



Nanopore-Based Detection of Viral RNA Modifications

Jonathan S. Abebe,^a Ruth Verstraten,^{b,c}  Daniel P. Depledge^{a,b,c}

^aDepartment of Microbiology, New York University School of Medicine, New York, New York, USA

^bInstitute of Virology, Hannover Medical School, Hannover, Germany

^cGerman Center for Infection Research (DZIF), partner site Hannover-Braunschweig, Hannover, Germany

ABSTRACT The chemical modification of ribonucleotides plays an integral role in the biology of diverse viruses and their eukaryotic host cells. Mapping the precise identity, location, and abundance of modified ribonucleotides remains a key goal of many studies aimed at characterizing the function and importance of a given modification. While mapping of specific RNA modifications through short-read sequencing approaches has powered a wealth of new discoveries in the past decade, this approach is limited by inherent biases and an absence of linkage information. Moreover, in viral contexts, the challenge is increased due to the compact nature of viral genomes giving rise to many overlapping transcript isoforms that cannot be adequately resolved using short-read sequencing approaches. The recent emergence of nanopore sequencing, specifically the ability to directly sequence native RNAs from virus-infected host cells, provides not just a new methodology for mapping modified ribonucleotides but also a new conceptual framework for what can be derived from the resulting sequencing data. In this minireview, we provide a detailed overview of how nanopore direct RNA sequencing works, the computational approaches applied to identify modified ribonucleotides, and the core concepts underlying both. We further highlight recent studies that have applied this approach to interrogating viral biology and finish by discussing key experimental considerations and how we predict that these methodologies will continue to evolve.

KEYWORDS RNA modifications, m⁶A, nanopore sequencing, virus

Since the discovery of pseudouridine (ψ) in 1957 (1), over 170 chemically modified ribonucleotides have been identified (2). Despite this, our knowledge of the function and modulation of modified ribonucleotides is generally sparse, with a few important exceptions. Within the eukaryotic cell, all major RNA species harbor modified ribonucleotides (3–9), with rRNAs and tRNAs among the most heavily modified species (reviewed in references 10 and 11). To date, the best characterized and most widely studied RNA modification is N⁶-methyladenosine (m⁶A), first discovered in 1975 (12, 13). This modification is found in mRNA (12), long intergenic noncoding RNAs (lincRNAs) (14), primary microRNAs (pri-miRNAs) (7), and rRNAs (15, 16). For mRNAs, miRNAs, and lincRNAs, m⁶A is installed cotranscriptionally in the nucleus by a methyltransferase complex minimally comprising METTL3 and METTL14 and enhanced by WT1 associated protein (WTAP) and a range of RNA polymerase II accessory proteins (17–19). m⁶A is recognized and bound by multiple YTH family members (YTHDF1–3, YTHDC1–2) along with a variety of hnRNPs and multiple members of the IGF2BP family (20–23). m⁶A is considered a reversible modification and may be removed by the demethylases ALKBH5 and FTO (24–26). While installation of m⁶A generally occurs in DRACH (D = A/G/U; R = G/A; H = A/C/U) sequence contexts, most DRACH sequences are not modified (27). m⁶A has been implicated in regulating numerous diverse cellular processes involved in RNA maturation and function, including splicing, polyadenylation, export, translation, and decay (24, 28–32). Beyond m⁶A, other important mRNA modifications include pseudouridine (ψ), N⁶,2'-O-dimethyladenosine (m⁶Am) (26, 33–36), 5-methylcytidine (m⁵C) (37, 38), 1-methyladenosine (m¹A) (39–41), 7-methylguanosine (m⁷G) (42–44), and 2'-O-methylation (Nm)

Editor Jacob Yount, Ohio State University

Copyright © 2022 Abebe et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Daniel P. Depledge, depledge.daniel@mh-hannover.de.

The authors declare no conflict of interest.

Published 17 May 2022

(45–47). For an extended overview of these, we direct the reader to the MODOMICS database resource (<http://genesilico.pl/modomics>) (2).

Unsurprisingly, the influence of RNA modifications on viral biology is complex, and many viruses have evolved exquisite mechanisms for interacting with RNA modification pathways. For viruses which replicate in the nucleus, m⁶A installed on host and viral mRNAs has been shown to have both proviral and antiviral effects (48). These include regulating viral gene expression in the context of HIV-1 (49–51), impairing nuclear processing and export in simian virus 40 (SV40) (52, 53), modulating the splicing ability of adenoviral transcripts (54), preventing m⁶A installation on beta interferon-encoding mRNAs to inhibit type 1 interferon responses (55, 56), and regulating the life cycle of Kaposi's sarcoma-associated herpesvirus (KSHV) (57–60). The installation of m⁶A on the RNAs of viruses which replicate in the cytoplasm similarly confers proviral or antiviral effects such as interfering with replication of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) (61, 62) and enterovirus 71 (63). Other RNA modifications, including pseudouridine, m⁵C, Nm, m⁶Am, N¹-methylguanosine (m¹G), and N⁴-acetylcytidine (ac⁴C), have all been shown to be installed on viral mRNAs (reviewed in references 64 and 65), while modifications such as m⁵C, m¹A, Nm, m⁷G, m¹G, and m⁶Am have variously been identified on the genomes of several RNA viruses (66). These modifications all play fundamental roles in viral life cycles. As with the analysis of eukaryotic cells, the location and abundance of modified ribonucleotides have typically been mapped using short-read sequencing and/or mass spectrometry approaches. For one of the most comprehensive overviews to date on the roles of RNA modifications in viral biology, we recommend a recent review by Baquero-Perez et al. (65).

While several techniques enable the identification of RNA modifications on individual RNA isoforms, many of these are low throughput and thus useful only in certain contexts (67). This changed in 2012, when the first high-throughput sequencing approaches for transcriptome-wide mapping of m⁶A were published (27, 68). Termed meRIP-Seq (methylated RNA immunoprecipitation sequencing) or m⁶A-Seq (m⁶A sequencing), these techniques combine immunoprecipitation with antibodies against m⁶A and subsequent short-read (Illumina) sequencing to enrich for RNA fragments containing m⁶A, prepared in parallel with either a nonspecific or no immunoprecipitation control. Once sequenced, reads from m⁶A-containing RNA fragments are enriched over the background control at m⁶A sites and provide a resolution range of 50 to 200 nucleotides (nt). These studies were the first to determine that m⁶A modifications are both abundant and enriched within long internal exons and in 3' untranslated regions (UTRs). This was later followed by newer antibody-dependent protocols such as miCLIP-Seq (m⁶A individual-nucleotide-resolution cross-linking and immunoprecipitation sequencing) (3), which was used to map m⁶A sites at nucleotide-level resolution. Despite the great success of these protocols, it should be noted that antibody-based approaches suffer from artifacts associated with off-target capture (69). The development of antibody-independent methods such as DART-Seq (deamination adjacent to RNA modification target sequencing) (70), Sequencing of RNA digested via m⁶A sensitive RNase (MAZTER-Seq) (71), m⁶A-label-seq (a metabolic labeling method to detect mRNA m⁶A transcriptome-wide at base resolution) (72), m⁶A-SEAL (an antibody-free, FTO-assisted chemical labeling method) (73), and m⁶A-SAC-seq (m⁶A-selective allyl chemical labeling and sequencing) (74) provide a useful set of orthologous m⁶A-specific methodologies for mapping the locations at which m⁶A is installed. While impressive, these techniques remain constrained by the biases and nonbiological variation associated with short-read sequencing approaches (69, 75), as well as the loss of linkage that comes from fragmenting individual RNA molecules. Moreover, in a viral context, short-read sequencing is generally not suitable for disentangling the complex overlapping RNA isoforms that are characteristic of most viral species (76). While this does not preclude the use of the above discussed approaches to interrogate viral biology, it does limit the conclusions that can be made. With this in mind, we define the ultimate aim of RNA modification detection approaches as follows: to detect, at single-level nucleotide resolution, the presence and identity of all modified ribonucleotides installed on a single RNA. By nature, short-read sequencing approaches cannot fulfill this goal. Indeed, the optimal way to accomplish this would require direct full-length

sequencing of RNAs in such a way that the position and identity of any modified ribonucleotides with a given RNA can be discerned.

