Nucleic Acids Research, 2020 **1**
doi: 10.1093/nar/gkaa692

m^{6}A-Atlas: a comprehensive knowledgebase for unraveling the *N*^{6}-methyladenosine (m^{6}A) epitranscriptome

Yujiao Tang[1,2,†], Kunqi Chen [1,3,*,†], Bowen Song[2,4,†], Jiongming Ma[1,†], Xiangyu Wu[1,3,†], Qingru Xu[1], Zhen Wei[1], Jionglong Su[4], Gang Liu[4], Rong Rong[1,2], Zhiliang Lu[1,2], João Pedro de Magalhães[3], Daniel J. Rigden[2] and Jia Meng [1,2,5]

[1]Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, China, [2]Institute of Systems, Molecular and Integrative Biology, University of Liverpool, L7 8TX Liverpool, UK, [3]Institute of Ageing & Chronic Disease, University of Liverpool, L7 8TX Liverpool, UK, [4]Department of Mathematical Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, China and [5]AI University Research Centre, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, China

Received June 20, 2020; Revised August 05, 2020; Editorial Decision August 06, 2020; Accepted August 09, 2020

ABSTRACT

***N*^{6}-Methyladenosine (m^{6}A) is the most prevalent RNA modification on mRNAs and lncRNAs. It plays a pivotal role during various biological processes and disease pathogenesis. We present here a comprehensive knowledgebase, m^{6}A-Atlas, for unraveling the m^{6}A epitranscriptome. Compared to existing databases, m^{6}A-Atlas features a high-confidence collection of 442 162 reliable m^{6}A sites identified from seven base-resolution technologies and the quantitative (rather than binary) epitranscriptome profiles estimated from 1363 high-throughput sequencing samples. It also offers novel features, such as; the conservation of m^{6}A sites among seven vertebrate species (including human, mouse and chimp), the m^{6}A epitranscriptomes of 10 virus species (including HIV, KSHV and DENV), the putative biological functions of individual m^{6}A sites predicted from epitranscriptome data, and the potential pathogenesis of m^{6}A sites inferred from disease-associated genetic mutations that can directly destroy m^{6}A directing sequence motifs. A user-friendly graphical user interface was constructed to support the query, visualization and sharing of the m^{6}A epitranscriptomes annotated with sites specifying their interaction with post-transcriptional machinery (RBP-binding, microRNA interaction and splicing sites) and interactively display the landscape of multiple RNA modifications. These resources provide fresh opportunities for unraveling the m^{6}A epitranscriptomes. m^{6}A-Atlas is freely accessible at: www.xjtlu.edu.cn/biologicalsciences/atlas.**

INTRODUCTION

A variety of chemical modifications are naturally decorated on cellular RNAs, modulating their biogenesis, stability and functions (1). To date, >150 types of RNA modifications have been identified (2), among which, *N*^{6}-methyladenosine (m^{6}A) is the most pervasive and the most intensively studied non-cap reversible marker present on eukaryotic messenger RNAs (mRNAs) and long non-coding RNAs (lncRNAs) (3). Recent studies suggest that m^{6}A plays a pivotal role during various biological processes including virus infection (4), stress (5), heat shock (6) and DNA damage (7), and regulates molecular functions such as RNA–protein interaction (8), RNA stability (9) and translation efficiency (10). Additionally, dysregulation of m^{6}A has been linked to a number of diseases including multiple cancers, such as leukemia and glioma (11).

A number of high-throughput experimental approaches have been developed for profiling the transcriptome-wide distribution of m^{6}A RNA modification, including, most notably, the antibody-based approach m^{6}A-seq (or MeRIP-seq) (12,13). With m^{6}A-seq, it is possible to identify condition-specific m^{6}A sites (14–16), quantify the m^{6}A methylation levels (17,18), or compare between experimental conditions (17–19). Despite the limits of m^{6}A-seq regarding the reproducibility, data quality and mediocre resolution (around 100bp) (20,21), this technology has been widely applied to characterize the m^{6}A epitranscriptome under various biological contexts in more than 30 organ-

*To whom correspondence should be addressed. Tel: +86 512 81880492; Fax: +86 512 88161899; Email: kunqi.chen@xjtlu.edu.cn
†The authors wish it to be known that, in their opinion, the first five authors should be regarded as Joint First Authors.

isms since its invention in 2012. Besides m^{6}A-seq, there are also recent techniques such as PA-m6A-seq (22), miCLIP (23) and m^{6}A-CLIP (24), that offer improved or even base-resolution epitranscriptome determination. However, these approaches report primarily the precise location of m^{6}A sites in physical, and are unsuitable for quantification of m^{6}A methylation levels.

To date, several bioinformatics websites and databases have been constructed aiming to properly collect, annotate, share and interpret the rapidly growing knowledge in RNA modifications (2,25–29). Among them, MODOMICS (2) is an RNA modification database concerning primarily the RNA chemical structure, corresponding biosynthetic pathways and RNA-modifying enzymes. RNAmod (29) can be used for visualizing, annotating and comparing RNA modification profiles in 21 species. MetDB (25) is the first epitranscriptome database collecting m^{6}A sites on mRNAs and lncRNAs (rather than small RNAs) reported from high-throughput sequencing approaches. The most recent release (version 2.0) of MetDB hosted m^{6}A sites in seven species collected from 185 m^{6}A-seq experiments. RMBase (26) is currently the most comprehensive epitranscriptome database containing ∼1 397 000 RNA modification sites among 13 species including m^{6}A and other RNA modifications such as m^{5}C and m^{1}A. Additionally, m6AVar collected disease-associated genetic mutations that may functions through their impact on reversible m^{6}A RNA methylation (30). These works addressed various aspects of RNA modifications, and together greatly improved our understanding of the epitranscriptome. Nevertheless, the m^{6}A site collections in existing epitranscriptome databases (MetDB and RMBase) suffer from the following two major limitations:

- **Limited reliability**: m6A-seq is today the most widely adopted experimental approach to profiling the m6A epitranscriptome, and is responsible for most available m6A data. Although m6A-seq offers only a mediocre resolution (∼100 bp), the precise location of the m6A sites may be narrowed down to base-resolution by searching for the m6A forming motif DRACH within the regions enriched with m6A signal (or m6A peaks). Most existing epitranscriptome databases, including MeTDB and RMBase, relied on this very simple strategy to enhance the resolution of m6A-seq. However, this strategy cannot differentiate random-occurring motifs from real m6A-containing ones, resulting in false positive predictions. For this reason, it is not surprising that both MeT-DB (426,544sites) and RMBase (477,452 sites) reported a very large number of transcriptome m6A sites in human, many of which may be false positive due to the possibility of the RRACH motif (where R = G or A, H = A, C or U, and the A stands for adenosine being modified to m6A) (31) being located close to a real m6A site (or within an m6A peak). In a previous study (32), we systematically evaluated the positional information in existing m6A databases highlighting major concerns regarding its reliability, with only limited AUROC (Area under Receiver Operating Characteristic) between 0.736 and 0.798. This is even less reliable than the classic sequence-based m6A site predictor SRAMP (33).
- **Binary profiles**: Condition-specific epitranscriptome profiles are available in MetDB (25) and REPIC (28). However, they collected only binary profiles, i.e. whether an m6A methylation site exists or not under a specific biological condition, but did not provide quantitative measurements with regards to the degree of methylation levels. Although prone to various artifacts and noise such as batch effect, GC content bias and varying efficiency of immunoprecipitation (18,20,21), the most abundant epitranscriptome m^{6}A-seq data does contain the necessary information for a rough quantification of the methylation levels for m^{6}A sites. This is often presented as the fold enrichment of reads in the IP samples compared to the input control samples. A number of statistical models have been developed based on the quantitative information contained in m^{6}A-seq for detecting and assessing the significance of m^{6}A sites (peaks) (15,16,34), or for comparing the changes in terms of the methylation levels in a case-control study (17,35,36). Clearly, a quantitative epitranscriptome profile with robust estimates of the methylation levels for every single m^{6}A site would provide more insights than the binary profiles hosted by existing epitranscriptome databases.

