RNA modifications: an overview of select web-based tools

JILLIAN RAMOS

Department of Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, School of Medicine, Aurora, Colorado 80045, USA

Editor's Note:
Jillian Ramos was the recipient of the RNA Society/Cold Spring Harbor Laboratory Press Award for Research Excellence by an Underrepresented Scientist presented by the RNA Society.

ABSTRACT

The field of epitranscriptomics has expanded dramatically in recent years, both in the number of identified RNA modifications and the number of researchers studying them. As knowledge of post-transcriptional modifications continues to expand, numerous new methods have been developed to detect these modifications. Additionally, modifications are being extended to therapeutic settings, such as with recent mRNA vaccines. With this increase in knowledge and use, the community is recognizing the necessity for user-friendly databases to (i) store information from both high- and low-throughput studies and (ii) provide prediction software on how RNA modifications contribute to RNA function and disease. This mini-review highlights select RNA modification databases and their key attributes with the aim of providing a resource to researchers in the field of epitranscriptomics.

INTRODUCTION

In 1957, the first modified RNA nucleoside was identified through the isolation and digestion of ribonucleic acids from yeast *Saccharomyces cerevisiae* (*S. cerevisiae*) (Davis and Allen 1957). This "fifth Nucleotide" would later be termed pseudouridine (Ψ) (Cohn 1960). Since this discovery, hundreds of RNA modifications have been identified on multiple RNA types (Fig. 1) in all domains of life as well as viruses. As detection methods become more sophisticated, we continue to find more (Grosjean 2015).

The increase in identification of ribonucleoside modifications has been driven by advancements in detection methods. Some of the first modified ribonucleosides were identified by thin-layer chromatography (TLC) and high-performance liquid chromatography (HPLC) (Grosjean et al. 2004). Since then, other methods for RNA modification detection have been pioneered, including liquid chromatography–mass spectrometry (LC–MS) and RNA reverse transcription (Kowalak et al. 1993; Motorin et al. 2007; Cai et al. 2015). While each technique has advantages, such as accurate quantification of a given modification or the specific location of the known modified residue, each also has pitfalls (Schaefer et al. 2017). Identifying the location of specific modifications often remains difficult due to certain modifications blocking the reverse transcriptase (RT) that is essential for sequencing RNA. Methods to overcome this limitation include using RTs with higher processivity capable of reading through modifications, AlkB treatment to demethylate select modifications that inhibit RT followed by sequencing, and now direct nanopore sequencing of an RNA (Cozen et al. 2015; Xu et al. 2019; Leger et al. 2021; Motorin and Marchand 2021). Improvement in detection and mapping increases the rate of RNA modification identification, and whole genome sequencing allows for identification of disease-causing variants in RNA modification enzymes. Functionally perturbed RNA modification enzymes can lead to several diseases, such as cancer and neurodevelopmental disorders, highlighting the importance of specific modifications to RNA function and to human health (Jonkhout et al. 2017).

In addition to their role in biology, RNA modifications can be repurposed for functions such as improving RNA-based therapies. The addition of post-transcriptional modifications can reduce the immunogenicity of RNA therapies by preventing activation of Toll-like receptors, a component of the innate immune system. These modifications

Corresponding author: jillian.ramos@cuanschutz.edu
Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.079443.122. Freely available online through the *RNA* Open Access option.

© 2022 Ramos This article, published in *RNA*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