NANOPORE DIRECT RNA SEQUENCING

In 2017, Oxford Nanopore Technologies (ONT) described a first approach to sequencing the full length of native RNA molecules extracted from a cellular context (77). They termed this technique direct RNA sequencing (dRNA-Seq or DRS). The construction of DRS libraries remains remarkably quick and simple at ~ 2 h. The standard approach utilizes 250 to 500 ng of isolated poly(A) RNA as input for a ligation reaction in which a double-stranded DNA adaptor is ligated to the 3' end of the poly(A) tail (77). Variations to this approach include the use of custom adaptors that allow targeting of specific nonpolyadenylated RNAs (78) or the use of poly(A) and rRNA depletion prior to artificial polyadenylation and capture, e.g., to capture nascent transcription (79). Following ligation, a reverse transcription step produces an RNA-cDNA duplex that stabilizes the RNA strand and reduces secondary structure formation before a second ligation allows the attachment of an engineered motor protein that allows nanopore docking and unwinding of the RNA-cDNA duplex. The DRS library is subsequently loaded onto a flow cell for sequencing. The flow cell contains an electrically resistant polymer membrane separating two electrodes, across which can be applied an electrical current. A series of protein nanopores are embedded within the polymer, each with its own sensor chip that continuously records the flow of ions (ionic current) passing through the nanopore. Upon docking with a nanopore, the motor protein unwinds the RNA-cDNA duplex, ratcheting the RNA molecule through the nanopore in the 3'-to-5' direction at a constant rate (~ 70 nt/s) (Fig. 1a). The flow of ionic current is measured within the narrowest part of the nanopore by an ammeter with contact points set a distance apart that roughly corresponds to 5 nucleotides (Fig. 1b). Thus, at any given time, the ionic current recorded generally represents a function of 5 nucleotides (5-mer). The composition of nucleotides in this 5-mer is subsequently derived by "base calling." Here, ionic current changes are segmented as discrete "events" in which the duration (dwell time), amplitude (signal intensity), and variance are recorded. Machine-learning (ML) (Fig. 2) algorithms (for an excellent review of ML geared toward biologists, see reference 80) subsequently convert "events" into "base calls," with each identified nucleotide being assigned a probability score that denotes the predicted accuracy of the call. While several base calling algorithms exist, the most commonly used is Guppy. Developed by ONT, Guppy uses a recurrent neural network (a form of deep learning [81]) trained on synthetic data sets comprising 5-mers of all possible A, C, G, and T or U combinations to predict the underlying nucleotide sequences with variable accuracy: 94 to 98% for DNA and 86 to 91% for RNA (82–84). At the DNA level of Guppy base calling, the training data sets have recently undergone further expansion to enable the native detection of specific DNA modifications (6mA and 5mC). However, such support is not yet available for RNA modifications. This is due in part to the vast numbers and diversity of RNA modifications (> 170) that exist. Consider, for example, a 5-mer of unmodified ribonucleotides (A, C, G, U). Here, there are 4^5 (1,024) potential sequence combinations that must be considered during base calling. Add in just a few modified ribonucleotides such as m^1A , m^6A , and pseudouridine, and the number of sequence combinations changes to 7^5 (16,807). There thus remains a burden on finding alternative approaches to detecting RNA modifications in DRS data sets, a challenge that has been embraced by the wider scientific community. At the time of writing, nanopore DRS approaches had been used to detect a wide range of RNA modifications, including m^6A , internal m^7G , m^5C , Ψ , Nm, and 5-hydroxymethylcytosine (5hmC) (85). This was made possible by distinct analytical approaches that we broadly classify as "error rate" and "signal-level" methodologies (Fig. 1c to f; Table 1).

INFERRING RNA MODIFICATIONS THROUGH BASE CALL ERROR RATES

The presence of modified ribonucleotides within an RNA has a marked effect on the ionic flow passing through a nanopore, a feature that directly affects the accuracy of downstream base calling. This generally manifests as an increased probability of one or more wrong nucleotides being assigned to a given 5-mer. By calculating the error rate (% incorrect base calls) at

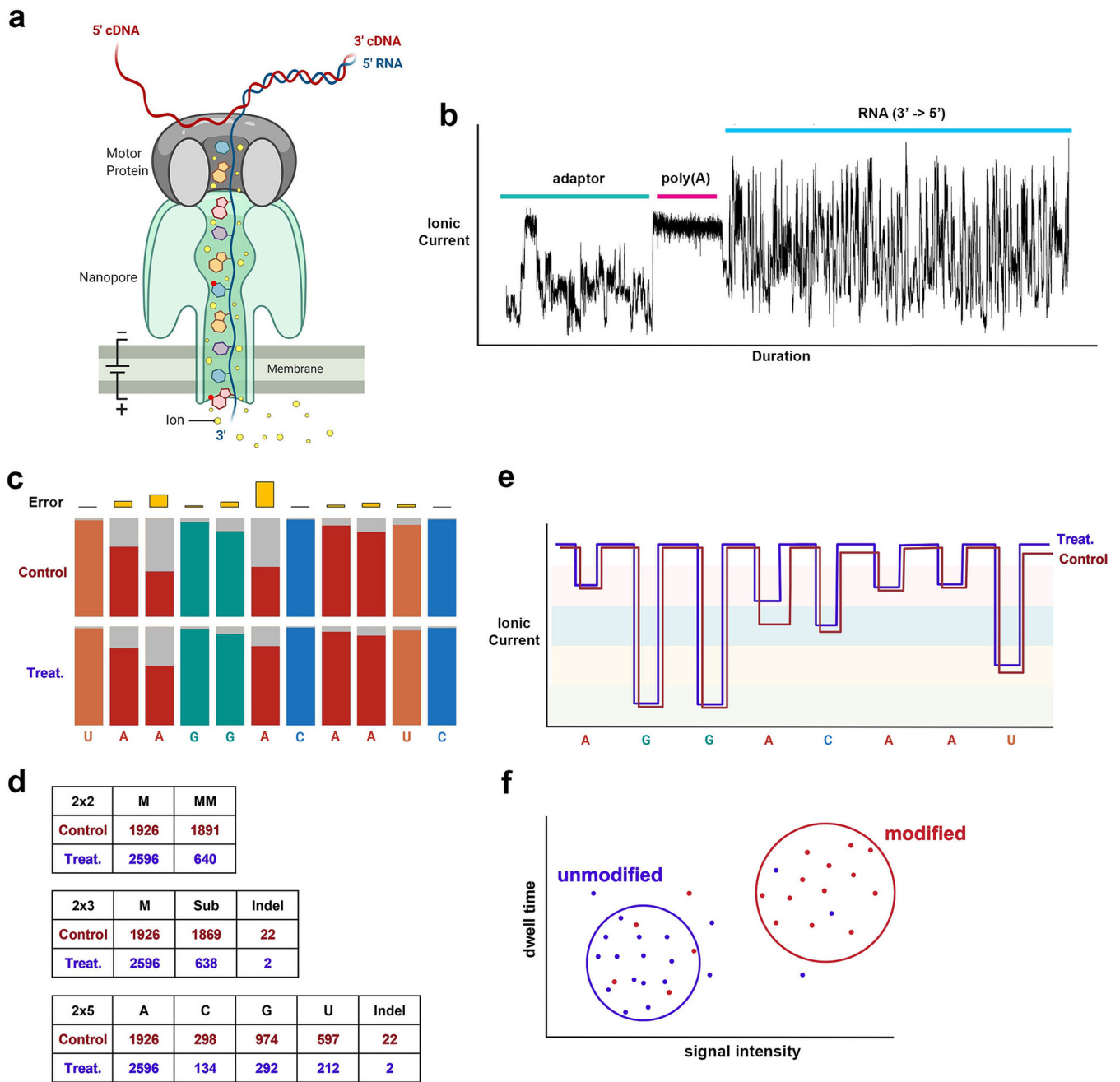


FIG 1 Overview of the nanopore direct RNA sequencing approach and RNA modification detection strategies. (a and b) DRS libraries comprise RNA-cDNA hybrids that are unwound by a helicase when docked to a membrane-embedded nanopore. (a) The RNA strand is selectively pushed through the nanopore in a 3'-to-5' direction, disturbing the general flow of ions through the pore. (b) Ammeters placed within the narrowest part of the pore measure changes in ionic flow (current), which can subsequently be partitioned into sections attributed to the DNA adaptor, poly(A) tail, and body of the sequenced RNA, here visualized as a squiggle plot (124). (c and d) Putatively modified ribonucleotides are identified through either error rate or signal-level analyses. (c) Error rate analyses identify individual positions within RNAs that show significantly different error profiles between two conditions (e.g., control versus treatment). Here, the proportion of reads containing the correct nucleotide at a given position is shown in colors other than gray, while the proportion of reads containing erroneous bases at the same position is shown in gray. The difference in calculated error rate is shown in gold. (d) Error profiles undergo statistical analyses in the form of 2×2 , 2×3 , or 2×5 contingency tables that are subsequently corrected for multiple testing. M, match; MM, mismatch. (e) For signal-level analyses, the initial process of resquigging assigns ionic flow features (signal intensity, dwell time, etc.) to their corresponding nucleotides within individual reads. (f) These features either undergo a comparative analysis (between data sets) or are evaluated using data from a prebuilt model generated by machine and/or deep learning.

