among the most prevalent post-transcriptional modifications in ncRNA. In this study, we have shown that m¹A is abundant in human mRNA and applied the m¹A-ID-seq technique for the first transcriptome-wide characterizations of m¹A. One of the most surprising features of m¹A revealed by m¹A-ID-seq is its enrichment in the 5' UTR of mRNA transcripts. The 5' UTR contains various elements involved in translation regulation, including high GC content and secondary structures³⁶. Indeed, regions harboring m¹A sites have

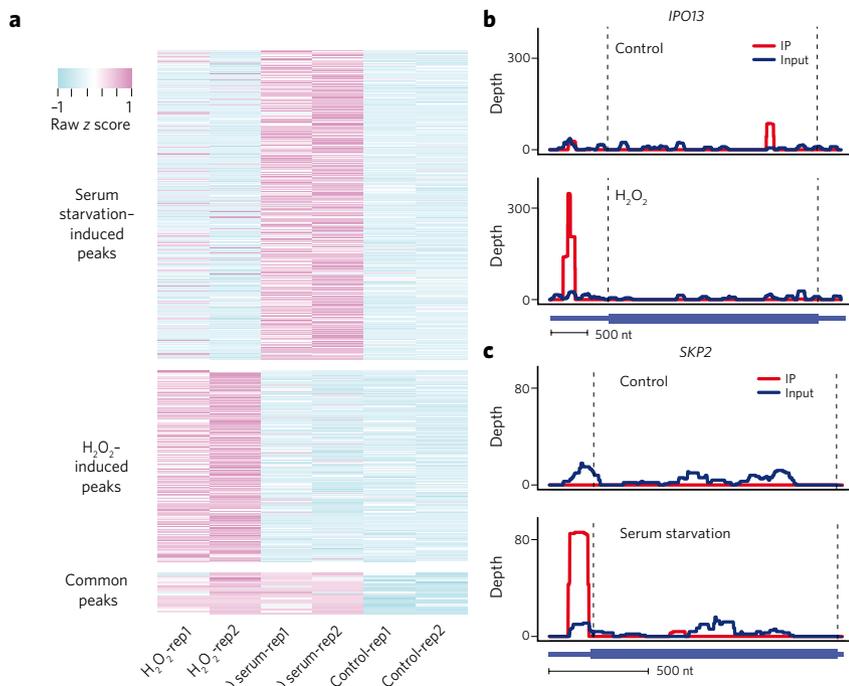


Figure 4 | m¹A is dynamically regulated by different stimuli. (a) Stress-induced m¹A sites. Inducible peaks were grouped into serum starvation-specific, H₂O₂-specific or common peaks. Fold change of each peak was normalized and is denoted with a z score. 'rep1' and 'rep2' denote two biological replicates in each instance. (b) Representative views of a H₂O₂-specific m¹A peak in human mRNA. (c) Representative views of a serum starvation-specific m¹A peak.

higher GC content and greater tendency to form stable secondary structures as compared to the non-m¹A controls. Very recent studies have shown that heat shock-induced m⁶A in the 5' UTR of newly transcribed mRNAs promotes cap-independent translation^{37,38}. Thus, whether or not m¹A methylation can also affect translation would be of functional interests for future investigations. It should be noted that a recent study³⁹ (published in *Nature* concurrently with this paper) shows that m¹A associates with translation initiation sites and correlates positively with protein production. Ideally, sequencing methods that both allow single-base resolution and provide quantitative information about m¹A modification would further aid future functional studies. In contrast to m¹A, m⁶A is preferentially found in 3' UTRs; the presence of m⁶A in the stem region is also believed to destabilize RNA hairpin structures^{40,41}. Additionally, it has recently been shown that the presence of m⁶A can influence RNA fate in various ways: selective binding of m⁶A-containing RNAs by the 'reader' protein YTHDF2 results in mRNA degradation, whereas recognition by YTHDF1 promotes translation^{42,43}; and within the nucleus, the m⁶A-binding HNRNPA2B1 protein promotes primary miRNA processing⁴⁴. Therefore, potential reader proteins that selectively recognize m¹A are also highly desirable as aids to a full understanding of the biological roles of m¹A.

We also showed that m¹A is reversible by ALKBH3, an AlkB-family DNA/RNA demethylase. Initially, ALKBH3 was shown to be a DNA repair enzyme guarding the genome against alkylation damage^{21,22}. It is confined mainly to the nucleoplasm and, to a lesser extent, the cytoplasm; the DNA repair function of ALKBH3 in cancer cells is also dependent on ASCC3, a 3'-5' DNA helicase²⁴. These observations do not, however, exclude the possibility that ALKBH3 could act on RNA and perform additional cellular functions. As a matter of fact, we have shown in this study that a large fraction of m¹A peaks are shared between WT and ALKBH3^{-/-} cells and that hundreds of specific mRNA m¹A peaks are observed in ALKBH3^{-/-} HEK293T cells; instead of being randomly distributed throughout

the transcriptome (as one would expect for nucleic acid damage), these ALKBH3^{-/-}-specific m¹A sites are enriched in the 5' UTR and also show consensus sequences. Such observations are highly suggestive of a regulatory role of ALKBH3 that is distinct from its established function as a repair enzyme.