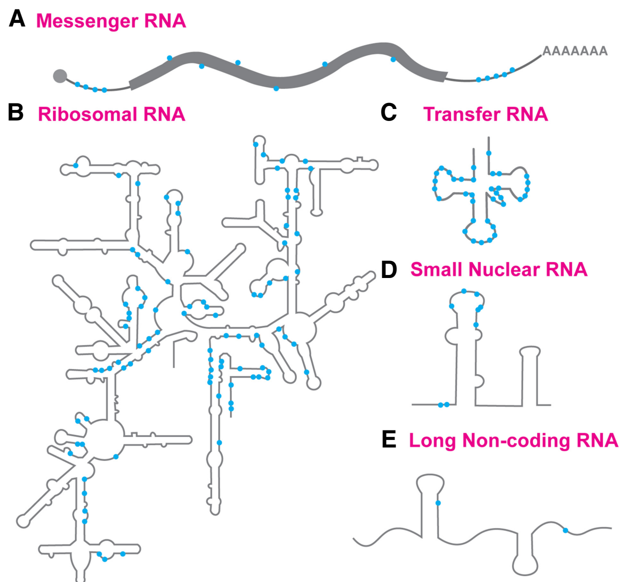

FIGURE 1. (*A*) A representative messenger RNA (gray) with conceptual locations of RNA modifications (blue dots) in the UTRs and coding region. (*B*) A representative ribosomal RNA (gray) from the small subunit. The modifications (blue dots) encapsulate all modifications identified on all rRNAs associated with the SSU for all organisms in MODOMICS. (*C*) A representative transfer RNA (gray). The modifications (blue dots) encapsulate all modifications identified on tRNAs for all organisms in MODOMICS. (*D*) The small nuclear RNA represented is the U5 snRNA, with the modifications (blue dots) highlighting locations of pseudouridines and ribose methylations. Figure adapted from Karijolich and Yu (2010). (*E*) A schematic of the long noncoding RNA HOTAIR with conceptual locations of modifications (blue dots) highlighting m^6A. Figure adapted from Porman et al. (2021).

include pseudouridine (Ψ), 2-thiouridine (2sU), 5-methylcytidine (m^5C), N6-methyladenosine (m^6A), and 5-methyluracil (m^5U) (Kariko et al. 2005). Additionally, the well-known N1-methyl-pseudouridine (m1Ψ) modification, which is currently used in COVID-19 mRNA vaccines, results in increased protein expression (Andries et al. 2015).

The advancement of detection methods, discoveries of more naturally occurring RNA modifications and modifying enzymes, and improvements in methods to predict modified nucleoside locations and their impacts has led to the creation of databases to compile the growing pool of information. The purpose of this mini-review is to highlight select databases, providing a source to help researchers navigate these resources. Described below are three "general modification" databases that address many types of post-transcriptional modifications and five additional databases containing information on four specific modifications (N6-methyladenosine [m^6A], 5-methylcytosine [m^5C], pseudouridine [Ψ], and inosine [I]) (Table 1). This review focuses on databases that link multiple pieces of information (e.g., modification status across multiple species, position of the modified residue, and links to publications), but does not discuss single-feature databases (e.g., modification site prediction tools). Finally, this review concludes with suggestions to increase the utility of these databases.

MODOMICS

MODOMICS was the first comprehensive RNA modification database, founded in 2006 by Stanislaw Dunin-Horkawicz and colleagues, and has been continuously updated, most recently in 2021 (Dunin-Horkawicz et al. 2006; Boccaletto et al. 2022). In 2006, the three primary features of the database included (i) providing modification chemical structures, (ii) displaying the predicted modification pathways, and (iii) linking the associated known RNA modification enzymes with their post-transcriptional modifications. In 2006 the RNA modification enzymes cataloged were solely from *Escherichia coli* (*E. coli*) and *S. cerevisiae,* while the 2021 update contains modification enzymes from 50+ species. From the original three menu options, the 2021 iteration contains nine menu options. Highlighted below are three useful features, but users should

TABLE 1. Table of all the databases discussed in this review.

Database	Modifications	# of species	Reference
MODOMICS	150+	50+	Boccaletto et al. (2022)
RMBase	m^6A, m^1A, m^5C Ψ, 2′O-Me,"other"	13	Xuan et al. (2018)
RMDisease	m^6A, m^1A, m^5C, D, f^5C Ψ, 2′O-Me, m^5U, hm^5C m^7G, m^6A_m, I, ac^4C	20	Song et al. (2022b)
m6A2Target	m^6A	Mouse, human	Deng et al. (2021).
m6A-TSHub	m^6A	Human	Song et al. (2022a)
m5C-Atlas	m^5C	13	Ma et al. (2022)
PIANO	Ψ	Human	Song et al. (2020)
Rediportal	I	Mouse, human	Mansi et al. (2021)

navigate MODOMICS and find those that specifically benefit their research.

Navigating putative and known modification pathways

For many RNA modifications, the enzymes that catalyze their formation and the reaction order are unknown. The PATHWAY tab in MODOMICS allows easy visualization of the potential reaction pathways through different line colors and patterns (Fig. 2A). Choosing a nucleoside reveals an array of modifications, arranged in layers and representing additional modifications formed after the precursor modification. Arrows are placed between the modifications, with colors (red, orange, etc.) and types of arrows (solid or dashed) representing the type of reaction and whether it is known or predicted, allowing for examination of the reaction that produces the modification. Clicking on a specific modification shows its chemical structure, the type of RNAs in which it has been identified, and the enzymes that catalyze the modification.