To address these limitations, we constructed m^{6}A-Altas, a comprehensive knowledgebase for unraveling the *N*^{6}-methyladenosine (m^{6}A) epitranscriptome. Compared to existing databases, m^{6}A-Atlas features a high-confidence collection of reliable m^{6}A sites identified from base-resolution technologies only (rather than non-base-resolution technique combined with motif analysis) and the quantitative (rather than binary) condition-specific epitranscriptome profiles estimated from a large number of high-throughput sequencing samples covering various tissues and cell lines. It also offers novel features, such as, the conservation of m^{6}A sites among seven vertebrate species (including human, mouse and chimp), the epitranscriptomes of 10 virus species (including HIV, KSHV and DENV) and their host cells, the putative biological functions of individual m^{6}A sites predicted from epitranscriptome data with the guilt-by-association principle, and the potential pathogenesis of human m^{6}A sites inferred from disease-associated genetic mutations that can destroy m^{6}A specifying sequence motif. A user-friendly graphical user interface was constructed to support the query, visualization and sharing of the m^{6}A epitranscriptomes annotated with potential sites relating to downstream processes (TF binding, RBP-binding, microRNA interaction and splicing sites) and the positions of other known RNA modifications. These resources provide researchers with new tools for unraveling the m^{6}A epitranscriptomes. The design of m^{6}A-Atlas is shown in Figure 1.

MATERIALS AND METHODS

High-confidence collection of m^{6}A sites

High-confidence m^{6}A sites were collected from 67 datasets generated by seven different base-resolution technologies, covering seven species, including human, mouse, rat, zebrafish, fly, Arabidopsis and yeast. The positional information was directly obtained from the original papers or relevant NCBI GEO datasets (37) (Supplementary Table S1).

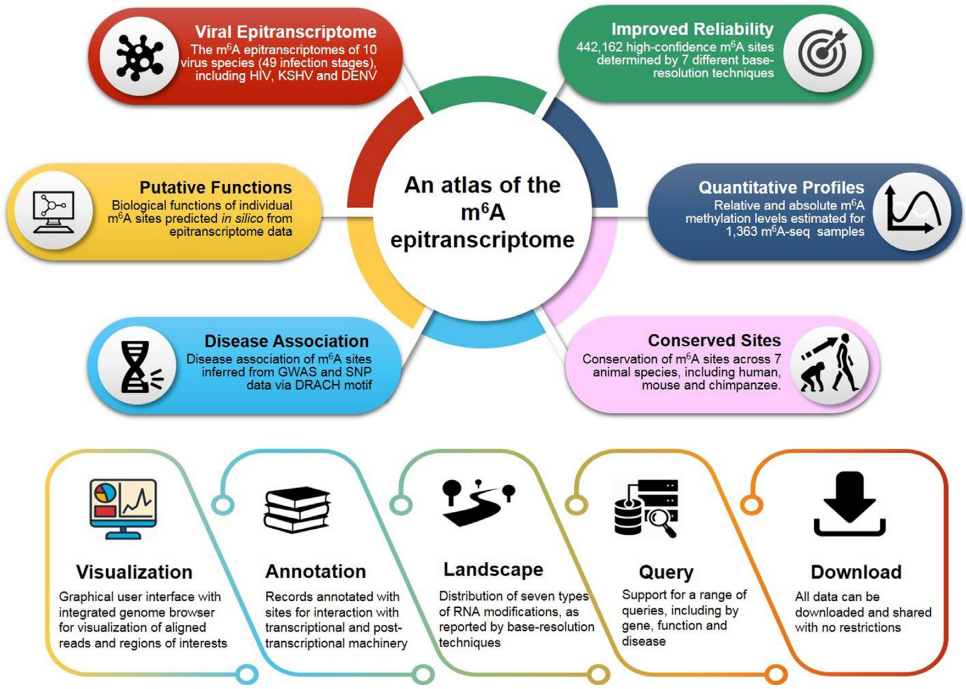

Figure 1. The overall design of m6A-Atlas. m^{6}A-Atlas features a high-confidence collection of reliable m^{6}A sites and quantitative epitranscriptome profiles. It also provides the conservation, sites for post-transcriptional machinery, putative biological functions and disease-association of individual m^{6}A sites, and the epitranscriptomes of host and virus during infection.

Quantification of m^{6}A methylation levels

A total of 1363 m^{6}A-seq samples profiling the epitranscriptomes in seven species were collected (Supplementary Tables S2–S4), and used to quantify the methylation levels of the high-confidence m^{6}A sites detected from base-resolution techniques. The raw m^{6}A-seq datasets were directly downloaded from NCBI GEO (37) and Genome Sequence Archive in BIG Data Center (38). Adaptors and low quality nucleotides were removed by Trim Galore, and the processed reads were aligned to the reference genome by HISAT2 (39). The m^{6}A level is calculated by the fold enrichment of reads in the IP samples compared with the input control samples within the 200 bp window of a given m^6

A site. It has been the most widely adopted way to quantify the methylation status, and is closely related to the peak calling process that usually seeks genomic regions with fold enrichment of >1 at a pre-defined significance level. Besides m^{6}A methylation levels, the matched transcriptional expression profiles (gene RNA expression levels in FPKM) were also estimated from the input control samples of m^{6}A-seq data with Cufflinks (40) to facilitate comparative analysis.

Viral and host epitranscriptomes

Recent studies suggested critical roles of m^{6}A on viral transcripts such as escaping from recognition by the RNA sensor RIG-I (41); unfortunately, this information has not been available in existing epitranscriptome databases. A total of 278 m^{6}A-seq samples profiling the epitranscriptomes during virus infection were collected (Supplementary Table S4). The data covered different infection stages of 10 virus species. The m^{6}A sites located on viral transcripts were detected by exomePeak2 (16) from m^{6}A-seq data, and are thus not at base-resolution as previously explained.

Conservation of m^{6}A sites in vertebrate

Conservation analysis is a powerful way for identifying the functionally important m^{6}A sites (42,43). By integrating and comparing a large number of m^{6}A sites across different species, we provided the first comprehensive profile of the conserved epitranscriptome in vertebrates at single-base resolution. During the cross-species comparison, the m^{6}A sites of one species were mapped to the homologous coordinates of a target organism using the LiftOver tool from the UCSC genome browser (44). Where the converted sites are also m^{6}A sites or located within the peaks enriched with m^{6}A signal in the target organism (for pig, monkey and chimpanzee, where only m^{6}A-seq data is available), these sites were considered conserved between the two species. Besides, the phastCons score (45) of the queried locus was also returned, showing the conservation of the corresponding genome region of the RNA methylation site.