Our discovery that m¹A is reversible also indicates that m¹A methylation may be involved in regulatory processes. Methylation of DNA and histones, being well-known epigenetic marks, are subject to the so-called active demethylation pathways. In fact, m⁶A is reversible by FTO and ALKBH5 (refs. 27,28), two homologs of ALKBH3. It remains to be seen whether other m¹A demethylases are present in human cells. Additionally, the methylation machinery of m¹A—the 'writer'—has yet to be characterized. m⁶A in mRNA is installed by a dedicated methyltransferase complex involving METTL3, METTL14 and WTAP^{45–47}; whether or not specific methylation machinery could also exist for m¹A is still open to future study. Alternatively, known m¹A methyltransferases may also be worth investigating. The interplay and regulation between these different components will need to be investigated in future to fully establish m¹A as a regulatory modification in the human transcriptome.

In summary, our study demonstrates that m¹A is a widespread RNA modification in human cells and also provides tools for functional studies of m¹A methylation. Analogous to histone code in which different histone modifications regulate transcription of genetic information, the discovery of m¹A in mRNA adds to the expanding repertoire of mRNA modifications and hence regulation through epitranscriptomic marks.

Received 20 October 2015; accepted 2 February 2016; published online 10 February 2016

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Accession codes. Sequencing data have been deposited into the Gene Expression Omnibus (GEO) under the accession code [GSE73941](#).

References

- Machnicka, M.A. *et al.* MODOMICS: a database of RNA modification pathways—2013 update. *Nucleic Acids Res.* **41**, D262–D267 (2013).
- Ozanick, S., Krecic, A., Andersland, J. & Anderson, J.T. The bipartite structure of the tRNA m1A58 methyltransferase from *S. cerevisiae* is conserved in humans. *RNA* **11**, 1281–1290 (2005).
- Chujo, T. & Suzuki, T. Trmt61B is a methyltransferase responsible for 1-methyladenosine at position 58 of human mitochondrial tRNAs. *RNA* **18**, 2269–2276 (2012).
- Schevitz, R.W. *et al.* Crystal structure of a eukaryotic initiator tRNA. *Nature* **278**, 188–190 (1979).
- Saikia, M., Fu, Y., Pavon-Eternod, M., He, C. & Pan, T. Genome-wide analysis of N¹-methyl-adenosine modification in human tRNAs. *RNA* **16**, 1317–1327 (2010).
- Helm, M. *et al.* The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. *Nucleic Acids Res.* **26**, 1636–1643 (1998).
- Jühling, F. *et al.* tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.* **37**, D159–D162 (2009).
- Suzuki, T., Nagao, A. & Suzuki, T. Human mitochondrial tRNAs: biogenesis,