Publications linking modifications to human disease

The HUMAN DISEASE section is a recent addition to MODOMICS. With recent identification of human diseases linked to perturbed RNA modifications, Boccaletto and colleagues provide a collection of publications that support these links. This section allows the user to search for either the enzyme, modification, or disease of interest and displays corresponding publications. Additionally, one can easily export the disease-associated publications of the given enzyme, modification, or disease. Boccaletto and colleagues mention that the list is nonexhaustive and invite the community to assist them in the expansion of this section.

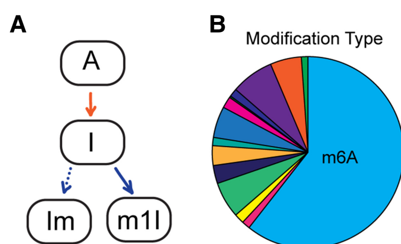

FIGURE 2. (*A*) Example of MODOMICS PATHWAYS tab. Choosing the ADENINE base will show modifications that are catalyzed from adenosine. Solid arrows represent known pathways (see orange and blue) and dashed is predicted (see blue). Orange arrow indicates group exchange and blue represents methylation. (*B*) A representative pie chart from RMDisease showing percentage of modification sites recorded in the database. Currently, m6A (light blue) has the greatest percentage of modification sites recorded in the database.

Visualizing modifications on RNA, tRNA, or rRNA secondary structure

MODOMICS has a collection of RNA sequences from 100 + organisms in the RNA SEQUENCES menu. The readout displays the type and location of modifications found in that sequence. For transfer RNA (tRNA) and ribosomal RNA (rRNA), the user can display modifications on the secondary structure (Fig. 1, rRNA and tRNA). The user has the option to display the secondary structure and modifications for a specific organism or for all organisms. When choosing to visualize for all organisms, the modifications are labeled so a user can quickly determine how likely that nucleotide is to be modified.

RMBase

The original version of RMBase was published in 2016 and the newest release (RMBase v2.0) was updated in 2018 (Sun et al. 2016; Xuan et al. 2018). The purpose for RMBase is to integrate the large epitranscriptome data sets for identification of specific modification sites. This site contains data from 47 studies and 566 samples from 13 species. Modifications included in RMBase are primarily N6-methyladenosine (m^6A), N1-methyladenosine (m^1A), 5-methylcytosine (m^5C), pseudouridine (Ψ), and ribose methylation (2′-O-Me), with an additional section for "other modifications" that contains primarily tRNA modifications. When choosing a specific modification for a given species, the database reports the position (chromosome # and specific location on the chromosome), gene name, and mRNA region (intron, exon, UTR) where the modification is found.

Capability to see all modifications found on an mRNA

One useful feature among many in RMBase is found under the MODGENE menu tab, which allows users to search for a specific gene and quickly identify how many modifications are present on the transcript (sorted by m^6A, m^1A, m^5C, Ψ, 2′-O-Me, and other). Choosing one of the modifications on that mRNA provides the modification's sequence location. The database also provides a SUPPORT NUMBER which indicates the number of publications that support the presence of the modification.

RMDisease v2.0

The first version of RMDisease combined 303,426 RNA modification sites and 40,915,548 somatic and germline single nucleotide polymorphisms (SNPs) to identify 202,307 genetic variants that impact RNA modification (Chen et al. 2021). This has since been expanded in the second version (Song et al. 2022b). The database now contains 873,819 experimentally validated RNA modification

sites, greatly expanding the number of variants that may impact modification status. The original 2021 release of RMDisease included eight modifications: N6-methyladenosine (m^6A), N1-methyladenosine (m^1A), 5-methylcytosine (m^5C), pseudouridine (Ψ), ribose methylation (2′-O-Me), 5-methyluridine (m^5U), N6,2′-O-dimethyladenosine (m^6Am), and 7-methylguanosine (m^7G). The newest release expands the number of modifications to 16 modifications, including inosine (I), N4-acetylcytidine (ac^4C), 5-hydroxymethylcytidine (hm^5C), dihydrouridine (D), and 5-formylcytidine (f^5C). Furthermore, RMDisease predicts how perturbation of these modifications has implications on miRNA binding, splicing sites, and protein binding to the RNA. Finally, identifying modification sites is not limited to just human data sets. A user can identify modified positions in up to 20 species (human, mouse, rat, zebrafish, maize, fly, yeast, fission yeast, *Arabidopsis*, rice, chicken, goat, sheep, pig, cow, rhesus, tomato, chimpanzee, green monkey, and COVID-19). Unfortunately, data sets do not exist for all 16 modifications for every species. The authors state that every potential modification should be experimentally validated, and their website contains detailed instructions on how to use RMDisease; some features are described below.