each position across a set of identical reads, one can identify hot spots of increased error rate. However, hot spots may also originate from natural within-sample polymorphisms and the presence of homopolymers, as well as stochastic factors. To correct for this, most error rate approaches utilize a comparative strategy in which hot spots are considered real only if they appear in a data set known to contain a given modification (e.g., m⁶A) compared to an

Machine learning (ML) approaches use special algorithms to 'learn' relationships between different data features within a multivariate data set (two or more data features) and apply this knowledge to make predictions on similar data sets. **Unsupervised ML** approaches cluster user-defined data features in the absence of data labels. A simple example is the clustering of data points according to signal intensity and dwell time across two dimensions (Figure 1f). By contrast, **Supervised ML** approaches examine the relationships between user-defined data features (e.g., signal intensity, k-mer sequence) and user-defined data labels (e.g., modified, unmodified) in a given data set and can subsequently be made to use predictions on new data sets. **Ensemble learning methods** use predictions obtained from multiple supervised learning algorithms to produce a consensus. ML approaches are more powerful than standard univariate (single data feature) statistical approaches such as *t*-tests. For multivariate data sets with large numbers of data features, the application of **dimensional reduction techniques** simplifies the analysis by identifying and using only the most important features.

FIG 2 Brief overview of machine learning.

identical data set in which the modification is absent (Fig. 1c). This latter strategy is achieved in many ways, including inhibition, silencing, or knockout of specific proteins, or else through the generation of *in vitro*-transcribed (IVT) RNA. Putative modified ribonucleotides are subsequently identified through one or more statistical tests, usually based on contingency tables (Fig. 1d), and the inclusion of biological replicates. At the time of writing, five softwares applying the error rate approach had been released (Table 1).

EpiNano was the first software to make use of the error rate approach by relying on a supervised learning approach (in this case, a support vector machine algorithm trained on synthetic m⁶A-modified and unmodified *in vitro*-transcribed constructs) to predict m⁶A modifications at RRACH motifs within yeast strains (86). This has since been replaced by a comparative approach termed EpiNano-Error, which uses features such as base quality score, the presence of indels, and/or mismatches to identify differences in error patterns between two matched samples (i.e., control versus treatment) and was recently used to detect m⁶A within circular RNAs (circRNAs) (87). Another tool, Eligos2, predicts modified sites by considering the difference in error rate (between two conditions) at a given position as well as at the two upstream and downstream flanking positions. Each position is tested using a Fisher exact test on a 2 × 2 contingency table containing matches and mismatches (substitutions, insertions, and deletions) (Fig. 1d). *P* values from each test are corrected via the Benjamini-Hochberg

TABLE 1 Nanopore-based RNA modification detection softwares and their characteristics

Analysis type and tool	Modification(s) reported ^a	Modifications targeted ^b	Approach	Stoichiometry estimates	Reference
Error rate					
DRUMMER	m ⁶ A	Any	Comparative	No	54
Eligos2	m ⁶ A	Any	Individual/comparative	No	88
JACUSA2	m ⁶ A	Any	Comparative	No	90
DiffErr	m ⁶ A	Any	Comparative	No	89
EpiNano	m ⁶ A, Ψ	Any	Individual/comparative	No	86
Signal level					
Tombo	m ⁶ A/m ⁵ C	Any	Individual/comparative	No	92
MINES	m ⁶ A	m ⁶ A only	Individual	No	100
nanom6A	m ⁶ A	m ⁶ A only	Individual	No	101
nanoRMS	Ψ, Nm	Any	Comparative	Yes	98
Nanocompore	m ⁶ A, Ψ	Any	Comparative	No	96
m6Anet	m ⁶ A	m ⁶ A only	Individual	Yes	102
xPore	m ⁶ A	Any	Individual/comparative	Yes	97
Penguin	Ψ	Ψ only	Individual	No	105
Yanocomp	m ⁶ A	Any	Comparative	Yes ^c	99
NanoPsu	Ψ	Ψ only	Individual	Yes	106
DENA	m ⁶ A	m ⁶ A only	Individual	No	104

^am⁶A, N⁶-methyladenosine; m⁵C, 5-methylcytosine; Ψ, pseudouridine; Nm, 2'-O-methylation.

^bAny, theoretically capable of detecting any RNA modification that alters ionic current.

^cStoichiometry estimates claimed but not evidenced at time of writing.

method and an odds ratio test used to identify putatively modified positions. Synthetically modified IVT RNA was initially used to make predictions on nine different RNA modifications (m^6A , m^1A , ψ , m^7G , inosine, hm^5C , m^5C , 5-methoxyuridine [5moU], and 5-formylcytosine [f^5C]), with resulting area under the receiver operating characteristic (ROC) curves (AUCs) ranging from ~ 0.54 to 0.92 (i.e., poor to great). Eligos2 was subsequently used to demonstrate that the sequence contexts of putative m^6A sites identified in yeast and human data sets were enriched for DRACH motifs (88). DRUMMER parses each position within a comparative alignment into a 2×5 contingency table (Fig. 1d) that is evaluated using a G test and an odds ratio test. The resulting P values are corrected for multiple testing, and a list of candidate sites are returned. Further filtering of data can be performed by estimating and correcting for background noise. An early version of DRUMMER was used to identify m^6A within adenovirus type 5 and mouse transcriptome data sets (54). Conceptually similar to DRUMMER, DiffErr also parses comparative alignments into a series of 2×5 contingency tables, one per position in the alignment. A G test is applied and, where P is < 0.05 , a G test for homogeneity between replicates of the same condition is performed. Positions at which the G test score between replicates is higher than the G test score between conditions are excluded, and the remaining positions undergo multiple testing correction. Sites with an adjusted P value of < 0.05 and a log fold change (logFC) of > 1 in control versus treatment samples are classified as methylated. The authors applied DiffErr to study m^6A in *Arabidopsis thaliana* with 66% of predicted sites being within five bases of a miCLIP peak (89).

While each of the above tools can theoretically detect any RNA modification that produces current fluctuations during nanopore sequencing, JACUSA2 is presently limited to m^6A detection. JACUSA2 considers three potential events that may arise from errors (substitutions, insertions, deletions) and examines only 5-mers with a central A (e.g., NNANN). This results in a 30-feature matrix (3 events, 5 k-mers, 2 conditions) that subsequently undergoes dimensionality reduction to identify characteristic patterns of m^6A modifications. Using a comparative approach, the authors identified multiple m^6A sites located in DRACH sequence contexts, and when this was extended to focus on the *EEF2* gene locus, they reported both a novel and previously identified m^6A site (90).

INFERRING RNA MODIFICATIONS THROUGH CHANGES IN SIGNAL INTENSITY

Signal-level approaches detect modified ribonucleotides by directly leveraging the raw electrical signal information collected during nanopore sequencing runs. As a first step, every signal-level approach requires “resquigglng” of the data. Resquigglng describes the process by which the raw signal for each individual read is dissected into “events” that are overlaid onto read alignments by using specialist algorithms included in the nanopolish (91) and Tombo (92) packages (Fig. 1e). The resulting signal information is used to either (i) train a model or (ii) perform comparative analyses between two experimental conditions (Table 1). Training a model uses the k-mer information and signal intensities as features and requires a list of known modification statuses to make a prediction. Software packages invoking this method thus have the benefit of running without a control sample, if the experimentally desired model already exists. However, such training models require input data that can be obtained only through orthologous data sets (e.g., miCLIP) or manual curation from online databases (e.g., REPIC [93]). By contrast, the comparative approach compares signal data from two distinct data sets (e.g., treatment versus control) to identify sites at which signal information deviates. Signal-level analyses are considered to have far greater potential than error rate approaches, and this has resulted in a wider range of tools being developed, again producing a divide between (i) those that can theoretically detect any RNA modification that changes the ionic flux and (ii) those that are specific for individual modifications (e.g., m^6A).