Putative functions of individual m^{6}A sites

The putative biological functions of individual m^{6}A sites were predicted for the conserved sites between human and mouse according to the guilt-by-association principle, as previously described (46). This method is based on the association of methylation patterns (or co-methylation) among functionally related m^{6}A sites inferred from the collected quantitative epitranscriptome data (detailed in Supplementary Materials: Functional Prediction of Individual m6A

Sites and Supplementary Figure S2). The predicted biological functions of the m^{6}A sites may help generate hypotheses for subsequent experimental validation. Please note that GO prediction was made from the epitranscriptome profiles of individual m^{6}A site (rather than entire gene), and thus can only predict site-specific GO functions.

Potential involvement of individual m^{6}A sites in pathogenesis

Inspired by m6AVar (30), the potential involvement of individual m^{6}A sites in pathogenesis was inferred from disease-associated genetic mutations that directly destroy the m^{6}A forming motif DRACH. For such mutations pathogenesis may result from the impact on epitranscriptome regulation of elimination of an m^{6}A site. The germline and the somatic mutations required for this analysis were downloaded from dbSNP (47) and The Cancer Genome Atlas (TCGA) (48), respectively; while the disease-association of mutations were obtained from the GWAS catalog (49) and ClinVar (50).

Basic annotation for m^{6}A sites

The splicing sites, miRNA target sites, RBP-binding sites and subcellular location were integrated to help understand the regulatory roles of m^{6}A. The splicing sites and subcellular location were obtained from the UCSC database (51) and RNALocate, respectively. The miRNA target sites and RBP-binding sites in human were obtained from starBase (52) and POSTAR (53). Transcriptional regulation data (ChIP-seq and DNase-seq) were obtained from the GTRD database (54) and TRlnc (55), respectively.

Landscape of m^{6}A and other RNA modifications

Existing databases only focus on individual sites. To help understand the clustering effects (56–58) and the synergy (59,60) of m^{6}A and other RNA modifications, we provide the landscape of multiple RNA modifications on the same gene so as to provide a more comprehensive overview. To do so, a total of 96 datasets generated from base-resolution technologies were collected (Supplementary Table S5), covering six other RNA modifications in human, including 5-Methylcytosine (m^{5}C), Pseudouridine (Ψ), $N^{6,2}$-O-dimethyladenosine (m^{6}Am), N^1-methyladenine (m^{1}A), N^7-methylguanosine (m^{7}G) and 2′-O-methylation (Nm). For mouse, we integrated m^{5}C, m^{1}A and Ψ with m^{6}A modification; and we integrated m^{1}A and Ψ for yeast epitranscriptomes.

Database and web interface implementation

MySQL tables were exploited for the storage and management of the metadata in m^{6}A-Atlas. Hyper Text Markup Language (HTML), Cascading Style Sheets (CSS) and Hypertext Preprocessor (PHP) were used to construct the web interface. The multiple statistical diagrams were presented by EChars, and the Jbrowse genome browser (61) was implemented for interactive exploration and visualization of relevant records.

RESULT

A high-confidence collection of 442,162 unique m^{6}A sites identified from seven base-resolution technologies were collected, covering seven organisms, including human (178,049 sites), mouse (110 959 sites), rat (6348 sites), zebrafish (63 998 sites), fly (36 913 sites), Arabidopsis (35 332 sites) and yeast (10 563). Please see Supplementary Table S6 for a pair-wise comparison of data generated from seven base-resolution technologies. For the epitranscriptome landscape view, we integrated 96 profiles generated from base-resolution technologies only, covering six other RNA modifications, including m^{5}C (95 391 sites), m^{1}A (16 346 sites), Ψ (3137 sites), m^{6}Am (2447 sites), m^{7}G (2525 sites) and Nm (1835 sites). This is, to date, the most comprehensive collection of high-confidence transcriptome modification sites generated from base-resolution technologies. We also collected the m^{6}A sites on viral transcripts from 10 virus species during 8 infection stages, including, hMPV, KSHV, HCV, RSV, HIV-1, DENV, WNV, ZIKV, HCMV and YFV. Analysis revealed a total of 5388 genetic mutations, linked to pathogenesis, which can destroy the m^{6}A forming motif DRACH of these reliable m^{6}A sites, potentially indirectly linking m^{6}A function to 134 different types of disease.

Quantitative m^{6}A profiles (methylation levels of every single m^{6}A sites) were provided for 109 experimental conditions (different cell line/tissue or treatments) in human, covering 46 human cell lines, 28 adult tissues and 8 fetal tissues. For mouse, there are a total of 80 epitranscriptome profiles covering 25 cell lines and 10 mouse tissues. The matched gene expression profiles are also available for comparative analysis, as it is often beneficial to focus on the methylation sites located on readily expressed genes.

Conservation analysis was performed among human, mouse, rat, zebrafish, pig, monkey and chimpanzee to identify the conserved m^{6}A sites between two vertebrate species (see Table 1). Around 22 000 m^{6}A sites (located on 6193 human genes) were found to be conserved between human and mouse, representing 12.45% and 20.15% of the m^{6}A sites in human and mouse, respectively, and are strongly enriched near the stop codons on mRNAs (Supplementary Figure S1). Putative GO (62) functions were predicted as previously described (46) for the conserved m^{6}A sites, which are presumably functionally more important (detailed in Supplementary Materials).

An atlas of m^{6}A epitranscriptome (m^{6}A-Atlas)

We constructed m^{6}A-Atlas, a comprehensive knowledgebase for unraveling the N^6-methyladenosine (m^{6}A) epitranscriptome. The user-friendly web interfaces provided in m^{6}A-Atlas enable users to search, browse, visualize and download the comprehensively annotated m^{6}A data (see Table 2) by gene, disease, function, genomic location and specie. A genome browser was integrated for an interactive exploration of genome regions of interest. All data provided in the m^{6}A-Atlas database can be freely downloaded or shared. We also integrated two powerful tools: the m6A-conservation-finder, which is used to identify the conserved m^{6}A sites among a user-provided list of candidate sites, and the differential-m6A-finder, which returns

Table 1. Conserved m^{6}A sites between vertebrates

	Human (178 049 sites)	**Mouse** (110 959 sites)	**Rat** (6348 sites)	**Zebrafish** (63 998 sites)	**Chimpanzee** (37 456 peaks)	**Monkey** (38 838 peaks)	**Pig** (121 409 peaks)
Human		22 359 (20.15%)	1117 (17.60%)	2823 (4.411%)	20 757 (55.41%)	16 832 (43.33%)	22 245 (18.32%)
Mouse	22170 (12.45%)		1709 (26.92%)	1725 (2.695%)	8749 (23.35%)	7864 (20.24%)	15 636 (12.87%)
Rat	1132 (0.635%)	1729 (1.558%)		76 (0.119%)	245 (0.654%)	256 (0.659%)	N/A
Zebrafish	2922 (1.641%)	2063 (1.859%)	86 (1.354%)		N/A	N/A	N/A

Note: Base-resolution epitranscriptome datasets are available only for human, mouse, rat and zebrafish. For pig, monkey and chimpanzee, only m^{6}A peaks called from m^{6}A-seq data are available. N/A means that the UCSC LiftOver tool does not support the conversion of the homologous coordinates between the two species, and conservation analysis was not performed.