Comparison between modifications and species

When the user hovers over the MODIFICATION tab and chooses a modification, three pie charts are displayed; see example of one in Fig. 2B. The first shows how many modification sites are present in the collected data sets and how this number compares to other modifications in the database. For example, m^6A currently makes up 60% of all modification sites in the database (Fig. 2B, light blue). The second pie chart is a comparison of the number of modification sites identified relative to other species (example, human vs. mouse vs. cow). Finally, the third pie chart indicates what percentage of SNPs potentially could cause loss and/or gain of modification.

Quick identification of potential impacts on protein binding and microRNAs

After choosing a modification from a species, the interface will display a list of RNAs that contain a SNP that may impact the modification status. This list can be further FILTERED into different RNAs, such as mRNAs and tRNAs. The list also indicates whether the SNP could influence the binding of RNA binding proteins (RBPs) and microRNAs (miRNAs). While the chart provides the number of RBP and miRNA binding sites that could be impacted, the JBROWSER button provides a more in-depth view of the SNP versus binding, including lists of protein binding sites and miRNAs which may warrant additional investigation.

SINGLE MODIFICATION DATABASES

There are many databases for specific modifications, especially abundant modifications such as N6-methyladenosine (m^6A), N1-methyladenosine (m^1A), 5-methylcytosine (m^5C), pseudouridine (Ψ), inosine (I), and ribose methylation (2′-O-Me). The presence of functional links and ease of use varies between these databases. Highlighted below are selected databases for a few well-studied modifications.

N6-methyladenosine (m^6A)

The m^6A modification is a prevalent and reversible modification found on multiple types of RNA (Zaccara et al. 2019). While there are multiple m^6A databases, highlighted below are M6A2Target and M6A-TSHub.

M6A2Target is a repository of m^6A writers, readers and erasers and their targets (Deng et al. 2021). The user can identify "VALIDATED TARGETS" which have been determined through low-throughput experiments or "POTENTIAL TARGETS" which were identified through high-throughput studies. This site contains information for two species, human and mouse, and a variety of cell lines. The user can sort by gene name or by the ten readers, ten writers, or two erasers. Exploring both VALIDATED TARGETS and POTENTIAL TARGETS provides similar information, but a key difference is that the VALIDATED TARGETS are linked to a PUBMED ID. Furthermore, the DETAILS button informs the user of the target mRNA site and provides information for studies that were performed perturbing the writer, reader, or eraser. If perturbation experiments were performed, observed phenotypes are listed.

The purpose of M6A-TSHub is to connect the differences between m^6A methylation in diverse tissues and diseases. Information is provided for 23 different tissues (with approximately ∼185,000 mapped sites) and 25 tumor samples with 500,000 mapped sites. This database allows for identification of SNPs in human tissues and could be used to help users identify if variants in a specific gene could impact m^6A sites (Song et al. 2022a).

5-Methylcytosine (m^5C) and pseudouridine (Ψ)

m^5C and Ψ are prevalent RNA modifications that influence many biological functions (Spenkuch et al. 2014; Xue et al. 2020). Published in 2022, the database m^5C-Atlas provides information on m^5C sites on mRNA, tRNA, and rRNA. This information can be sorted by 12 different species and one virus (human immunodeficiency virus), or a user can search a gene of interest (Ma et al. 2022). Released in 2020, the database PIANO (pseudouridine site identification and functional annotation) has many similar features. Specifically, PIANO is a repository of

identified Ψ sites. Information provided includes the TECHNIQUE in which the site was identified and a PUBMED ID. Additionally, predictions are provided if a Ψ site is associated with known protein binding sites and/or miRNA-targeting sequences (Song et al. 2020).