Starting with the former category, Tombo supports the detection of both DNA and RNA modifications (92). Three specific methods are available depending on the available data sets: alternative, *de novo*, and sample comparison. For DRS data sets, the sample comparison approach is specifically recommended and provides two models for analysis (i) `model_sample_compare`, and (ii) `level_sample_compare`. These methods compare two sets of reads to identify differences at each reference position based on signal-level distributions by using

standard univariate statistical tests, i.e., the KS test, the U test, and the t test. While the original presentation of this tool focused on DNA modifications, Tombo has recently been used to look at m⁵C modifications in the transcriptomes of porcine reproductive and respiratory syndrome virus (94) and *Arabidopsis thaliana* (95). Nanocompore uses a comparative approach to make RNA modification calls by initially grouping reads belonging to each reference transcript and then assigning their corresponding median signal intensity and dwell time to each transcript position (96). The two conditions are compared on a positional basis using either a univariate pairwise test or a bivariate classification method based on a two-component (dwell time versus signal intensity) Gaussian mixture model (GMM), a form of supervised learning. GMM clustering is followed by a logistical regression test which determines the significance in difference of reads between conditions. Finally, P values are corrected for multiple testing. Synthetically modified oligonucleotides were used to benchmark the accuracy of Nanocompore for numerous different modifications (including m⁶A, inosine, ψ , m⁵C, and m¹G) before analysis of m⁶A in yeast and human RNAs. Conceptually similar to Nanocompore, xPore also uses a GMM to compare differences in signal intensity (but not dwell time) between modified and unmodified data sets (97). However, it also adds in prior information of a theoretical signal distribution of unmodified RNA to further guide the Gaussian model. Using this model, signal properties such as the mean and variance are used to estimate the fraction of modified reads and assign a modification probability. A two-step intermediate filtering step is included to reduce false positives, with the remaining sites undergoing a Z-test on differentially modified rates between samples. In their analysis of m⁶A modified sites, the authors focused on positions with a central adenine (NNANN), using DRS from wild-type (WT) and METTL3 knockout (KO) HEK293 T cells, and they obtained an AUC of 0.86. However, when the analysis was extended to all k-mers, the rate of false positives increased. nanoRMS utilizes signal intensity, dwell time, and base quality/probability information to perform either unsupervised or supervised clustering that assigns reads into modified and unmodified groups. When running a supervised clustering method, which is generally suited to low-stoichiometry contexts such as m⁶A on mRNA, comparative data sets (e.g., treatment versus control) are required. By contrast, unsupervised clustering is used to predict RNA modifications in a single sample alone. The limitation is that this is generally able to predict only specific RNA modifications in high-stoichiometry contexts, e.g., pseudouridine in rRNAs. In the original publication, nanoRMS processed comparative DRS data sets to predict pseudouridine modifications within yeast RNAs and, further, identified two new pseudouridines in yeast 15s mitochondrial rRNA (98). Finally, Yanocomp uses a comparative analysis approach to identify modifications by using a 5-mer sliding window to fit a multivariate GMM to separate data into modified and unmodified distributions (99). Yanocomp has been successfully applied to the identification of m⁶A modifications within the *A. thaliana* transcriptome, which showed a significant overlap with sites identified by the orthologous miCLIP method.

In the category of modification-specific tools, four are designed specifically for the detection of m⁶A while two are designed to identify pseudouridine. The common feature of modification-specific tools is that they use machine learning approaches to identify putatively modified sites. This strategy obviates the need for a comparative analysis (e.g., treatment versus control) and may be particularly useful for screening data sets and generating preliminary analyses. The first of these, MINES, specifically detects m⁶A modifications by harnessing Tombo's *de novo* algorithm by filtering putatively modified sites to keep only those located within a 30-nucleotide window centered within a DRACH motif (100). This approach reduces computational load but restricts the analysis to m⁶A modifications and indeed will not detect m⁶A installed outside of DRACH motifs (e.g., DRACG). The stoichiometric values obtained postfiltering were combined with publicly available miCLIP results to train a classifier. The prediction accuracy of the classifier ranged from 67 to 83% and was heavily reliant on DRACH sequence contexts. Nanom6A extracts features such as the median, mean, and standard deviation of signal for each 5-mer and uses these to train an ensemble learning model on RRACH motifs (101). To validate their model, they used Nanom6A to predict modified and unmodified reads within a synthesized mRNA data set with known m⁶A

and non-m⁶A ratios. Validation was extended in HEK293 cells by quantifying the loss of m⁶A following METTL3 silencing. In addition, they used Nanom6A across different biological systems with known locations of m⁶As and compared their results against those of different biological/computational approaches. Here, 81% of m⁶A-modified mRNAs predicted by Nanom6A were validated by meRIP-Seq. At single-base resolution, 49% of m⁶A predictions correlated with meRIP-Seq peaks. m6Anet uses a combination of signal- and sequence-level features of DRS data sets to predict a probability of modification at each individual site using a multiple instance learning framework (102). To train their model, the authors used the HCT116 cell line with matched m⁶A-cross-linking-exonuclease sequencing (m⁶ACE-Seq) (103)-obtained modification labels. This approach yielded an AUC of 0.83 across DRACH motifs in two different human cell lines. DENA uses signal features (mean, median, standard deviation, dwell time, and base quality) obtained from Tombo resquigging to train a neural network to make m⁶A modification calls (104). Training data were obtained by analysis of *A. thaliana* DRS data sets, in which DiffErr (99) was initially used to identify sites at RRACH motifs that contained a differential “error” within 5 nucleotides between the WT and KO. Electrical signals at these sites from either WT or KO reads were used as training labels, m⁶A or A, respectively. For the 12 possible RRACH motifs, the AUC ranged from 0.90 to 0.97 on the test data set, along with 90% accuracy by comparing these sites to m⁶A sites detected by miCLIP.

Penguin focuses on detection of pseudouridine and extracts 5-mers, mean, standard deviation, and length of signal. 5-mers are further filtered to retain only those containing a uracil at the center (105). A gold standard set of modified bases contained within HEK293 and HeLa cell lines were used as labels and were merged with the output from nanopolish. Three unsupervised models were trained and used to benchmark the two cell lines, achieving accuracies between 85 and 93%. Similar results were achieved when training was done specifically on HEK293 T cells and tested on an independent cell line (i.e., HeLa). NanoPsu uses 12 different base calling error features to identify pseudouridine by using supervised learning (106). To train the model, rRNAs isolated from human, yeast, *Caenorhabditis elegans*, *Drosophila melanogaster*, and human fecal bacteria were split into two subsets. For each sample, half underwent Illumina bisulfite sequencing to map pseudouridine sites (2,142 sites) which were used as training labels, while the other half underwent nanopore DRS sequencing to obtain training features. After filtering for pseudouridine sites that passed a read coverage of 20, 640 modified and 689 randomly selected unmodified U sites were used for training. This resulted in an AUC of 0.938 on the test data set, with downstream analysis on HeLa cells (in the presence/absence of interferon treatment) further validating their results.

(CURRENT) LIMITS IN THE DETECTION OF MODIFIED RIBONUCLEOTIDES BY NANOPORE SEQUENCING

Taken together, it is clear that a wealth of methodologies exist by which DRS data sets can be used to identify putatively modified nucleotides. The choice of error rate versus signal-level approaches offers researchers semiorthologous approaches with which to analyze the same data set. However, caution must be applied here, as a recent study demonstrated that both error rate and signal-level approaches have “blind spots” that may result in potential false negatives (i.e., the presence of a modified nucleotide reported by error rate but not signal-level approaches and vice versa) (98). Moreover, several of the above-mentioned tools (xPore, m6Anet, nanoRMS, NanoPsu) offer stoichiometry estimates at the level of individual isoforms, although the accuracy of these estimates remains questionable. In part, this is due to the lack of an available “ground truth” data set in which the location and abundance of all RNA modifications is known. Given the technical challenges and excessive cost associated with generating such a data set, it seems unlikely that this issue will be rectified soon. When considering which methodology/tool to use, it is worth noting that while powerful and generally easier to implement, error rate methodologies are limited to the level of RNA isoforms rather than individual RNAs and their power is greatest where read depths are very high, as sensitivity generally scales with depth of sequencing (107). Signal-level methodologies are, at present, also limited to the level of RNA isoforms but require significantly lower read

depths to make accurate predictions and retain the promise of allowing analysis of individual reads in the future. While none of these methods yet achieves the ultimate goal of identifying the presence and position of modified ribonucleotides within individual RNAs, their potential remains far greater than short-read sequencing approaches. Indeed, the primary difference remains that the experimental sides of short-read approaches (meRIP-Seq, miCLIP-Seq, etc.) are long and complex while the subsequent computational analysis of sequencing data is generally simple due to the existence of many well-established and robust pipelines. By contrast, DRS data sets can be generated in just a few hours with little technical expertise required. However, computational analyses are currently more complex due to a relative lack of robust, well-validated pipelines.