- function, structural aspects, and diseases. *Annu. Rev. Genet.* **45**, 299–329 (2011).
9. Sharma, S., Watzinger, P., Kötter, P. & Entian, K.D. Identification of a novel methyltransferase, Bmt2, responsible for the N¹-methyl-adenosine base modification of 25S rRNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **41**, 5428–5443 (2013).
 10. Srivastava, R. & Gopinathan, K.P. Ribosomal RNA methylation in *Mycobacterium smegmatis* SN2. *Biochem. Int.* **15**, 1179–1188 (1987).
 11. Peifer, C. *et al.* Yeast Rrp8p, a novel methyltransferase responsible for m¹A 645 base modification of 25S rRNA. *Nucleic Acids Res.* **41**, 1151–1163 (2013).
 12. Chan, C.T. *et al.* A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. *PLoS Genet.* **6**, e1001247 (2010).
 13. Helm, M. & Alfonzo, J.D. Posttranscriptional RNA modifications: playing metabolic games in a cell's chemical Legoland. *Chem. Biol.* **21**, 174–185 (2014).
 14. Ballesta, J.P. & Cundliffe, E. Site-specific methylation of 16S rRNA caused by p_{ct}, a pactamycin resistance determinant from the producing organism, *Streptomyces pactum*. *J. Bacteriol.* **173**, 7213–7218 (1991).
 15. Motorin, Y., Muller, S., Behm-Ansmant, I. & Branlant, C. Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.* **425**, 21–53 (2007).
 16. Behm-Ansmant, I., Helm, M. & Motorin, Y. Use of specific chemical reagents for detection of modified nucleotides in RNA. *J. Nucleic Acids* **2011**, 408053 (2011).
 17. Hodgkinson, A. *et al.* High-resolution genomic analysis of human mitochondrial RNA sequence variation. *Science* **344**, 413–415 (2014).
 18. Hauenschild, R. *et al.* The reverse transcription signature of N¹-methyladenosine in RNA-Seq is sequence dependent. *Nucleic Acids Res.* **43**, 9950–9964 (2015).
 19. Falnes, P.O., Johansen, R.F. & Seeberg, E. AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* **419**, 178–182 (2002).
 20. Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T. & Sedgwick, B. Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature* **419**, 174–178 (2002).
 21. Duncan, T. *et al.* Reversal of DNA alkylation damage by two human dioxygenases. *Proc. Natl. Acad. Sci. USA* **99**, 16660–16665 (2002).
 22. Aas, P.A. *et al.* Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature* **421**, 859–863 (2003).
 23. Ougland, R. *et al.* AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Mol. Cell* **16**, 107–116 (2004).
 24. Dango, S. *et al.* DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation. *Mol. Cell* **44**, 373–384 (2011).
 25. Li, X. *et al.* Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* **11**, 592–597 (2015).
 26. Macon, J.B. & Wolfenden, R. 1-Methyladenosine. Dimroth rearrangement and reversible reduction. *Biochemistry* **7**, 3453–3458 (1968).
 27. Jia, G. *et al.* N⁶-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **7**, 885–887 (2011).
 28. Zheng, G. *et al.* ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **49**, 18–29 (2013).
 29. Maden, B.E. The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **39**, 241–303 (1990).
 30. Zheng, G. *et al.* Efficient and quantitative high-throughput tRNA sequencing. *Nat. Methods* **12**, 835–837 (2015).
 31. Cozen, A.E. *et al.* ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. *Nat. Methods* **12**, 879–884 (2015).
 32. Dominissini, D. *et al.* Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature* **485**, 201–206 (2012).
 33. Meyer, K.D. *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–1646 (2012).
 34. Ke, S. *et al.* A majority of m⁶A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* **29**, 2037–2053 (2015).
 35. Linder, B. *et al.* Single-nucleotide-resolution mapping of m⁶A and m⁶Am throughout the transcriptome. *Nat. Methods* **12**, 767–772 (2015).
 36. Araujo, P.R. *et al.* Before it gets started: regulating translation at the 5' UTR. *Comp. Funct. Genomics* **2012**, 475731 (2012).
 37. Zhou, J. *et al.* Dynamic m⁶A mRNA methylation directs translational control of heat shock response. *Nature* **526**, 591–594 (2015).
 38. Meyer, K.D. *et al.* 5' UTR m⁶A promotes cap-independent translation. *Cell* **163**, 999–1010 (2015).
 39. Dominissini, D. *et al.* The dynamic N¹-methyladenosine methylome in eukaryotic messenger RNA. *Nature* <http://dx.doi.org/10.1038/nature16998> (2016).
 40. Liu, N. *et al.* N⁶-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564 (2015).
 41. Roost, C. *et al.* Structure and thermodynamics of N⁶-methyladenosine in RNA: a spring-loaded base modification. *J. Am. Chem. Soc.* **137**, 2107–2115 (2015).
 42. Wang, X. *et al.* N⁶-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–120 (2014).
 43. Wang, X. *et al.* N⁶-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388–1399 (2015).
 44. Alarcón, C.R. *et al.* HNRNPA2B1 is a mediator of m⁶A-dependent nuclear RNA processing events. *Cell* **162**, 1299–1308 (2015).
 45. Liu, J. *et al.* A METTL3-METTL14 complex mediates mammalian nuclear RNA N⁶-adenosine methylation. *Nat. Chem. Biol.* **10**, 93–95 (2014).
 46. Schwartz, S. *et al.* Perturbation of m⁶A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Reports* **8**, 284–296 (2014).
 47. Ping, X.L. *et al.* Mammalian WTAP is a regulatory subunit of the RNA N⁶-methyladenosine methyltransferase. *Cell Res.* **24**, 177–189 (2014).

Acknowledgments

The authors would like to thank G. Jia and C. Zhang for assistance with recombinant ALKBH3 protein. This work was supported by the National Basic Research Foundation of China (no. 2014CB964900) and the National Natural Science Foundation of China (nos. 21472009 and 21522201).

Author contributions

X.L., X.X. and C.Y. conceived the project, designed the experiments and wrote the manuscript. X.L. performed the experiments with the help of K.W., L.W. and X.S.; X.X. designed and performed the bioinformatics analysis; S.M. participated in discussion. All authors commented on and approved the paper.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Any supplementary information, chemical compound information and source data are available in the [online version of the paper](#). Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>. Correspondence and requests for materials should be addressed to C.Y.