Inosine (I)

The inosine RNA modification is catalyzed through an adenosine deamination reaction and is important in many RNAs. For example, inosine is an essential tRNA modification influencing wobble decoding (Gerber and Keller 1999). Inosine is also found abundantly in Alu elements in mRNA, influencing the innate immune system (Mannion et al. 2014; Chung et al. 2018).

One database that provides A-to-I editing information is Rediportal, originally developed in 2017 with a 2021 update (Picardi et al. 2017; Mansi et al. 2021). The original database cataloged 4.5 million A-to-I editing events across 55 body sites categorized into 30 tissues (∼2500 RNA-seq data sets). This has been expanded to 16 million A-to-I editing events from ∼9600 RNA-seq data sets and the updated database now includes mouse. When the user searches for a gene of interest, the database returns all isoforms and annotates important regions such as UTRs, exons, and introns. The RNA editing profile is located below the isoforms, showing the editing distribution across the multiple isoforms. Clicking on a single isoform provides additional information on the editing event such as the location (chromosome # and location on that chromosome) and where in the mRNA the editing event is found (example, 5′ UTR). Additional information includes the number of samples in which the editing event was detected. The window can be expanded, revealing the editing levels of that specific event in all body sites. This data can be displayed in multiple ways, such as heat maps or box plots.

OUTLOOK

The epitranscriptomics field is expanding. New modifications are still being discovered, such as phosphorylation of tRNA and the discovery of glycoRNAs (Flynn et al. 2021; Ohira et al. 2022). Methods to identify modifications, both at a whole genome scale (sequencing methods) and at specific nucleotide resolution, continue to improve. Researchers are recognizing the need for collaborative databases to share the abundant emerging information. Unfortunately, while many databases are developed, many links to these are not functional or the databases are not continuously maintained. The databases discussed in this mini-review have an excellent user-friendly interface and rich information from a modification location, link to disease, and information about the discovery. Importantly, this list is nonexhaustive and users are encouraged to explore other databases that may benefit their needs.

Looking ahead, public repositories of new data where the authors can quickly update the databases with new information will continue to be critical resources. This is especially true as differences in modifications across species and tissues are discovered. Additionally, databases that connect a given modification to the method of discovery will be particularly useful, including linking newly developed methods for specific modifications. Finally, we anticipate advancements in modeling which can hypothesize how modifications on an RNA may influence binding of another RNA (e.g., mRNA and tRNA) or protein (e.g., RNA binding proteins), and thus provide ideas for additional experiments. These tools will be important as the exciting field of epitranscriptomics continues to develop and expand.

ACKNOWLEDGMENTS

Thank you to those who read and provided edits on the manuscript: Dragony Fu (University of Rochester), Jeffrey Kieft (CU Anschutz), and members of the Kieft Laboratory (Conner Langeberg, Parker Nichols, David Costantino). Thank you to the RNA Society for the opportunity to write this review.

REFERENCES

Andries O, Mc Cafferty S, De Smedt SC, Weiss R, Sanders NN, Kitada T. 2015. N^1-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. *J Control Release* **217:** 337–344. doi:10.1016/j.jconrel.2015.08.051

Boccaletto P, Stefaniak F, Ray A, Cappannini A, Mukherjee S, Purta E, Kurkowska M, Shirvanizadeh N, Destefanis E, Groza P, et al. 2022. MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res* **50:** D231–D235. doi:10.1093/nar/gkab1083.

Cai WM, Chionh YH, Hia F, Gu C, Kellner S, McBee ME, Ng CS, Pang YL, Prestwich EG, Lim KS, et al. 2015. A platform for discovery and quantification of modified ribonucleosides in RNA: application to stress-induced reprogramming of tRNA modifications. *Methods Enzymol* **560:** 29–71. doi:10.1016/bs.mie.2015.03.004

Chen K, Song B, Tang Y, Wei Z, Xu Q, Su J, de Magalhaes JP, Rigden DJ, Meng J. 2021. RMDisease: a database of genetic variants that affect RNA modifications, with implications for epitranscriptome pathogenesis. *Nucleic Acids Res* **49:** D1396–DD404. doi:10.1093/nar/gkaa790

Chung H, Calis JJA, Wu X, Sun T, Yu Y, Sarbanes SL, Dao Thi VL, Shilvock AR, Hoffmann HH, Rosenberg BR, et al. 2018. Human ADAR1 prevents endogenous RNA from triggering translational shutdown. *Cell* **172:** 811–24.e14. doi:10.1016/j.cell.2017.12.038