NANOPORE-MEDIATED DETECTION OF RNA MODIFICATIONS IN VIRAL CONTEXTS

At the time of writing, just six of the above-described tools had been applied to the analysis of viral DRS data sets: DRUMMER, Nanocompore, Eligos2, Tombo, MINES, and m6Anet. The first reported use of DRS to examine RNA modifications in viral contexts was published in 2019 by Viehweger and colleagues (108). Here, several subgenomic RNAs (sgRNAs) of the human alphacoronavirus HCoV-229E were reported by Tombo to contain 5-methylcytosine (m^5C) in both the conserved leader sequence (LS) and downstream open reading frames (ORFs). The validity of the reported signal has not, at the time of this writing, been confirmed by orthologous approaches (e.g., bisulfite sequencing). The subsequent emergence of the betacoronavirus SARS-CoV-2 in late 2019 led to a rash of studies aiming to identify and characterize the role of RNA modification in coronavirus biology (109–114). First among these was a delineation of the SARS-CoV-2 transcriptome architecture that included a description of 41 locations containing putatively modified ribonucleotides (109). These sites were identified by Tombo through a comparative analysis of native SARS-CoV-2 sgRNAs with 15 partially overlapping IVT fragments that spanned the entire SARS-CoV-2 gRNA. Sequence analysis identified the most common motif at these locations to be AAGAA-like, and these were particularly enriched at the 3' end of viral sgRNAs and more commonly on longer viral transcripts (gRNA, S, 3a, E, and M) than shorter ones (6, 7a, 7b, 8, and N). Of note, the authors also performed a comparative m^5C analysis and observed numerous signal-level changes using pretrained models. However, these changes were also observed in the unmodified (IVT) controls, suggesting that these are false positives. A second study soon followed in which a small number of sgRNAs from SARS-CoV-2 and the related human betacoronavirus HCoV-OC43 were shown to harbor m^6A modifications. The first of these utilized DRUMMER and Eligos2 to perform comparative analyses of DRS data sets derived from SARS-CoV-2 (A549^{+ACE2}) and HCoV-OC43 (MRC-5)-infected cells that were treated with either a specific METTL3 inhibitor or a noninhibiting control (110, 115). These data were validated using the orthogonal meRIP-Seq methodology (61, 110) and also served to show that inhibiting the enzymatic activity of METTL3 led to reduced titers of both SARS-CoV-2 and HCoV-OC43. The installation of m^6A on a small number of SARS-CoV-2 sgRNAs was further validated by additional studies (111, 112) using comparative DRS approaches allied with Tombo, MINES, and m6Anet. Another recent study utilized Tombo to show that the SARS-CoV-2 epitranscriptome remained stable during infection of Vero cells (113). While the identity of RNA modifications at the individual modified sites was not reported, the authors noted that at least two sites appeared unique to the gRNA and were not found on sgRNAs. The validity of these findings remains uncertain, however, due to a lack of orthologous confirmations. Finally, the presence of five pseudouridine modifications in the SARS-CoV-2 transcriptome was revealed by comparative analysis of SARS-CoV-2 DRS data sets and synthetic RNAs, two of which were confirmed through *in vitro* assays (114).

Beyond human coronaviruses, the use of nanopore DRS in detecting viral RNA modifications has thus far been limited, with just two studies published to date (54, 116). The first examined the role of m^6A in the life cycle of adenovirus serotype 5. Here, DRS data sets were derived from adenovirus-infected A549 cells in which METTL3 was present or knocked out (54). A comparative analysis of these data sets, which included two biological replicates per

condition, was performed using alpha versions of DRUMMER and Nanocompore and at both the exome and isoform levels. Both tools identified a large number of putatively modified sites in both early and late viral transcripts, with the majority (~80%) reported by both tools and located within DRACH motifs. Further analysis of the DRS data sets generated in the absence of METTL3 showed a significant reduction in the splicing efficiency of late viral RNAs. The second study examined the role of m⁶A in the life cycle of herpes simplex virus 1 (HSV-1) (116). Remarkably, HSV-1 induces relocalization of multiple host proteins associated with RNA modifications from the nucleus to the cytoplasm and, with this, a loss of RNA modifications, including m⁶A, from both host and viral mRNAs in a manner mediated by the viral ICP27 protein. The losses from human RNAs were determined by a comparative analysis of DRS data sets derived from HSV-1-infected fibroblasts and IVT RNAs representing the human gamma actin (ACTG1), while the loss of modified ribonucleotides from the viral RNAs was determined by a comparative analysis of infections of normal human dermal fibroblasts (NHDFs) with an HSV-1 wild type or an ICP27 null mutant. The comparative analysis was performed using an alpha version of DRUMMER, which reported that only around 25% of modified sites were located within or proximal to AC dinucleotides within DRACH motifs. It was thus inferred, but not validated, that the remaining sites represent other modified ribonucleotides.

EXPERIMENTAL CONSIDERATIONS WHEN USING NANOPORE SEQUENCING TO IDENTIFY RNA MODIFICATIONS

As is demonstrated above, the breadth and depth of computational approaches for detecting modified ribonucleotides in nanopore DRS data present virologists with a wide range of analytical options but also require a careful consideration of the underlying aims of individual experiments and the resulting data quality. Based on our own experiences with adenoviruses, herpesviruses, and betacoronaviruses, we have curated a list of the following parameters that we consider to be critical for consideration in the design and execution of nanopore-based RNA modification studies in viral contexts.

(i) Annotation quality. RNA modification detection tools require that DRS data sets are aligned against the transcriptome of the organism of interest. In viral contexts, this remains problematic, as very few viruses possess high-quality annotations. In the absence of such annotations, alignment of DRS reads often produces a subset of artefactual results in which the annotation for a given sequence read does not exist and leads to a subset of reads being assigned to an alternative transcript isoform instead. This typically hinders RNA modification analysis through the generation of false-positive and/or false-negative results. This can be remedied by either (i) generating a new high-quality transcriptome annotation or (ii) limiting the analysis to the exome level (i.e., aligning DRS data to the genome rather than the transcriptome). The former remains a complex, time-consuming, but ultimately valuable undertaking that can be partially accomplished using computational annotation tools such as StringTie2 (117) and FLAIR (118). The latter approach is simple and quick and can be accomplished with tools such as DRUMMER, Eligos, and Yanocomp, albeit at the expense of sensitivity and specificity (54).

(ii) Prefiltering of aligned read data. Regardless of whether pursuing exome- or isoform-level alignments, filtering of DRS read alignments prior to RNA modification detection can significantly improve sensitivity and accuracy. Standard measures should include screening for and removing adapter sequences from DRS reads, as well as removing secondary and supplementary alignments. This latter measure may also include rejecting individual reads with multiple primary alignments (i.e., they cannot be uniquely assigned to a single transcript isoform). A further step that is particularly relevant for isoform-level (i.e., transcriptome) analyses involves removing all primary read alignments that do not extend to within 50 nt of the defined 5' end of a given transcript.

(iii) Biological replicates. The use of at least two biological replicates per condition is recommended. Several tools have inbuilt support for such analyses (DRUMMER, Eligos, Nanocompore, xPore, EpiNano), while others require additional scripting to merge results from individual analyses. The principal benefit of including biological replicates is to enhance the identification of consistently modified sites while also providing measures of stochastic variation.

(iv) Manipulating RNA modification abundance. Comparative analyses rely on comparing conditions in which a given modified ribonucleotide is present at differing levels. In our experience with m⁶A, inhibition of METTL3 enzymatic activity or the generation of METTL3 conditional knockout cell lines significantly enhances sensitivity relative to RNA silencing strategies. Error rate analysis tools are particularly insensitive to RNA silencing strategies, presumably due to reduced stoichiometric differences between conditions. The use of IVT RNAs is another increasingly popular strategy, although this generally limits comparative approaches to reporting the presence/or absence, but not the identity, of all modifications within the RNA isoform being compared.

(v) The right tool for the right job. Choosing the right tool for a given analysis is a complex process. Read depth is a critical factor for several tools (DRUMMER, Eligos) (88, 107). Others favor specificity over sensitivity (EpiNano, m6Anet, MINES). Some (Nanocompore, xPore, m6Anet) even allow for estimates of stoichiometry, although these require careful orthogonal validation. Where possible, we recommend performing analyses with at least two or more distinct nanopore RNA modification tools and comparing outputs. This can help identify both false positives and false negatives that may arise from the intrinsic biases of a given tool. Due to the ease of installation, comprehensive documentation, and flexibility regarding how transcriptome alignments can be preprocessed, we have had particular success in viral contexts with DRUMMER, Nanocompore, and nanoRMS, although a comprehensive formal comparison of all tools has not been undertaken.