ONLINE METHODS

Cell culture. U2OS, HEK293T, WPMY-1, DU145, LNCap, and A549 were used for m¹A analysis in this study. U2OS, HEK293T, WPMY-1, DU145, and LNCap were maintained in DMEM medium (Gibco) supplemented with 10% FBS and 1% penicillin/streptomycin. A549 was maintained in RPMI 1640 medium (Gibco) supplemented with 10% FBS and 1% penicillin/streptomycin. Mycoplasma contamination tests were performed routinely using the GMyc-PCR Mycoplasma Test Kit from YEASEN (cat. #40601).

Antibodies. Monoclonal mouse anti-m¹A antibody was purchased from MBL (D345-3) (1:2,000 dilution for dot blot assays). Polyclonal rabbit anti-m⁶A antibody was purchased from Synaptic Systems (202 003) (dissolved to 1 mg/μL using RNase-free water and 1:2,000 dilution for dot blot assays). Polyclonal rabbit anti-ALKBH3 antibody was purchased from Millipore (09-882) (1:2,000 dilution for western blot assays). Monoclonal mouse anti-β actin antibody was purchased from CWBiotech (CW0096) (1:5,000 dilution for western blot assays). The secondary antibodies used are anti-mouse-IgG-HRP (CW0102; CWBiotech) (1:8,000 dilution for western blot assays) and anti-rabbit-IgG-HRP (CW0103; CWBiotech) (1:8,000 dilution for western blot assays).

RNA isolation. Total RNA was isolated from cell lines with TRIzol according to the manufacturer's instructions (Invitrogen). An additional DNase I treatment step was added to avoid DNA contamination. For poly(A)⁺ RNA purification, small RNA was depleted first using MEGAclear Transcription Clean-Up Kit (Invitrogen) and then by two successive rounds of poly(A)⁺ selection using oligo(dT)₂₅ Dynabeads (Invitrogen).

Quantification of m¹A, m¹dA, i⁶A and m⁶A by LC-MS/MS. 200 ng RNA or ssDNA was digested into nucleosides by 0.5 U nuclease P1 (Sigma, N8630) in 20 μL buffer containing 10 mM ammonium acetate, pH 5.3 at 42 °C for 6 h followed by the addition of 2.5 μL 0.5 M MES buffer, pH 6.5 and 0.5 U alkaline phosphatase (Sigma, P4252). The mixture was incubated at 37 °C for another 6 h and diluted to 50 μL, and 5 μL of the solution was injected into LC-MS/MS. The nucleosides were separated by ultra-performance liquid chromatography on a C18 column, and then detected by triple-quadrupole mass spectrometer (AB SCIEX QTRAP 5500) in the positive ion multiple reaction-monitoring (MRM) mode. The mass transitions of *m/z* 282.0 to 150.1 (m¹A), *m/z* 282.0 to 150.1 (m⁶A), *m/z* 336.0 to 136.1 (i⁶A), *m/z* 268.0 to 136.0 (A), *m/z* 266.0 to 150.1 (m¹dA), and *m/z* 252.1 to 136 (dA) were monitored and recorded. A series of concentrations of pure authentic nucleoside standards (A, from 50 nM to 2,000 nM; m¹A, from 0.1 nM to 50 nM; m⁶A, from 0.1 nM to 50 nM; and i⁶A, from 0.01 nM to 50 nM) were run for every batch of experiments to obtain their corresponding stand curves. Concentrations of nucleosides in mRNA samples were deduced by fitting the signal intensities into the stand curves. The ratios of m¹A/A, m⁶A/A, i⁶A/A and m¹dA to dA were subsequently calculated.

Dot blot assays. Three synthetic m¹A, A and m⁶A oligos (5'-UGUCCUUAAX_UUAGGACUU-3', X = m¹A, A or m⁶A) were used to test the specificity of the m¹A antibody used in this study. Fivefold serial dilution RNAs were denatured and spotted on the Amersham Hybond-N⁺ membrane followed by UV crosslinking. The membranes were blocked with 5% non-fat milk in 0.1% PBST at room temperature for 1 h, then probed with anti-m¹A antibody (1:2,000) or anti-m⁶A antibody (1:2,000) in 5% non-fat milk at room temperature for 1 h and then washed three times with 0.1% PBST. Then the membranes were incubated with anti-mouse-IgG-HRP (for anti-m¹A antibody) or anti-rabbit-IgG-HRP (for anti-m⁶A antibody). Blots were washed four times with 0.1% PBST and developed with the ECL Western Blotting Detection Kit (Pierce).