Cohn WE. 1960. Pseudouridine, a carbon-carbon linked ribonucleoside in ribonucleic acids: isolation, structure, and chemical characteristics. *J Biol Chem* **235:** 1488–1498. doi:10.1016/S0021-9258(18)69432-3

Cozen AE, Quartley E, Holmes AD, Hrabeta-Robinson E, Phizicky EM, Lowe TM. 2015. ARM-seq: alkB-facilitated RNA methylation

sequencing reveals a complex landscape of modified tRNA fragments. *Nat Methods* **12:** 879–884. doi:10.1038/nmeth.3508

Davis FF, Allen FW. 1957. Ribonucleic acids from yeast which contain a fifth nucleotide. *J Biol Chem* **227:** 907–915. doi:10.1016/S0021-9258(18)70770-9

Deng S, Zhang H, Zhu K, Li X, Ye Y, Li R, Liu X, Lin D, Zuo Z, Zheng J. 2021. M6A2Target: a comprehensive database for targets of m^{6}A writers, erasers and readers. *Brief Bioinform* **22:** bbaa055. doi:10.1093/bib/bbaa055

Dunin-Horkawicz S, Czerwoniec A, Gajda MJ, Feder M, Grosjean H, Bujnicki JM. 2006. MODOMICS: a database of RNA modification pathways. *Nucleic Acids Res* **34:** D145–D149. doi:10.1093/nar/gkj084

Flynn RA, Pedram K, Malaker SA, Batista PJ, Smith BAH, Johnson AG, George BM, Majzoub K, Villalta PW, Carette JE, et al. 2021. Small RNAs are modified with N-glycans and displayed on the surface of living cells. *Cell* **184:** 3109–3124.e22. doi:10.1016/j.cell.2021.04.023

Gerber AP, Keller W. 1999. An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* **286:** 1146–1149. doi:10.1126/science.286.5442.1146

Grosjean H. 2015. RNA modification: the Golden Period 1995-2015. *RNA* **21:** 625–626. doi:10.1261/rna.049866.115

Grosjean H, Keith G, Droogmans L. 2004. Detection and quantification of modified nucleotides in RNA using thin-layer chromatography. *Methods Mol Biol* **265:** 357–391. doi:10.1385/1-59259-775-0:357

Jonkhout N, Tran J, Smith MA, Schonrock N, Mattick JS, Novoa EM. 2017. The RNA modification landscape in human disease. *RNA* **23:** 1754–1769. doi:10.1261/rna.063503.117

Karijolich J, Yu YT. 2010. Spliceosomal snRNA modifications and their function. *RNA Biol* **7:** 192–204. doi:10.4161/rna.7.2.11207

Kariko K, Buckstein M, Ni H, Weissman D. 2005. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* **23:** 165–175. doi:10.1016/j.immuni.2005.06.008

Kowalak JA, Pomerantz SC, Crain PF, McCloskey JA. 1993. A novel method for the determination of post-transcriptional modification in RNA by mass spectrometry. *Nucleic Acids Res* **21:** 4577–4585. doi:10.1093/nar/21.19.4577

Leger A, Amaral PP, Pandolfini L, Capitanchik C, Capraro F, Miano V, Migliori V, Toolan-Kerr P, Sideri T, Enright AJ, et al. 2021. RNA modifications detection by comparative Nanopore direct RNA sequencing. *Nat Commun* **12:** 7198. doi:10.1038/s41467-021-27393-3

Ma J, Song B, Wei Z, Huang D, Zhang Y, Su J, de Magalhaes JP, Rigden DJ, Meng J, Chen K. 2022. m^{5}C-Atlas: a comprehensive database for decoding and annotating the 5-methylcytosine (m^{5}C) epitranscriptome. *Nucleic Acids Res* **50:** D196–D203. doi:10.1093/nar/gkab1075

Mannion NM, Greenwood SM, Young R, Cox S, Brindle J, Read D, Nellaker C, Vesely C, Ponting CP, McLaughlin PJ, et al. 2014. The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. *Cell Rep* **9:** 1482–1494. doi:10.1016/j.celrep.2014.10.041