OUTLOOK

The field of nanopore sequencing is moving rapidly, pushed by constant biological and computational innovations. As noted above, native support for the identification of several DNA modifications was recently integrated into Guppy, and it is reasonable to expect similar support for some RNA modifications (e.g., m⁶A) in the coming years. This is driven by ongoing improvements in the DRS chemistry that will further reduce error rates and make modified ribonucleotides easier to distinguish from their unmodified counterparts. Whether such improvements will ultimately limit the utility of error rate approaches to RNA modification detection remains uncertain. We further anticipate that computational approaches for RNA modification detection will increasingly focus on the accurate measurement of stoichiometry, through both refinement of existing tools such as Nanocompore, xPore, and nanoRMS and the development and release of novel tools.

In looking further toward the future, we expect that comparative analysis approaches for profiling RNA modifications will eventually become defunct, replaced by a combination of native detection during base calling and the use of advanced signal-level algorithms capable of reporting multiple different types of RNA modifications within a single analysis (e.g., m⁶A, m¹A, pseudouridine, etc.). Such developments would also place within reach the ultimate goal of identifying the presence, identity, and specific location of modified ribonucleotides within individual RNA molecules.

To conclude in the context of viral RNA biology, the key limitation at present remains the general lack of well-annotated transcriptomes. Without these, viral RNA modification studies will be restricted to exome-level analyses that lack resolution. While resolving this will take time, it is notable that an increasing number of studies are addressing this issue across the viral spectrum (109, 119–123). Where high-quality transcriptome annotations already exist, it is clear that RNA modification detection using nanopore sequencing has provided significant advances that would simply not be possible using other approaches. This will only be enhanced further in the future.

ACKNOWLEDGMENTS

We thank Beate Sodeik, Abel Viejo-Borbolla, Jens Bohne, and Carolin Vogt for feedback and critical reading of the manuscript.

REFERENCES

1. Davis FF, Allen FW. 1957. Ribonucleic acids from yeast which contain a fifth nucleotide. *J Biol Chem* 227:907–915. [https://doi.org/10.1016/S0021-9258\(18\)70770-9](https://doi.org/10.1016/S0021-9258(18)70770-9).
2. Boccaletto P, Stefaniak F, Ray A, Cappannini A, Mukherjee S, Purta E, Kurkowska M, Shirvanizadeh N, Destefanis E, Groza P, Avşar G, Romitelli A, Pir P, Dassi E, Conticello SG, Aguiló F, Bujnicki JM. 2022. MODOMICS: a database

- of RNA modification pathways. 2021 update. *Nucleic Acids Res* 50:D231–D235. <https://doi.org/10.1093/nar/gkab1083>.
3. Linder B, Grozhnik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. 2015. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods* 12:767–772. <https://doi.org/10.1038/nmeth.3453>.
 4. Bringmann P, Lührmann R. 1987. Antibodies specific for N6-methyladenosine react with intact snRNPs U2 and U4/U6. *FEBS Lett* 213:309–315. [https://doi.org/10.1016/0014-5793\(87\)81512-0](https://doi.org/10.1016/0014-5793(87)81512-0).
 5. Saneyoshi M, Harada F, Nishimura S. 1969. Isolation and characterization of N6-methyladenosine from *Escherichia coli* valine transfer RNA. *Biochim Biophys Acta* 190:264–273. [https://doi.org/10.1016/0005-2787\(69\)90078-1](https://doi.org/10.1016/0005-2787(69)90078-1).
 6. Liu N, Parisien M, Dai Q, Zheng G, He C, Pan T. 2013. Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA* 19:1848–1856. <https://doi.org/10.1261/ma.041178.113>.
 7. Alarcón CR, Lee H, Goodarzi H, Halberg N, Tavazoie SF. 2015. N6-methyladenosine marks primary microRNAs for processing. *Nature* 519:482–485. <https://doi.org/10.1038/nature14281>.
 8. Piekna-Przybylska D, Decatur WA, Fournier MJ. 2008. The 3D rRNA modification maps database: with interactive tools for ribosome analysis. *Nucleic Acids Res* 36:D178–183. <https://doi.org/10.1093/nar/gkm855>.
 9. Sergiev PV, Golovina AY, Osterman IA, Nesterchuk MV, Sergeeva OV, Chugunova AA, Evfratov SA, Andreianova ES, Pletnev PI, Laptev IG, Petriukov KS, Navalayeu TI, Koteliansky VE, Bogdanov AA, Dontsova OA. 2016. N6-methylated adenosine in RNA: from bacteria to humans. *J Mol Biol* 428:2134–2145. <https://doi.org/10.1016/j.jmb.2015.12.013>.
 10. Li S, Mason CE. 2014. The pivotal regulatory landscape of RNA modifications. *Annu Rev Genomics Hum Genet* 15:127–150. <https://doi.org/10.1146/annurev-genom-090413-025405>.
 11. Roundtree IA, Evans ME, Pan T, He C. 2017. Dynamic RNA modifications in gene expression regulation. *Cell* 169:1187–1200. <https://doi.org/10.1016/j.cell.2017.05.045>.
 12. Desrosiers R, Friderici K, Rottman F. 1974. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc Natl Acad Sci U S A* 71:3971–3975. <https://doi.org/10.1073/pnas.71.10.3971>.
 13. Rottman F, Shatkin AJ, Perry RP. 1974. Sequences containing methylated nucleotides at the 5' termini of messenger RNAs: possible implications for processing. *Cell* 3:197–199. [https://doi.org/10.1016/0092-8674\(74\)90131-7](https://doi.org/10.1016/0092-8674(74)90131-7).
 14. Fazi F, Fatica A. 2019. Interplay between N6-methyladenosine (m6A) and non-coding RNAs in cell development and cancer. *Front Cell Dev Biol* 7:116. <https://doi.org/10.3389/fcell.2019.00116>.
 15. Maden BE. 1986. Identification of the locations of the methyl groups in 18 S ribosomal RNA from *Xenopus laevis* and man. *J Mol Biol* 189:681–699. [https://doi.org/10.1016/0022-2836\(86\)90498-5](https://doi.org/10.1016/0022-2836(86)90498-5).
 16. Maden BE. 1988. Locations of methyl groups in 28 S rRNA of *Xenopus laevis* and man. Clustering in the conserved core of molecule. *J Mol Biol* 201:289–314. [https://doi.org/10.1016/0022-2836\(88\)90139-8](https://doi.org/10.1016/0022-2836(88)90139-8).