Table 2. Contents of m^{6}A-Atlas

	Reliable m^{6}A sites	Quantitative profiles	Basic annotation	Conservation in vertebrates	Putative functions	Landscape with six other RNA modifications	Disease association
Human	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mouse	Yes	Yes	Yes	Yes	Yes	Yes	-
Rat	Yes	Yes	Yes	Yes	-	-	-
Zebrafish	Yes	Yes	Yes	Yes	-	-	-
Fly	Yes	Yes	Yes	-	-	-	-
Arabidopsis	Yes	Yes	Yes	-	-	-	-
Yeast	Yes	Yes	Yes	-	-	Yes	-

the differentially methylated m^{6}A sites between two experimental conditions. For the convenience of the users, detailed instructions on how to use m^{6}A-Atlas were placed in the 'help' page. m^{6}A-Atlas is freely accessible online at: www.xjtlu.edu.cn/biologicalsciences/atlas.

Case study on lncRNA: MALAT1

Of interests here are the m^{6}A RNA methylation sites on MALAT1, which encodes a 7.9 kb long non-coding RNA and is located at chromosome 11q13. Searching by gene 'MALAT1' at the front page of m6A-Atlas database returns a total of 142 m^{6}A records related to MALAT1. This is consistent with our knowledge that MALAT1 is highly m6A-modified (63–66). It is possible to filter these records by technology (Figure 2A) or cell line/tissue (Figure 2B). After clicking an MALAT1 m^{6}A site, more information is provided including the source of the record such as PubMed ID, GEO accession number, technology, cell line and treatment. The overall distribution of all the m^{6}A sites located on MALAT1 together with their annotations is visualized in the 'landscape' section (Figure 2C), and can be downloaded when switching to its 'Table' view. It is clear from there that four MALAT1 m^{6}A sites (located on chr11: 65267646, 65267843, 65267952 and 65270276, respectively, hg19 assembly) were repetitively detected by more than 10 base-resolution experiments, and all the four of them are conserved between human and monkey, suggesting their prevalence and potentially important functions. Specifically, the MALAT1 m^{6}A site located at chr11:65257646 (supported by 13 experiments) were conserved between monkey (supported by one m^{6}A-seq experiment) and mouse (supported by six base-resolution experiments) (Figure 2D). The biological functions of this site was predicted *in silico* and shown in the GO prediction panel. We can see that the m^{6}A site exhibits co-methylation pattern with the m^{6}A sites located on USP10, AGO2, TRAF7, etc. (Figure 2E), and network-based analysis predicts, according to the guilt-by-association principle, that it may be associated with a number of functions, such as, DNA synthesis involved in DNA repair (P-value = 0.01297), cell fate determination (p-value = 0.0454) and negative regulation of cell development (P-value = 0.00366) (Figure 2F). This is consistent with recent finding that m^{6}A regulation of MALAT1 regulates tumorigenesis (67) and embryonic development (68). The methylation profiles of individual m^{6}A site under >100 different biological contexts (cell line, tissue, treatment, virus infection, etc.) can be found under the 'methylation level' panel. It is possible to filter for a specific cell line or tissue of interests. The highest methylation level was found in iSLK.219 cell line. It is worth noting that the methylation level of this site is higher than 0 on almost all conditions covered, suggesting its prevalent existence and functionality. (Figure 2G). The matched RNA expression profiles of MALAT1 are also available in the gene expression panel for comparative analysis (Figure 2H). Additionally, post-transcriptional annotation suggests that the m^{6}A site is in proximity to the binding position of microRNA hsa-miR-670-3p and 21 RNA binding proteins, including HRNPC, which may be immediately regulated through the m^{6}A site. Interestingly, the regulation of interaction between MATLAT1 and HNRNPC via m^{6}A methylation has been previously confirmed (69).

Case study on protein coding gene: SMAD3

SMAD3 (SMAD Family Member 3) has been shown previously shown to promote binding of the m^{6}A methyltransferase complex to a subset of transcripts involved in early cell fate decisions (70). Of interests here is to check the m^{6}A methylation status of this gene. Searching by gene 'SMAD3' at the front page of m^{6}A-Atlas database returns a total of 44 m^{6}A records related to MALAT1. It is possible to further filter these records by technology (Figure 3A) or cell

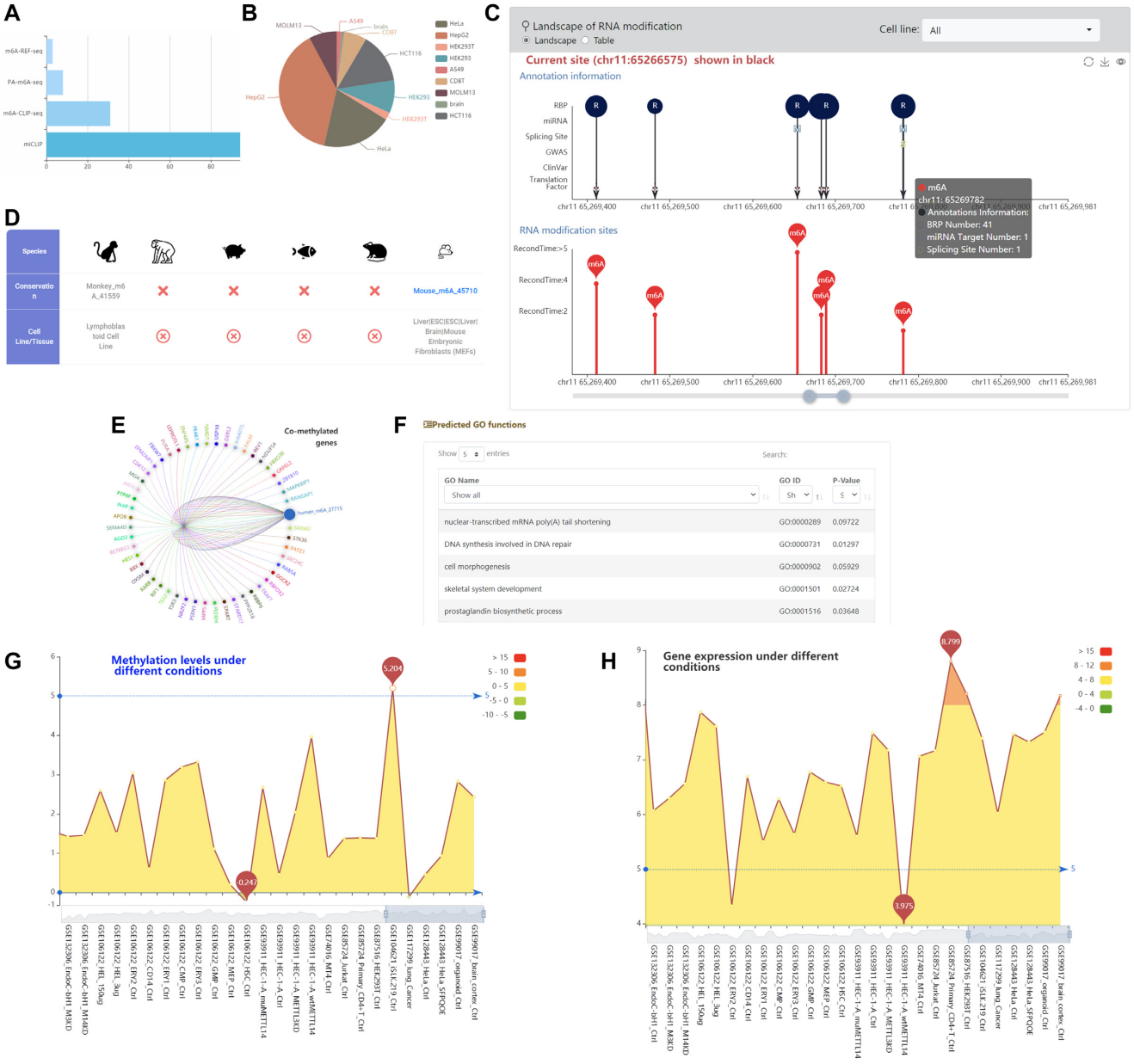

Figure 2. The m^{6}A sites of MALAT1 and related information. (**A** and **B**) More than 100 m^{6}A records of MALAT1 were detected from different technologies and in different cell lines/tissues. (**C**) The overall distribution pattern of m^{6}A sites on MALAT1 transcript is illustrated together with their annotations. (**D**) An m^{6}A site is conserved between human and mouse. (**E**) The m^{6}A site is co-methylated with sites located on a few genes, which may suggest their functional relevance. (**F**) The predicted GO functions of an individual m^{6}A site. (**G**) The methylation levels of an m^{6}A site under different biological conditions. (**H**) The matched gene expression profiles of the m^{6}A-carrying gene.