m¹A immunoprecipitation. We tested the ability of m¹A antibody to enrich m¹A-containing RNA by LC-MS/MS and qRT-PCR analyses. For LC-MS/MS analysis: 50 mg poly(A)⁺ RNA was fragmented into ~150 nt using magnesium RNA fragmentation buffer (NEB). Fragmented RNA (as input) was denatured and incubated with 5 mg anti-m¹A antibody in IPP buffer (150 mM NaCl, 0.1% NP-40, 10 mM Tris, pH 7.4) at 4 °C overnight. 50 μL Protein A/G UltraLink Resin (Pierce) was added to the RNA antibody mixture and incubated for additional 2 h at 4 °C. The RNA in the supernatant was purified by phenol chloroform extraction and ethanol precipitation (as flowthrough). Resins were washed with IPP buffer and RNA was isolated by phenol chloroform extraction and ethanol precipitation (as IP, immunoprecipitated).

For qRT-PCR analyses, total RNA was isolated from HEK293T with TRIzol (Invitrogen) followed by an additional DNase I treatment to avoid DNA contamination. Small RNA depletion was performed using MEGAclear Transcription Clean-Up Kit (Invitrogen). 10 mg total RNA was fragmented into ~150 nt using magnesium RNA fragmentation buffer (NEB) and concentrated by ethanol precipitation. Fragmented RNA (as input) was denatured and incubated with 2 mg anti-m¹A antibody in IPP buffer (150 mM NaCl, 0.1% NP-40, 10 mM Tris, pH 7.4) at 4 °C overnight. 20 μL Protein A/G UltraLink Resin (Pierce) was added to the RNA antibody mixture and incubated for additional 2 h at 4 °C. Resins were washed with IPP buffer and RNA was eluted from resins with 3 mg/mL N¹-methyladenosine (Berry & Associates) in IPP buffer and purified by phenol chloroform extraction and ethanol

precipitation. Input and immunoprecipitated RNAs were reverse transcribed into cDNA using RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific) and quantified by qPCR using SYBR GREEN mix (Takara) on ABI ViiA 7 real-time PCR system. The qPCR primers for methylated sequence was designed adjacent and downstream m¹A1322 in 28S (FWD: 5'-CACGGACCAAGGAGTCTAACAC-3', RVS: 5'-GCCTTCACCTTCATTGCG-3'). 18S rRNA which doesn't contain m¹A was used as an unmethylated control and the qPCR primers are as follows: FWD: 5'-ACCCTTGAACCCCATTCGTGA-3', RVS: 5'-GCCTCACTAAACCATCCAATCGG-3'). The enrichment of immunoprecipitation was calculated relative to input sample.

m¹A-ID-seq. 50 mg poly(A)⁺ RNA was fragmented into ~150 nt using magnesium RNA fragmentation buffer (NEB). Fragmented poly(A)⁺ RNA was desalted and concentrated by ethanol precipitation. Fragmented poly(A)⁺ RNA was denatured at 65 °C for 5 min, followed by chilling on ice and then incubated with 5 mg anti-m¹A antibody in IPP buffer (150 mM NaCl, 0.1% NP-40, 10 mM Tris, pH 7.4) at 4 °C overnight. 50 μL Protein A/G UltraLink Resin (Pierce) was washed three times and resuspended in 200 μL IPP buffer and incubated with the RNA antibody mixture for additional 2 h at 4 °C. Resins were washed with IPP buffer, once with low salt buffer (75 mM NaCl, 0.1% NP-40, 10 mM Tris, pH 7.4), and twice with high salt buffer (200 mM NaCl, 0.1% NP-40, 10 mM Tris, pH 7.4) and once with TEN buffer (10 mM Tris, pH 7.4, 1 mM EDTA, pH 8.0, 0.05% NP-40). RNA containing m¹A was eluted from resins with 3 mg/mL N¹-methyladenosine (Berry & Associates) in IPP buffer and purified by phenol chloroform extraction and ethanol precipitation.

10 ng (~0.2 pmol) of the eluted m¹A-containing RNA was denatured at 65 °C for 5 min, chilled on ice and then demethylated in a 20 μL demethylation mixture containing 0.4 pmol purified AlkB. The optimized reaction buffer contained 50 mM MES, pH 6.5, 283 μM of (NH₄)₂Fe(SO₄)₂·6H₂O, 300 μM 2-ketoglutarate, 2 mM L-ascorbic acid, 0.4 U/μL RNase inhibitor. The demethylation reaction was incubated 2 h at 37 °C and quenched by the addition of 5 mM EDTA. Demethylated RNA was purified by phenol chloroform extraction and ethanol precipitation.