 17. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, Jia G, Yu M, Lu Z, Deng X, Dai Q, Chen W, He C. 2014. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol* 10:93–95. <https://doi.org/10.1038/nchembio.1432>.
 18. Schwartz S, Mumbach MR, Jovanovic M, Wang T, Maciag K, Bushkin GG, Mertins P, Ter-Ovanesyan D, Habib N, Cacchiarelli D, Sanjana NE, Freinkman E, Pacold ME, Satija R, Mikkelsen TS, Hacohen N, Zhang F, Carr SA, Lander ES, Regev A. 2014. Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep* 8:284–296. <https://doi.org/10.1016/j.celrep.2014.05.048>.
 19. Ke S, Pandya-Jones A, Saito Y, Fak JJ, Vågbo CB, Geula S, Hanna JH, Black DL, Darnell JE, Darnell RB. 2017. m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev* 31:990–1006. <https://doi.org/10.1101/gad.301036.117>.
 20. Hazra D, Chapat C, Graille M. 2019. m⁶A mRNA destiny: chained to the rYTHm by the YTH-containing proteins. *Genes (Basel)* 10:49. <https://doi.org/10.3390/genes10010049>.
 21. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. 2015. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* 518:560–564. <https://doi.org/10.1038/nature14234>.
 22. Zhou KI, Shi H, Lyu R, Wylder AC, Matuszek Z, Pan JN, He C, Parisien M, Pan T. 2019. Regulation of co-transcriptional pre-mRNA splicing by m6A through the low-complexity protein hnRNPG. *Mol Cell* 76:70–81.e9. <https://doi.org/10.1016/j.molcel.2019.07.005>.
 23. Huang H, Weng H, Sun W, Qin X, Shi H, Wu H, Zhao BS, Mesquita A, Liu C, Yuan CL, Hu Y-C, Hüttelmaier S, Skibbe JR, Su R, Deng X, Dong L, Sun M, Li C, Nachtergaele S, Wang Y, Hu C, Ferchen K, Greis KD, Jiang X, Wei M, Qu L, Guan J-L, He C, Yang J, Chen J. 2018. Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat Cell Biol* 20:285–295. <https://doi.org/10.1038/s41556-018-0045-z>.
 24. Bartosovic M, Molares HC, Gregorova P, Hrossova D, Kudla G, Vanacova S. 2017. N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3'-end processing. *Nucleic Acids Res* 45:11356–11370. <https://doi.org/10.1093/nar/gkx778>.
 25. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang C-M, Li CJ, Vågbo CB, Shi Y, Wang W-L, Song S-H, Lu Z, Bosmans RGP, Dai Q, Hao Y-J, Yang X, Zhao W-M, Tong W-M, Wang X-J, Bogdan F, Furu K, Fu Y, Jia G, Zhao X, Liu J, Krokan HE, Klungland A, Yang Y-G, He C. 2013. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell* 49:18–29. <https://doi.org/10.1016/j.molcel.2012.10.015>.
 26. Mauer J, Luo X, Blanjoie A, Jiao X, Grozhnik AV, Patil DP, Linder B, Pickering BF, Vasseur J-J, Chen Q, Gross SS, Elemento O, DeBart F, Kiledjian M, Jaffrey SR. 2017. Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature* 541:371–375. <https://doi.org/10.1038/nature21022>.
 27. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M, Sorek R, Rechavi G. 2012. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485:201–206. <https://doi.org/10.1038/nature11112>.
 28. Roundtree IA, Luo G-Z, Zhang Z, Wang X, Zhou T, Cui Y, Sha J, Huang X, Guerrero L, Xie P, He E, Shen B, He C. 2017. YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *Elife* 6:e31311. <https://doi.org/10.7554/eLife.31311>.
 29. Xiao W, Adhikari S, Dahal U, Chen Y-S, Hao Y-J, Sun B-F, Sun H-Y, Li A, Ping X-L, Lai W-Y, Wang X, Ma H-L, Huang C-M, Yang Y, Huang N, Jiang G-B, Wang H-L, Zhou Q, Wang X-J, Zhao Y-L, Yang Y-G. 2016. Nuclear m(6A) reader YTHDC1 regulates mRNA splicing. *Mol Cell* 61:507–519. <https://doi.org/10.1016/j.molcel.2016.01.012>.
 30. Fu Y, Dominissini D, Rechavi G, He C. 2014. Gene expression regulation mediated through reversible m⁶A RNA methylation. *Nat Rev Genet* 15:293–306. <https://doi.org/10.1038/nrg3724>.
 31. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, Weng X, Chen K, Shi H, He C. 2015. N(6)-methyladenosine modulates messenger RNA translation efficiency. *Cell* 161:1388–1399. <https://doi.org/10.1016/j.cell.2015.05.014>.
 32. Lin S, Choe J, Du P, Triboulet R, Gregory RI. 2016. The m(6A) methyltransferase METTL3 promotes translation in human cancer cells. *Mol Cell* 62:335–345. <https://doi.org/10.1016/j.molcel.2016.03.021>.
 33. Sun H, Zhang M, Li K, Bai D, Yi C. 2019. Cap-specific, terminal N6-methylation by a mammalian m6Am methyltransferase. *Cell Res* 29:80–82. <https://doi.org/10.1038/s41422-018-0117-4>.
 34. Boulias K, Toczyłowska-Socha D, Hawley BR, Liberman N, Takashima K, Zaccara S, Guez T, Vasseur J-J, DeBart F, Aravind L, Jaffrey SR, Greer EL. 2019. Identification of the m6Am methyltransferase PCIF1 reveals the location and functions of m6Am in the transcriptome. *Mol Cell* 75:631–643.e8. <https://doi.org/10.1016/j.molcel.2019.06.006>.
 35. Sendinc E, Valle-Garcia D, Dhali A, Chen H, Henriques T, Navarrete-Perea J, Sheng W, Gygi SP, Adelman K, Shi Y. 2019. PCIF1 catalyzes m6Am mRNA methylation to regulate gene expression. *Mol Cell* 75:620–630.e9. <https://doi.org/10.1016/j.molcel.2019.05.030>.
 36. Akichika S, Hirano S, Shichino Y, Suzuki T, Nishimasa H, Ishitani R, Sugita A, Hirose Y, Iwasaki S, Nureki O, Suzuki T. 2019. Cap-specific terminal N6-methylation of RNA by an RNA polymerase II-associated methyltransferase. *Science* 363:eaav0080. <https://doi.org/10.1126/science.aav0080>.
 37. Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, Suter CM, Preiss T. 2012. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res* 40:5023–5033. <https://doi.org/10.1093/nar/gks144>.
 38. Wyatt GF. 1950. Occurrence of 5-methylcytosine in nucleic acids. *Nature* 166:237–238. <https://doi.org/10.1038/166237b0>.
 39. Dunn DB. 1961. The occurrence of 1-methyladenine in ribonucleic acid. *Biochim Biophys Acta* 46:198–200. [https://doi.org/10.1016/0006-3002\(61\)90668-0](https://doi.org/10.1016/0006-3002(61)90668-0).
 40. Seo KW, Kleiner RE. 2020. YTHDF2 recognition of N1-methyladenosine (m1A)-modified RNA is associated with transcript destabilization. *ACS Chem Biol* 15:132–139. <https://doi.org/10.1021/acscchembio.9b00655>.
 41. Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, Dai Q, Di Segni A, Salmon-Divon M, Clark WG, Zheng G, Pan T, Solomon O, Eyal E, Hershkovitz V, Han D, Doré LC, Amariglio N, Rechavi G, He C. 2016. The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 530:441–446. <https://doi.org/10.1038/nature16998>.