line/tissue (Figure 3B). After clicking an MALAT1 m^{6}A site, more information is provided including the source of the record such as PubMed ID, GEO accession number, technology, cell line and treatment. The overall distribution of all the m^{6}A (and other RNA modification) sites located on MALAT1 together with their annotations can be visualized in the 'landscape' section (Figure 3C), and can be downloaded when switching to its 'Table' view. It is clear from there that four MALAT1 RNA modification sites (located on chr15: 67483526, 67358454, 67358455 and 67358461, respectively, hg19 assembly) were repetitively detected by more than 5 base-resolution experiments. However, only the first one (chr15: 67483526) is m^{6}A RNA methylation site; the others three are all m^{5}C sites. The m^{6}A site located at chr15: 67483526 (supported by seven experiments) is conserved in monkey and chimp (both supported by one m^{6}A-seq experiment) (Figure 3D). We can see from the 'methylation level' panel that, the methylation level (log2 fold enrichment) of this site is above 0 for almost all experimental conditions (Figure 3E), showing its prevalent existence and functionality. The highest methylation level was observed under A549 cell lines, and a strong demethy-

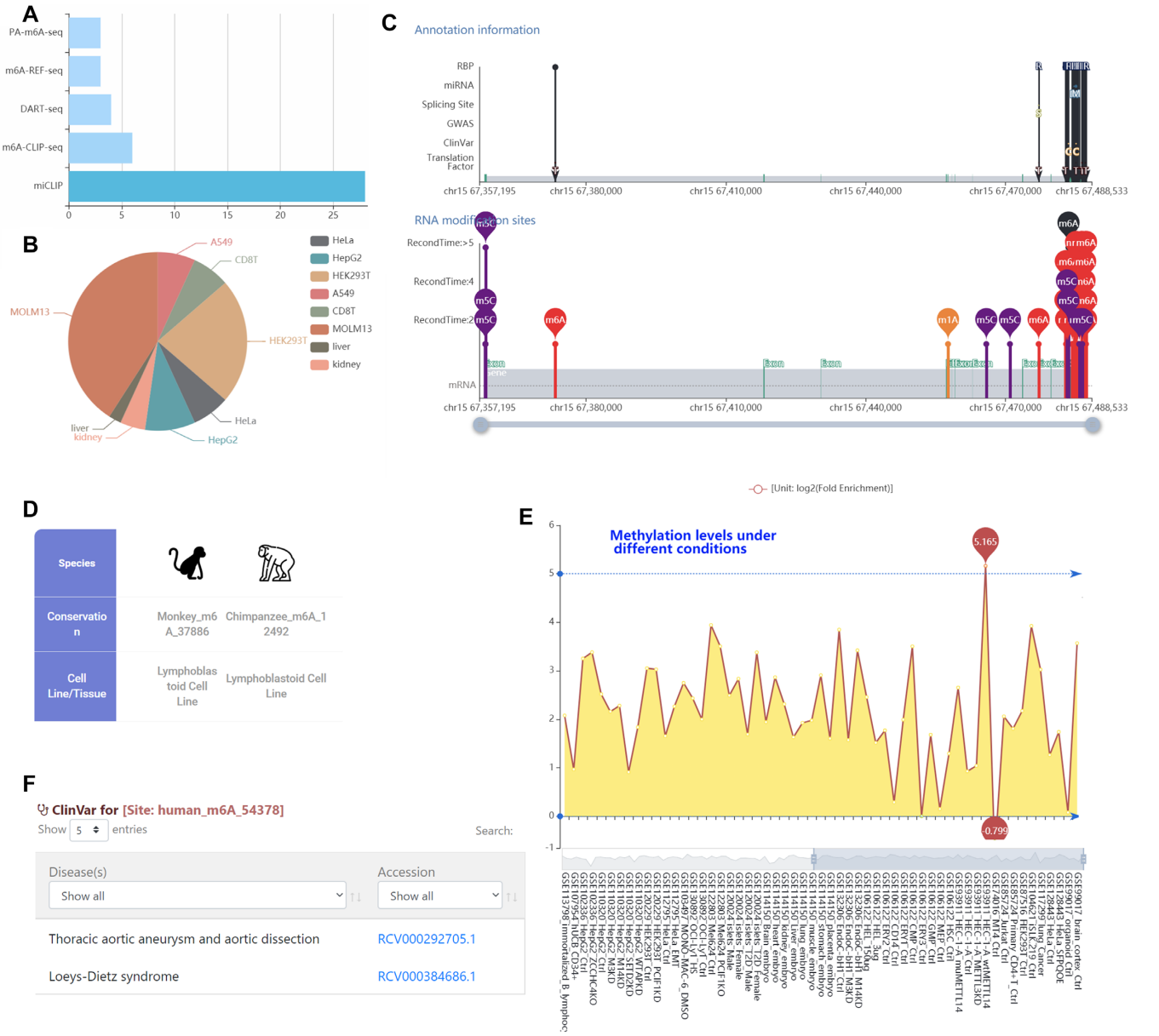

Figure 3. The m^{6}A sites of SMAD3 and related information. (**A** and **B**) More than 40 m^{6}A records of SMAD3 were detected from different technologies and in different cell lines/tissues. (**C**) The overall distribution pattern of m^{6}A sites on MALAT1 transcript is illustrated together with their annotations. (**D**) An m^{6}A site is conserved between human, mouse and Chimp. (**E**) The m^{6}A site shows stable m6A signal (with $\log_2$ fold enrichment greater than 0) under different experimental conditions, suggesting its universal existence and prevalence. The highest methylation level was detected under A549 cell lines, and a strong demethylation of the site was observed after METTL3 is knocked down, suggesting the methylation of this site is likely to be METTL3-dependent. (**F**) A known disease-relevant germline mutation (chr15: 67483526) can destroy the m^{6}A forming motif of this m6A site, linking it to epitranscriptome disease pathogenesis.

lation of the site was observed after METTL3 is knocked down, suggesting the methylation of this site is likely to be METTL3-dependent. Importantly, the m^{6}A-forming motif of this site DRACH may be destroyed by a known synonymous mutation (chr15: 67483526) relevant to Thoracic aortic aneurysm and aortic dissection (RCV000292705) and Loeys-Dietz syndrome (RCV000384686), which may indicate a disease-relevant mechanism that functions at the epitranscriptome layer (Figure 3F).