Mansi L, Tangaro MA, Lo Giudice C, Flati T, Kopel E, Schaffer AA, Castrignano T, Chillemi G, Pesole G, Picardi E. 2021. REDIportal: millions of novel A-to-I RNA editing events from thousands of RNAseq experiments. *Nucleic Acids Res* **49:** D1012–D1D19. doi:10.1093/nar/gkaa916

Motorin Y, Marchand V. 2021. Analysis of RNA modifications by second- and third-generation deep sequencing: 2020 update. *Genes (Basel)* **12:** 278. doi:10.3390/genes12020278

Motorin Y, Muller S, Behm-Ansmant I, Branlant C. 2007. Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol* **425:** 21–53. doi:10.1016/S0076-6879(07)25002-5

Ohira T, Minowa K, Sugiyama K, Yamashita S, Sakaguchi Y, Miyauchi K, Noguchi R, Kaneko A, Orita I, Fukui T, et al. 2022. Reversible RNA phosphorylation stabilizes tRNA for cellular thermotolerance. *Nature* **605:** 372–379. doi:10.1038/s41586-022-04677-2

Picardi E, D'Erchia AM, Lo Giudice C, Pesole G. 2017. REDIportal: a comprehensive database of A-to-I RNA editing events in humans. *Nucleic Acids Res* **45:** D750–DD57. doi:10.1093/nar/gkw767

Porman AM, Roberts JT, Duncan ED, Chrupcala ML, Levine AA, Kennedy MA, Williams MM, Richer JK, Johnson AM. 2021. A single N6-methyladenosine site in lncRNA HOTAIR regulates its function in breast cancer cells. bioRxiv doi:10.1101/2020.06.08.140954

Schaefer M, Kapoor U, Jantsch MF. 2017. Understanding RNA modifications: the promises and technological bottlenecks of the 'epitranscriptome'. *Open Biol* **7:** 170077. doi:10.1098/rsob.170077

Song B, Tang Y, Wei Z, Liu G, Su J, Meng J, Chen K. 2020. PIANO: a web server for pseudouridine-site (Psi) identification and functional annotation. *Front Genet* **11:** 88. doi:10.3389/fgene.2020.00088

Song B, Huang D, Zhang Y, Wei Z, Su J, Pedro de Magalhaes J, Rigden DJ, Meng J, Chen K. 2022a. m6A-TSHub: unveiling the context-specific m^{6}A methylation and m6A-affecting mutations in 23 human tissues. *Genomics Proteomics Bioinformatics.* doi:10.1101/2022.01.12.476117

Song B, Wang X, Liang Z, Ma J, Huang D, Wang Y, Pedro de Magalhaes J, Rigden DJ, Meng J, Liu G, et al. 2022b. RMDisease V2.0: an updated database of genetic variants that affect RNA modifications with disease and trait implication. *Nucleic Acids Res* doi:10.1093/nar/gkac750

Spenkuch F, Motorin Y, Helm M. 2014. Pseudouridine: still mysterious, but never a fake (uridine)!. *RNA Biol* **11:** 1540–1554. doi:10.4161/15476286.2014.992278

Sun WJ, Li JH, Liu S, Wu J, Zhou H, Qu LH, Yang JH. 2016. RMBase: a resource for decoding the landscape of RNA modifications from high-throughput sequencing data. *Nucleic Acids Res* **44:** D259–D265. doi:10.1093/nar/gkv1036

Xu H, Yao J, Wu DC, Lambowitz AM. 2019. Improved TGIRT-seq methods for comprehensive transcriptome profiling with decreased adapter dimer formation and bias correction. *Sci Rep* **9:** 7953. doi:10.1038/s41598-019-44457-z

Xuan JJ, Sun WJ, Lin PH, Zhou KR, Liu S, Zheng LL, Qu LH, Yang JH. 2018. RMBase v2.0: deciphering the map of RNA modifications from epitranscriptome sequencing data. *Nucleic Acids Res* **46:** D327–DD34. doi:10.1093/nar/gkx934

Xue C, Zhao Y, Li L. 2020. Advances in RNA cytosine-5 methylation: detection, regulatory mechanisms, biological functions and links to cancer. *Biomark Res* **8:** 43. doi:10.1186/s40364-020-00225-0

Zaccara S, Ries RJ, Jaffrey SR. 2019. Reading, writing and erasing mRNA methylation. *Nat Rev Mol Cell Biol* **20:** 608–624. doi:10.1038/s41580-019-0168-5