Fragmented poly(A)⁺ RNA (as "input"), immunoprecipitated RNA 11 and demethylated immunoprecipitated RNA [as "(+) demethylase" samples] were subjected to library construction. Reverse transcription was performed using AMV Reverse Transcriptase (NEB) and random primers. The reverse transcription was carried out in reverse transcription buffer (50 mM Tris, pH 7.0, 50 mM KCl, 10 mM DTT) with dNTPs (0.2 mM each final) and MgCl₂ (10 mM final) at 42 °C for 1 h, which was also optimized to maximize cDNA truncations at the site of m¹A methylation. The library construction steps were performed using NEBNext Ultra RNA Library Prep Kit for Illumina (E7530) (unstranded) according to the manufacturer's instructions. During the library preparation process, random primers were used both in the reverse transcription step and in the second strand synthesis step. Hence, the exact positions of RT stops by m¹A during the first strand synthesis are considered to be lost during the process of second strand synthesis and end repair. All libraries were sequenced on Illumina HiSeq 2500 with paired-end 2 × 100 bp read length.

Validation of m¹A peaks. 10 mg poly(A)⁺ RNA was divided into two parts: one part was fragmented using magnesium RNA fragmentation buffer (from NEB, hence regular fragmentation condition) as the untreated sample, and the other part was treated with 0.1 M Na₂CO₃ / NaHCO₃, pH 10.2 at 65 °C for 3 h as the alkaline treated sample. Under alkaline treatment, the endogenous m¹A were converted into m⁶A.

Immunoprecipitation of m⁶A-containing RNA was performed to the untreated and alkaline treated samples in parallel. Samples were denatured and incubated with 2.5 mg anti-m⁶A antibody in IPP buffer (150 mM NaCl, 0.1% NP-40, 10 mM Tris, pH 7.4) at 4 °C overnight. 20 μL Protein G Dynabeads (Invitrogen) were added to the RNA antibody mixture and incubated for additional 2 h at 4 °C. Beads were washed with IPP buffer, and RNA was eluted from beads with 2.5 mg/mL N⁶-methyladenosine in IPP buffer and purified by phenol chloroform extraction and ethanol precipitation. Input and immunoprecipitated RNAs were analyzed by qRT-PCR. The m⁶A-IP/input ratio of target regions in untreated and alkaline treated samples were calculated, respectively. And the m⁶A-IP/input ratio of alkaline treated samples was normalized by that of the untreated samples, giving fold change (y-axis in **Supplementary Fig. 7**). The primers used for qRT-PCR were listed in **Supplementary Table 1**.

Furthermore, m¹A-IP was also performed for the untreated and alkaline treated samples. The procedure was similar to the above except for the use of anti-m¹A antibody for immunoprecipitation and N¹-methyladenosine for elution. m¹A-IP/input ratios were also calculated to give fold change.

Cloning, expression and purification of AlkB and ALKBH3. A truncated AlkB with deletion of the N-terminal 11 amino acids was cloned into pET30a (Novagen), transformed to *E. coli* BL21(DE3), grown in LB medium at 37 °C until the OD₆₀₀ reached 0.6–0.8 and incubated at 30 °C for additional 4 h with the addition of 1 mM IPTG. The proteins were purified as reported with slight modifications⁴⁸. Proteins were purified using Ni-NTA chromatography (GE Healthcare) and gel-filtration chromatography (Superdex 200, GE Healthcare) followed by Mono-Q anion exchange chromatography

(GE Healthcare). Such purification procedure effectively avoided RNA contamination from *E. coli* (expression host).

A truncated ALKBH3 with deletion of the N-terminal 69 amino acids was cloned into pET30a (Novagen) and transformed to *E. coli* BL21(DE3) followed by growing at 37 °C in M9 medium until the OD₆₀₀ reached 0.6–0.8 and incubating at 30 °C for additional 4 h with the addition of 1 mM IPTG. Proteins were purified using Ni-NTA chromatography (GE Healthcare) and gel-filtration chromatography (Superdex 200, GE Healthcare) as reported²¹.

Biochemistry assay of ALKBH3 activity *in vitro*. The demethylation activity was performed in a 20 μL mixture containing 10 μM ssRNA or ssDNA containing a central N¹-methyladenosine, 1 μM ALKBH3, 50 mM Tris-Cl, pH 7.4, 283 μM (NH₄)₂Fe(SO₄)₂·6H₂O, 300 μM 2-ketoglutarate, 2 mM L-ascorbic acid, 50 μg/mL BSA, and 1 U/μL RNase inhibitor. The demethylation reaction was incubated at 37 °C and quenched by the addition of 5 mM EDTA followed by heating at 95 °C. RNA or DNA digested by nuclease P1 and alkaline phosphatase were analyzed by LC-MS/MS (AB SCIEX QTRAP 5500).