Case study: KSHV viral m^{6}A

The 'Virus' page of m^{6}A-Atlas provides the m^{6}A sites located on the viral transcripts. Of interests here is the m^{6}A methylation sites located on KSHV viral transcripts. After selecting KSHV at the 'Virus' page, a total 356 m^{6}A sites were returned. Various information related to those sites were provided, including the position, source, strain, gene, width, etc., and it is possible to further narrow down the results by the infected cell line (Figure 4).

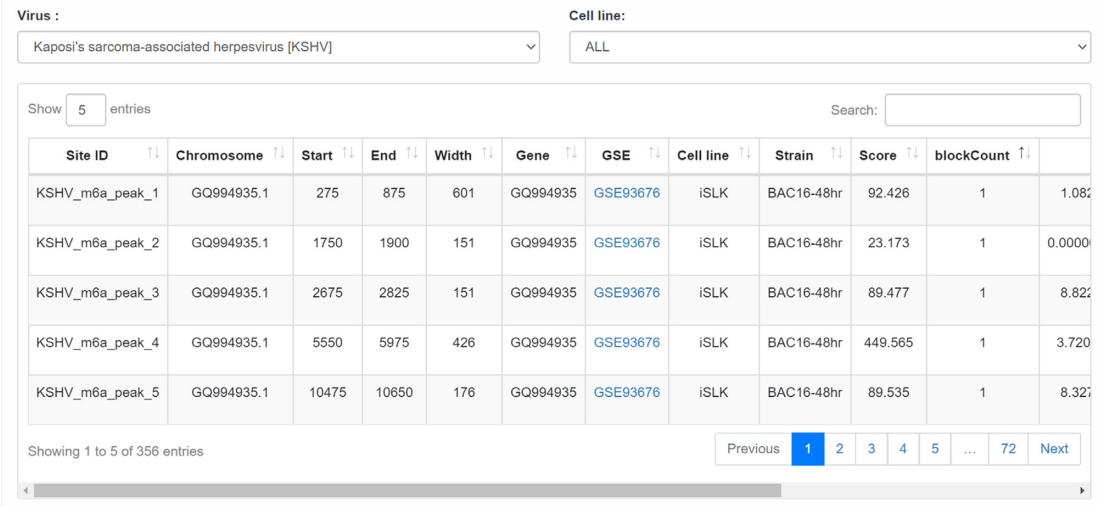

Figure 4. Viral m^{6}A sites on KSHV transcripts. The 356 m^{6}A sites located on KSHV viral transcripts detected under various conditions and infection stages.

CONCLUSION

As the most abundant and most intensively studied RNA modification, *N*^{6}-methyladenosine (m^{6}A) has been regarded as an important epigenetic mark regulating various stages of various aspects of RNA function. Centralized bioinformatics platforms are sorely needed to take advantage of the rapid accumulation of large amounts of epitranscriptome data. We present here m^{6}A-Atlas, a comprehensive knowledgebase for the m^{6}A epitranscriptome. Compared to existing epitranscriptome databases, m^{6}A-Atlas features a high-confidence collection of reliable m^{6}A sites and condition-specific quantitative (rather than binary) epitranscriptome profiles. It also offers novel features, such as, the conservation of m^{6}A sites among seven vertebrate species (including human, mouse and chimp), the viral epitranscriptomes of 10 virus species (including HIV, KSHV and DENV), the putative biological functions of individual m^{6}A sites predicted from epitranscriptome data, and the potential involvement in disease of individual m^{6}A sites inferred from disease-associated genetic mutations that can destroy m^{6}A specifying sequence motif. Please refer to Supplementary Table S7 for a brief comparison of m^{6}A-Atlas with existing epitranscriptome databases focusing on RNA modification site collection. A user-friendly graphical user interface was constructed to support the query, visualization and sharing of the m^{6}A epitranscriptomes annotated with potential interaction sites with the transcriptional and post-transcriptional machineries (TF binding, RBP-binding, microRNA interaction and splicing sites). The m^{6}A data are also presented in the context of sites of six other RNA modifications, enabling an overview of the RNA modification landscape. These resources together provide fresh opportunities for studying the complexities of m^{6}A epitranscriptomes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Author contributions: K. Chen conceived the idea and initialized the project; K. Chen and J. Ma collected and processed the epitranscriptome data of eukaryotes and viruses, respectively; B. Song generated post-transcriptional annotations, conducted disease-association analysis and conservation analysis; X. Wu performed GO function prediction; Y. Tang designed and built the m^{6}A-Atlas website; K. Chen drafted the manuscript. All authors read, critically revised and approved the final manuscript.

FUNDING

National Natural Science Foundation of China [31671373]; XJTLU Key Program Special Fund [KSF-T-01, KSF-E-51, KSF-P-02]. Funding for open access charge: National Natural Science Foundation of China [31671373]; XJTLU Key Program Special Fund [KSF-T-01, KSF-E-51, KSF-P-02]
Conflict of interest statement. Daniel J. Rigden is Executive Editor of *NAR*.

REFERENCES

1. Jones,J.D., Monroe,J. and Koutmou,K.S. (2020) A molecular-level perspective on the frequency, distribution, and consequences of messenger RNA modifications. *WIREs RNA*, **n/a**, e1586.
2. Boccaletto,P., Machnicka,M.A., Purta,E., Piatkowski,P., Baginski,B., Wirecki,T.K., de Crecy-Lagard,V., Ross,R., Limbach,P.A., Kotter,A. *et al.* (2018) MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.*, **46**, D303–D307.
3. Meyer,K.D. and Jaffrey,S.R. (2017) Rethinking m(6)A readers, writers, and erasers. *Annu. Rev. Cell Dev. Biol.*, **33**, 319–342.
4. Winkler,R., Gillis,E., Lasman,L., Safra,M., Geula,S., Soyris,C., Nachshon,A., Tai-Schmiedel,J., Friedman,N., Le-Trilling,V.T.K. *et al.* (2019) m(6)A modification controls the innate immune response to infection by targeting type I interferons. *Nat. Immunol.*, **20**, 173–182.
5. Engel,M., Eggert,C., Kaplick,P.M., Eder,M., Roh,S., Tietze,L., Namendorf,C., Arloth,J., Weber,P., Rex-Haffner,M. *et al.* (2018) The Role of m(6)A/m-RNA Methylation in Stress Response Regulation. *Neuron*, **99**, 389–403.
6. Zhou,J., Wan,J., Gao,X., Zhang,X., Jaffrey,S.R. and Qian,S.B. (2015) Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature*, **526**, 591–594.