Generation of ALKBH3^{-/-} 293T cells with CRISPR-Cas9. The plasmid PX330 containing two expression cassettes, hSpCas9 and the chimeric guide RNA was used in this study. The guide RNA was designed according to the website <http://crispr.mit.edu/>, and the target sequence of ALKBH3 is GAGCCCGAGTTTCAGGGAGCC. The plasmid containing the guide RNA sequence was transfected into HEK293T cells using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Cells were diluted and seeded into 96-well dish 48 h after transfection according to the limiting dilution protocol. After 10 days, single colonies were picked and detected by western blot.

shRNA knockdown and overexpression of ALKBH3. The oligos targeting ALKBH3 were annealed and cloned into the pLKO vector according to the TRC shRNA library protocol (<http://www.broadinstitute.org/rnai/public/>). The oligo sequences were listed: FWD: CCGGCGCACATTTGAGATGAGAAAGCTCGAGCTTCTCATCTCAAATGTCCGTTTTTG; RVS: AATTCAAAAACGCACATTTGAGATGAGAAAGCTCGAGCTTCTCATCTCAAATGTGCG. A scrambled shRNA was used as the mock control. Lentiviruses were packaged by co-transfecting HEK293T cells with pLKO, pCMV-dR8.91 and VSV-G plasmids, following the instructions from Broad Institute. The supernatants from transfected cells were harvested after 2 days and used to infect WPMY-1 and DU145 cells followed by puromycin selection for 2 days. Knockdown efficiency was verified by western blot.

ALKBH3 was cloned into pcDNA3.1 using the following primers: FWD: ATGAGAGAAAAAGACGGCGAG; RVS: CCAGGGTGCCCTCGAG. The plasmid was transfected into HEK293T, WPMY-1 and DU145 cells with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. The empty pcDNA3.1 vector was used as the mock control.

Stress conditions. Different stress conditions were used as follows: (1) Drug treatment, cells were incubated with hydrogen peroxide (50 μM) for 30 min, or cycloheximide (40 mg/mL) for 4 h, or IFN-γ (200 ng/mL) for 6 h respectively; (2) Heat shock, cells were incubated at 42 °C for 8 h; (3) Starvation, cells were washed with PBS twice before maintaining with serum free DMEM medium (serum starvation) or low glucose DMEM (glucose starvation) or serum free low glucose DMEM (serum and glucose starvation) 16 h before RNA isolation.

Pre-processing and alignment of reads. Raw sequencing data were subjected to Trim_galore for quality control and adaptor trimming (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The minimum quality threshold was set to 20, and reads shorter than 25 after trimming were discarded. Reference transcriptome were prepared based on the University of California, Santa Cruz (UCSC) RefSeq table in human (hg19). The redundant sequences with the same RefSeq ID were removed. The processed reads were mapped to the reference transcriptome using HISAT⁴⁹. Only those proper-paired mapped reads with no more than two mismatches were adopted for subsequent analysis. All performances related to the processing of sam/bam file were done by using SAMtools (<http://samtools.sourceforge.net/>).

Transcriptome-wide identification of m¹A peaks. Aligned reads were subjected to model-based analysis of ChIP-seq (MACS2) algorithm for peak calling. Peak calling was performed using “(-) demethylase” samples and “input” samples. The ‘effective genome size’ parameter was set to the calculated transcriptome size (5 × 10⁶). In order to identify the high-confidence m¹A peaks, we required the peaks to (1) be present in both replicates and (2) in either replicates, have a *q* < 10⁻¹⁰ and enrichment fold > 3. Only peaks that can meet both requirements were used for further analysis. Intersection between bed files was performed using BEDTools (<http://bedtools.readthedocs.org/en/latest/>).

Definition and calculations of the demethylase-sensitivity score (DS score). In order to more accurately extract the regions containing m¹A, we defined a “demethylase-sensitivity score” (DS score) to evaluate the increase in normalized

sequencing depth (due to the removal of m¹A by AlkB and hence accumulated full-length RT products) between the (+) and (-) demethylase samples. First, m¹A peaks (defined in the section of “Transcriptome-wide identification of m¹A peaks”) and the corresponding 50 nt 5' to each peak were used for DS score calculations (taking into consideration that m¹A could locate 5' to the peaks due to RT stops). Second, sequencing depth of each nucleotide within a peak were added up to obtain the total depth for this peak, which was denoted as “Area(Peak_{sample})”. Third, within each peak, 10-nt sliding windows covering the entire peak—with an increment of 5 nt—were defined, and sequencing depth for each sliding window were calculated, denoted as “Area(Window_{sample})”. Fourthly, sequencing depth of one sliding window was divided by the total depth of the corresponding peak to obtain the normalized depth of the sliding window, denoted as “Norm(Depth_{sample})”. Hence, the normalized depth of one particular sliding window within the “(-) demethylase” and “(+) demethylase” samples can be calculated:

$$Norm(Depth_{(+)demethylase}) = \frac{Area(Window_{(+)demethylase})}{Area(Peak_{(+)demethylase})}$$

$$Norm(Depth_{(-)demethylase}) = \frac{Area(Window_{(-)demethylase})}{Area(Peak_{(-)demethylase})}$$