7. Xiang,Y., Laurent,B., Hsu,C.H., Nachtergaele,S., Lu,Z., Sheng,W., Xu,C., Chen,H., Ouyang,J., Wang,S. *et al.* (2017) RNA m(6)A methylation regulates the ultraviolet-induced DNA damage response. *Nature*, **543**, 573–576.
8. Liu,N., Dai,Q., Zheng,G., He,C., Parisien,M. and Pan,T. (2015) N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature*, **518**, 560–564.
9. Wang,X., Lu,Z., Gomez,A., Hon,G.C., Yue,Y., Han,D., Fu,Y., Parisien,M., Dai,Q., Jia,G. *et al.* (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*, **505**, 117–120.
10. Wang,X., Zhao,B.S., Roundtree,I.A., Lu,Z., Han,D., Ma,H., Weng,X., Chen,K., Shi,H. and He,C. (2015) N(6)-methyladenosine modulates messenger RNA translation efficiency. *Cell*, **161**, 1388–1399.
11. Esteve-Puig,R., Bueno-Costa,A. and Esteller,M. (2020) Writers, readers and erasers of RNA modifications in cancer. *Cancer Lett.*, **474**, 127–137.
12. Dominissini,D., Moshitch-Moshkovitz,S., Schwartz,S., Salmon-Divon,M., Ungar,L., Osenberg,S., Cesarkas,K., Jacob-Hirsch,J., Amariglio,N., Kupiec,M. *et al.* (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*, **485**, 201–206.
13. Meyer,K.D., Saletore,Y., Zumbo,P., Elemento,O., Mason,C.E. and Jaffrey,S.R. (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell*, **149**, 1635–1646.
14. Feng,J., Liu,T., Qin,B., Zhang,Y. and Liu,X.S. (2012) Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.*, **7**, 1728–1740.
15. Zhang,Y. and Hamada,M. (2020) MoAIMS: efficient software for detection of enriched regions of MeRIP-Seq. *BMC Bioinformatics*, **21**, 103.
16. Meng,J., Cui,X., Rao,M.K., Chen,Y. and Huang,Y. (2013) Exome-based analysis for RNA epigenome sequencing data. *Bioinformatics*, **29**, 1565–1567.
17. Liu,L., Zhang,S.-W., Huang,Y. and Meng,J. (2017) QNB: differential RNA methylation analysis for count-based small-sample sequencing data with a quad-negative binomial model. *BMC Bioinformatics*, **18**, 387.
18. Li,J., Huang,Y., Cui,Q. and Zhou,Y. (2020) m6Acorr: an online tool for the correction and comparison of m6A methylation profiles. *BMC Bioinformatics*, **21**, 31.
19. Meng,J., Lu,Z., Liu,H., Zhang,L., Zhang,S., Chen,Y., Rao,M.K. and Huang,Y. (2014) A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. *Methods*, **69**, 274–281.
20. Zhang,T., Zhang,S.W., Zhang,L. and Meng,J. (2018) trumpet: transcriptome-guided quality assessment of m(6)A-seq data. *BMC Bioinformatics*, **19**, 260.
21. McIntyre,A.B.R., Gokhale,N.S., Cerchietti,L., Jaffrey,S.R., Horner,S.M. and Mason,C.E. (2020) Limits in the detection of m(6)A changes using MeRIP/m(6)A-seq. *Scientific reports*, **10**, 6590.
22. Chen,K., Lu,Z., Wang,X., Fu,Y., Luo,G.Z., Liu,N., Han,D., Dominissini,D., Dai,Q., Pan,T. *et al.* (2015) High-resolution N(6) -methyladenosine (m(6) A) map using photo-crosslinking-assisted m(6) A sequencing. *Angew. Chem. Int. Ed. Engl.*, **54**, 1587–1590.
23. Linder,B., Grozhik,A.V., Olarerin-George,A.O., Meydan,C., Mason,C.E. and Jaffrey,S.R. (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods*, **12**, 767–772.
24. Ke,S., Alemu,E.A., Mertens,C., Gantman,E.C., Fak,J.J., Mele,A., Haripal,B., Zucker-Scharff,I., Moore,M.J., Park,C.Y. *et al.* (2015) A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation. *Genes Dev.*, **29**, 2037–2053.
25. Liu,H., Wang,H., Wei,Z., Zhang,S., Hua,G., Zhang,S.W., Zhang,L., Gao,S.J., Meng,J., Chen,X. *et al.* (2018) MeT-DB V2.0: elucidating context-specific functions of N6-methyl-adenosine methyltranscriptome. *Nucleic Acids Res.*, **46**, D281–D287.
26. Xuan,J.J., Sun,W.J., Lin,P.H., Zhou,K.R., Liu,S., Zheng,L.L., Qu,L.H. and Yang,J.H. (2018) RMBase v2.0: deciphering the map of RNA modifications from epitranscriptome sequencing data. *Nucleic Acids Res.*, **46**, D327–D334.
27. Han,Y., Feng,J., Xia,L., Dong,X., Zhang,X., Zhang,S., Miao,Y., Xu,Q., Xiao,S., Zuo,Z. *et al.* (2019) CVm6A: a visualization and exploration database for m(6)As in cell lines. *Cells*, **8**, 168.
28. Liu,S., Zhu,A., He,C. and Chen,M. (2020) REPIC: a database for exploring N6-methyladenosine methylome. *Genome Biol.*, **21**, 100.
29. Liu,Q. and Gregory,R.I. (2019) RNAmod: an integrated system for the annotation of mRNA modifications. *Nucleic Acids Res.*, **47**, W555–W548.
30. Zheng,Y., Nie,P., Peng,D., He,Z., Liu,M., Xie,Y., Miao,Y., Zuo,Z. and Ren,J. (2018) m6AVar: a database of functional variants involved in m6A modification. *Nucleic Acids Res.*, **46**, D139–D145.
31. Chen,T., Hao,Y., Zhang,Y., Li,M., Wang,M., Han,W., Wu,Y., Lv,Y., Hao,J. and Wang,L. (2015) m6A RNA methylation is regulated by MicroRNAs and promotes reprogramming to pluripotency. *Cell Stem Cell*, **16**, 289–301.
32. Chen,K., Wu,X., Zhang,Q., Wei,Z., Rong,R., Lu,Z., Meng,J., de Magalhães,J.P., Su,J. and Rigden,D.J. (2019) WHISTLE: a high-accuracy map of the human N6-methyladenosine (m6A) epitranscriptome predicted using a machine learning approach. *Nucleic Acids Res.*, **47**, e41.
33. Zhou,Y., Zeng,P., Li,Y.-H., Zhang,Z. and Cui,Q. (2016) SRAMP: prediction of mammalian N6-methyladenosine (m6A) sites based on sequence-derived features. *Nucleic Acids Res.*, **44**, e91.
34. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M. and Li,W. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
35. Zhang,Z., Zhan,Q., Eckert,M., Zhu,A., Chryplewicz,A., De Jesus,D.F., Ren,D., Kulkarni,R.N., Lengyel,E., He,C. *et al.* (2019) RADAR: differential analysis of MeRIP-seq data with a random effect model. *Genome Biol.*, **20**, 294.
36. Cui,X., Zhang,L., Meng,J., Rao,M., Chen,Y. and Huang,Y. (2018) MeTDiff: a novel differential RNA methylation analysis for MeRIP-Seq data. *IEEE/ACM Trans. Comput. Biol. Bioinf.*, **15**, 526–534.
37. Wheeler,D.L., Church,D.M., Edgar,R., Federhen,S., Helmberg,W., Madden,T.L., Pontius,J., Schuler,G.D., Schriml,L.M. and Sequeira,E. (2004) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **32**, 13–16.
38. Tang,B. (2017) GSA: Genome Sequence Archive*. *Genomics Proteomics Bioinformatics*, **15**, 14–18.
39. Kim,D., Langmead,B. and Salzberg,S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods*, **12**, 357–360.
40. Roberts,A., Trapnell,C., Donaghey,J., Rinn,J.L. and Pachter,L. (2011) Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol.*, **12**, R22.
41. Lu,M., Zhang,Z., Xue,M., Zhao,B.S., Harder,O., Li,A., Liang,X., Gao,T.Z., Xu,Y., Zhou,J. *et al.* (2020) N6-methyladenosine modification enables viral RNA to escape recognition by RNA sensor RIG-I. *Nat. Microbiol.*, **5**, 584–598.
42. Ma,L., Zhao,B., Chen,K., Thomas,A., Tuteja,J.H., He,X., He,C. and White,K.P. (2017) Evolution of transcript modification by N(6)-methyladenosine in primates. *Genome Res.*, **27**, 385–392.
43. Liu,Z. and Zhang,J. (2018) Most m6A RNA modifications in protein-coding regions are evolutionarily unconserved and likely nonfunctional. *Mol. Biol. Evol.*, **35**, 666–675.
44. Speir,M.L., Zweig,A.S., Rosenbloom,K.R., Raney,B.J., Benedict,P., Parisa,N., Lee,B.T., Katrina,L., Donna,K. and Hinrichs,A.S. (2015) The UCSC Genome Browser database: 2016 update. *Nucleic Acids Res.*, **44**, D717–D725.
45. Siepel,A., Bejerano,G., Pedersen,J.S., Hinrichs,A.S., Hou,M., Rosenbloom,K., Clawson,H., Spieth,J., Hillier,L.W., Richards,S. *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.*, **15**, 1034–1050.
46. Wu,X., Wei,Z., Chen,K., Zhang,Q., Su,J., Liu,H., Zhang,L. and Meng,J. (2019) m6Acomet: large-scale functional prediction of individual m(6)A RNA methylation sites from an RNA co-methylation network. *BMC Bioinformatics*, **20**, 223.
47. Sherry,S.T., Ward,M.H., Kholodov,M., Baker,J., Phan,L., Smigielski,E.M. and Sirotkin,K. (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.*, **29**, 308–311.
48. Tomczak,K., Czerwinska,P. and Wiznerowicz,M. (2015) The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp. Oncol. (Poznan, Poland)*, **19**, A68–A77.