Finally, the DS score of this sliding window was obtained by subtracting the two normalized depth:

$$DS\ score = (Norm(Depth_{(+)demethylase}) - Norm(Depth_{(-)demethylase})) \times 100$$

To better illustrate how the DS score is calculated, we took the validated m¹A peak in SUMO3 as an example. The peak called by MACS2 ranges from nucleotide position 82 to 192, and thus the first 10-nt window starts from 32–41. The total depths of the peak (nt 32–192) were then calculated for the (+) and (-) demethylase samples, respectively. The depths for different sliding windows were also calculated. Using sliding window 122–131 as an example, the normalized depths of sliding window 122–131 were 0.144 and 0.089 for the (+) and (-) demethylase samples, respectively. Finally, the DS score of this sliding window was 5.51. Windows sensitive to demethylase treatment would receive a positive DS score, while windows distal to m¹A sites should theoretically receive a DS score below or around zero.

Identifications of demethylase-sensitive regions within called m¹A peaks. The DS scores of sliding windows were used to identify the demethylase-sensitive regions. The sliding window with the highest DS score within one m¹A peak was first selected, and extended toward both 5' and 3' directions. The extension was terminated when both of the following two criteria were met: (i) DS score of the current sliding window was less than 0.3 fold of the highest DS score within the peak; (ii) the sliding window next to the current one has a DS score less than 0.3 fold of the highest DS score, or it was out of the defined peak region. These continuous sliding windows then became the “demethylase-sensitive region.” The demethylase-sensitive region is expected to include the m¹A sites while exhibiting improved resolution in comparison to that of regular IP experiments. Peaks whose highest DS scores were less than 1.0 were excluded for further analyses. Demethylase-sensitive regions shorter than 60 nt were extended to 60 nt. The demethylase-sensitive regions were used for topology analyses in the manuscript. In the above-mentioned case of SUMO3, the DS scores are -1.16, 0.89 and 5.51 for the first sliding window (32–41), the first window with positive score (92–101) and the window with highest positive score (122–131), respectively. Based on such criteria, a DS-sensitive region (92–171) was then obtained for SUMO3, improving the resolution of m¹A detection from 111 nt to 80 nt.

Annotation of mRNA segments. Nonoverlapping mRNA segments “5' UTR”, “CDS” and “3' UTR” were used for m¹A peak analysis. Additionally, the region “near start codon” was included, defined as 25 nt upstream and 25 nt downstream the start codon. The location on the mRNA segments was performed using BEDTools, by calculating the intersection with the above-mentioned categories, which were based on annotation file of the transcripts. m¹A peaks found in non-coding RNAs were all classified as “ncRNA”.

Secondary structure analysis. RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) was used to evaluate the secondary structure of the 50 nt with highest DS scores in each peak. The control sequences, which are of the same length, were randomly selected from 5' UTR or “near start codon”. The minimum free energy (MFE) calculated by RNAfold was used to represent the stability of the folded sequence. Randomization was performed for 10 times to increase the reliability. The significance of MFE difference between the demethylase-sensitive regions and control sequences were evaluated using Mann-Whitney *U*-test. All statistical analyses were performed with built-in R scripts.

Defining *in vivo* targets of ALKBH3 and stress-induced peaks. In order to detect the reliable peaks containing m¹A induced by stress conditions or by ALKBH3 knockout, two stringent criteria were adopted: (i) the peak in both replicates of

H₂O₂/serum starvation/ALKBH3-KO had to fulfill the criteria of *q*-value <10⁻¹⁰ and enrichment fold >3; (ii) the peak could not have been reported in either replicate of the wide-type samples by MACS2.

Motif discovery and GO analysis. The MEME algorithm (<http://meme-suite.org/>) was used to *de novo* search for enriched motifs from the demethylase-sensitive regions. We used two sequence sets as controls to ensure the reliability of the identified motifs. The first one was randomly selected sequences from the 5' UTR or "near start codon"; the second one was shuffling the original sequences without changing the nucleotide

content. The motif search was performed with demethylase-sensitive regions.

All Gene Ontology (GO) analyses were performed using the DAVID web-based tool (<http://david.abcc.ncifcrf.gov/>).

48. Zhu, C. & Yi, C. Switching demethylation activities between AlkB family RNA/DNA demethylases through exchange of active-site residues. *Angew. Chem. Int. Edn Engl.* **53**, 3659–3662 (2014).
49. Kim, D., Langmead, B. & Salzberg, S.L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).