49. Buniello,A., MacArthur,J.A.L., Cerezo,M., Harris,L.W., Hayhurst,J., Malangone,C., McMahon,A., Morales,J., Mountjoy,E., Sollis,E. *et al.* (2019) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.*, **47**, D1005–D1012.
50. Landrum,M.J., Lee,J.M., Benson,M., Brown,G., Chao,C., Chitipiralla,S., Gu,B., Hart,J., Hoffman,D., Hoover,J. *et al.* (2016) ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.*, **44**, D862–D868.
51. Lawrence,M., Huber,W., Pages,H., Aboyoun,P., Carlson,M., Gentleman,R., Morgan,M.T. and Carey,V.J. (2013) Software for computing and annotating genomic ranges. *PLoS Comput. Biol.*, **9**, e1003118.
52. Li,J.H., Liu,S., Zhou,H., Qu,L.H. and Yang,J.H. (2014) starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.*, **42**, D92–D97.
53. Zhu,Y., Xu,G., Yang,Y.T., Xu,Z., Chen,X., Shi,B., Xie,D., Lu,Z.J. and Wang,P. (2019) POSTAR2: deciphering the post-transcriptional regulatory logics. *Nucleic Acids Res.*, **47**, D203–D211.
54. Yevshin,I., Sharipov,R., Valeev,T., Kel,A. and Kolpakov,F. (2016) GTRD: a database of transcription factor binding sites identified by ChIP-seq experiments. *Nucleic Acids Res.*, **45**, D61–D67.
55. Li,Y., Li,X., Yang,Y., Li,M., Qian,F., Tang,Z., Zhao,J., Zhang,J., Bai,X., Jiang,Y. *et al.* (2020) TRlnc: a comprehensive database for human transcriptional regulatory information of lncRNAs. *Brief. Bioinform.*, bbaa011.
56. Ke,S., Pandya-Jones,A., Saito,Y., Fak,J.J., Vagbo,C.B., Geula,S., Hanna,J.H., Black,D.L., Darnell,J.E. Jr. and Darnell,R.B. (2017) m(6)A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev.*, **31**, 990–1006.
57. Zou,Q., Xing,P., Wei,L. and Liu,B. (2019) Gene2vec: gene subsequence embedding for prediction of mammalian N (6)-methyladenosine sites from mRNA. *RNA*, **25**, 205–218.
58. Chen,K., Wei,Z., Zhang,Q., Wu,X., Rong,R., Lu,Z., Su,J., de Magalhaes,J.P., Rigden,D.J. and Meng,J. (2019) WHISTLE: a high-accuracy map of the human N6-methyladenosine (m6A) epitranscriptome predicted using a machine learning approach. *Nucleic Acids Res.*, **47**, e41.
59. Courtney,D.G., Tsai,K., Bogerd,H.P., Kennedy,E.M., Law,B.A., Emery,A., Swanstrom,R., Holley,C.L. and Cullen,B.R. (2019) Epitranscriptomic addition of m(5)C to HIV-1 transcripts regulates viral gene expression. *Cell Host Microbe*, **26**, 217–227.
60. Courtney,D.G., Chalem,A., Bogerd,H.P., Law,B.A., Kennedy,E.M., Holley,C.L. and Cullen,B.R. (2019) Extensive epitranscriptomic methylation of A and C residues on murine leukemia virus transcripts enhances viral gene expression. *MBio*, **10**, e01209-19.
61. Skinner,M.E., Uzilov,A.V., Stein,L.D., Mungall,C.J. and Holmes,I.H. (2009) JBrowse: a next-generation genome browser. *Genome Res.*, **19**, 1630–1638.
62. Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S. and Eppig,J.T. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
63. Li,L., Feng,T., Lian,Y., Zhang,G., Garen,A. and Song,X. (2009) Role of human noncoding RNAs in the control of tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 12956–12961.
64. Ji,Q., Zhang,L., Liu,X., Zhou,L., Wang,W., Han,Z., Sui,H., Tang,Y., Wang,Y. and Liu,N. (2014) Long non-coding RNA MALAT1 promotes tumour growth and metastasis in colorectal cancer through binding to SFPQ and releasing oncogene PTBP2 from SFPQ/PTBP2 complex. *Br. J. Cancer*, **111**, 736–748.
65. Latorre,E., Carelli,S., Raimondi,I., D'Agostino,V., Castiglioni,I., Zucal,C., Moro,G., Luciani,A., Ghilardi,G. and Monti,E. (2016) The ribonucleic complex HuR-MALAT1 represses CD133 expression and suppresses epithelial–mesenchymal transition in breast cancer. *Cancer Res.*, **76**, 2626–2636.
66. Zhang,X., Hamblin,M.H. and Yin,K.-J. (2017) The long noncoding RNA Malat1: its physiological and pathophysiological functions. *RNA biology*, **14**, 1705–1714.
67. Brown,J.A., Kinzig,C.G., DeGregorio,S.J. and Steitz,J.A. (2016) Methyltransferase-like protein 16 binds the 3′-terminal triple helix of MALAT1 long noncoding RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 14013–14018.
68. Mendel,M., Chen,K.-M., Homolka,D., Gos,P., Pandey,R.R., McCarthy,A.A. and Pillai,R.S. (2018) Methylation of structured RNA by the m6A writer METTL16 is essential for mouse embryonic development. *Mol. Cell*, **71**, 986–1000.
69. He,R.-Z., Jiang,J. and Luo,D.-X. (2020) The functions of N6-methyladenosine modification in lncRNAs. *Genes Dis*, doi:10.1016/j.gendis.2020.03.005.
70. Bertero,A., Brown,S., Madrigal,P., Osnato,A., Ortmann,D., Yiangou,L., Kadiwala,J., Hubner,N.C., de los Mozos,I.R., Sadée,C. *et al.* (2018) The SMAD2/3 interactome reveals that TGFβ controls m6A mRNA methylation in pluripotency. *Nature*, **555**, 256–259.