42. Shatkin AJ. 1976. Capping of eucaryotic mRNAs. *Cell* 9:645–653. [https://doi.org/10.1016/0092-8674\(76\)90128-8](https://doi.org/10.1016/0092-8674(76)90128-8).
43. Banerjee AK. 1980. 5'-Terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol Rev* 44:175–205. <https://doi.org/10.1128/mr.44.2.175-205.1980>.
44. Zhang L-S, Liu C, Ma H, Dai Q, Sun H-L, Luo G, Zhang Z, Zhang L, Hu L, Dong X, He C. 2019. Transcriptome-wide mapping of internal N7-methylguanosine methylome in mammalian mRNA. *Mol Cell* 74:1304–1316.e8. <https://doi.org/10.1016/j.molcel.2019.03.036>.
45. Sikorski PJ, Warminski M, Kubacka D, Ratajczak T, Nowis D, Kowalska J, Jemielity J. 2020. The identity and methylation status of the first transcribed nucleotide in eukaryotic mRNA 5' cap modulates protein expression in living cells. *Nucleic Acids Res* 48:1607–1626. <https://doi.org/10.1093/nar/gkaa032>.
46. Wang J, Alvin Chew BL, Lai Y, Dong H, Xu L, Balamkundu S, Cai WM, Cui L, Liu CF, Fu X-Y, Lin Z, Shi P-Y, Lu TK, Luo D, Jaffrey SR, Dedon PC. 2019. Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. *Nucleic Acids Res* 47:e130. <https://doi.org/10.1093/nar/gkz751>.
47. Dai Q, Moshitch-Moshkovitz S, Han D, Kol N, Amariglio N, Rechavi G, Dominissini D, He C. 2017. Nm-seq maps 2'-O-methylation sites in human mRNA with base precision. *Nat Methods* 14:695–698. <https://doi.org/10.1038/nmeth.4294>.
48. Gokhale NS, Horner SM. 2017. RNA modifications go viral. *PLoS Pathog* 13:e1006188. <https://doi.org/10.1371/journal.ppat.1006188>.
49. Tirumuru N, Wu L. 2019. HIV-1 envelope proteins up-regulate N6-methyladenosine levels of cellular RNA independently of viral replication. *J Biol Chem* 294:3249–3260. <https://doi.org/10.1074/jbc.RA118.005608>.
50. Jurczynski D, Zhang W, Terry SN, Kehrer T, Bermúdez González MC, McGregor E, Mulder LCF, Eckwahl MJ, Pan T, Simon V. 2020. HIV protease cleaves the antiviral m6A reader protein YTHDF3 in the viral particle. *PLoS Pathog* 16:e1008305. <https://doi.org/10.1371/journal.ppat.1008305>.
51. Kennedy EM, Bogerd HP, Kornepati AVR, Kang D, Ghoshal D, Marshall JB, Poling BC, Tsai K, Gokhale NS, Horner SM, Cullen BR. 2016. Posttranscriptional m(6)A editing of HIV-1 mRNAs enhances viral gene expression. *Cell Host Microbe* 19:675–685. <https://doi.org/10.1016/j.chom.2016.04.002>.
52. Finkel D, Groner Y. 1983. Methylations of adenosine residues (m6A) in pre-mRNA are important for formation of late simian virus 40 mRNAs. *Virology* 131:409–425. [https://doi.org/10.1016/0042-6822\(83\)90508-1](https://doi.org/10.1016/0042-6822(83)90508-1).
53. Tsai K, Courtney DG, Cullen BR. 2018. Addition of m6A to SV40 late mRNAs enhances viral structural gene expression and replication. *PLoS Pathog* 14:e1006919. <https://doi.org/10.1371/journal.ppat.1006919>.
54. Price AM, Hayer KE, McIntyre ABR, Gokhale NS, Abebe JS, Della Fera AN, Mason CE, Horner SM, Wilson AC, Depledge DP, Weitzman MD. 2020. Direct RNA sequencing reveals m6A modifications on adenovirus RNA are necessary for efficient splicing. *Nat Commun* 11:6016. <https://doi.org/10.1038/s41467-020-19787-6>.
55. Rubio RM, Depledge DP, Bianco C, Thompson L, Mohr I. 2018. RNA m6A modification enzymes shape innate responses to DNA by regulating interferon β . *Genes Dev* 32:1472–1484. <https://doi.org/10.1101/gad.319475.118>.
56. Winkler R, Gillis E, Lasman L, Safra M, Geula S, Soyris C, Nachshon A, Taischmiedel J, Friedman N, Le-Trilling VTK, Trilling M, Mandelboim M, Hanna JH, Schwartz S, Stern-Ginossar N. 2019. m6A modification controls the innate immune response to infection by targeting type I interferons. *Nat Immunol* 20:173–182. <https://doi.org/10.1038/s41590-018-0275-z>.
57. Tan B, Liu H, Zhang S, da Silva SR, Zhang L, Meng J, Cui X, Yuan H, Sorel O, Zhang S-W, Huang Y, Gao S-J. 2018. Viral and cellular N6-methyladenosine and N6,2'-O-dimethyladenosine epitranscriptomes in the KSHV life cycle. *Nat Microbiol* 3:108–120. <https://doi.org/10.1038/s41564-017-0056-8>.
58. Hesser CR, Karijolic J, Dominissini D, He C, Glaunsinger BA. 2018. N6-methyladenosine modification and the YTHDF2 reader protein play cell type specific roles in lytic viral gene expression during Kaposi's sarcoma-associated herpesvirus infection. *PLoS Pathog* 14:e1006995. <https://doi.org/10.1371/journal.ppat.1006995>.
59. Ye F, Chen ER, Nilsen TW. 2017. Kaposi's sarcoma-associated herpesvirus utilizes and manipulates RNA N6-adenosine methylation to promote lytic replication. *J Virol* 91:e00466-17. <https://doi.org/10.1128/JVI.00466-17>.
60. Baquero-Perez B, Antanaviciute A, Yonchev ID, Carr IM, Wilson SA, Whitehouse A. 2019. The Tudor SND1 protein is an m6A RNA reader essential for replication of Kaposi's sarcoma-associated herpesvirus. *Elife* 8:e47261. <https://doi.org/10.7554/eLife.47261>.
61. Liu J, Xu Y-P, Li K, Ye Q, Zhou H-Y, Sun H, Li X, Yu L, Deng Y-Q, Li R-T, Cheng M-L, He B, Zhou J, Li X-F, Wu A, Yi C, Qin C-F. 2021. The m6A methylome of SARS-CoV-2 in host cells. *Cell Res* 31:404–414. <https://doi.org/10.1038/s41422-020-00465-7>.
62. Li N, Hui H, Bray B, Gonzalez GM, Zeller M, Anderson KG, Knight R, Smith D, Wang Y, Carlin AF, Rana TM. 2021. METTL3 regulates viral m6A RNA modification and host cell innate immune responses during SARS-CoV-2 infection. *Cell Rep* 35:109091. <https://doi.org/10.1016/j.celrep.2021.109091>.
63. Hao H, Hao S, Chen H, Chen Z, Zhang Y, Wang J, Wang H, Zhang B, Qiu J, Deng F, Guan W. 2019. N6-methyladenosine modification and METTL3 modulate enterovirus 71 replication. *Nucleic Acids Res* 47:362–374. <https://doi.org/10.1093/nar/gky1007>.
64. Courtney DG. 2021. Post-transcriptional regulation of viral RNA through epitranscriptional modification. *Cells* 10:1129. <https://doi.org/10.3390/cells10051129>.
65. Baquero-Perez B, Geers D, Diez J. 2021. From A to m6A: the emerging viral epitranscriptome. *Viruses* 13:1049. <https://doi.org/10.3390/v13061049>.
66. McIntyre W, Netzband R, Bonenfant G, Biegel JM, Miller C, Fuchs G, Henderson E, Arra M, Canki M, Fabris D, Pager CT. 2018. Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection. *Nucleic Acids Res* 46:5776–5791. <https://doi.org/10.1093/nar/gky029>.
67. Zhu W, Wang J-Z, Xu Z, Cao M, Hu Q, Pan C, Guo M, Wei J-F, Yang H. 2019. Detection of N6-methyladenosine modification residues (Review). *Int J Mol Med* 43:2267–2278. <https://doi.org/10.3892/ijmm.2019.4169>.
68. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. 2012. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149:1635–1646. <https://doi.org/10.1016/j.cell.2012.05.003>.
69. McIntyre ABR, Gokhale NS, Cerchiotti L, Jaffrey SR, Horner SM, Mason CE. 2020. Limits in the detection of m6A changes using MeRIP/m6A-seq. *Sci Rep* 10:6590. <https://doi.org/10.1038/s41598-020-63355-3>.
70. Meyer KD. 2019. DART-seq: an antibody-free method for global m6A detection. *Nat Methods* 16:1275–1280. <https://doi.org/10.1038/s41592-019-0570-0>.
71. Garcia-Campos MA, Edelheit S, Toth U, Safra M, Shachar R, Viukov S, Winkler R, Nir R, Lasman L, Brandis A, Hanna JH, Rossmanith W, Schwartz S. 2019. Deciphering the “m6A code” via antibody-independent quantitative profiling. *Cell* 178:731–747.e16. <https://doi.org/10.1016/j.cell.2019.06.013>.
72. Shu X, Cao J, Cheng M, Xiang S, Gao M, Li T, Ying X, Wang F, Yue Y, Lu Z, Dai Q, Cui X, Ma L, Wang Y, He C, Feng X, Liu J. 2020. A metabolic labeling method detects m6A transcriptome-wide at single base resolution. *Nat Chem Biol* 16:887–895. <https://doi.org/10.1038/s41589-020-0526-9>.
73. Wang Y, Xiao Y, Dong S, Yu Q, Jia G. 2020. Antibody-free enzyme-assisted chemical approach for detection of N6-methyladenosine. *Nat Chem Biol* 16:896–903. <https://doi.org/10.1038/s41589-020-0525-x>.
74. Hu L, Liu S, Peng Y, Ge R, Su R, Seneviratne C, Harada BT, Dai Q, Wei J, Zhang L, Hao Z, Luo L, Wang H, Wang Y, Luo M, Chen M, Chen J, He C. 14 March 2022. m6A RNA modifications are measured at single-base resolution across the mammalian transcriptome. *Nat Biotechnol* <https://doi.org/10.1038/s41587-022-01243-z>.
75. Kukurba KR, Montgomery SB. 2015. RNA sequencing and analysis. *Cold Spring Harb Protoc* 2015:951–969. <https://doi.org/10.1101/pdb.top084970>.
76. Depledge DP, Mohr I, Wilson AC. 2019. Going the distance: optimizing RNA-Seq strategies for transcriptomic analysis of complex viral genomes. *J Virol* 93:e01342-18. <https://doi.org/10.1128/JVI.01342-18>.
77. Galarde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, Pantic N, Admassu T, James P, Warland A, Jordan M, Ciccone J, Serra S, Keenan J, Martin S, McNeill L, Wallace EJ, Jayasinghe L, Wright C, Blasco J, Young S, Brocklebank D, Juul S, Clarke J, Heron AJ, Turner DJ. 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods* 15:201–206. <https://doi.org/10.1038/nmeth.4577>.
78. Keller MW, Rambo-Martin BL, Wilson MM, Ridenour CA, Shepard SS, Stark TJ, Neuhaus EB, Dugan VG, Wentworth DE, Barnes JR. 2018. Direct RNA sequencing of the coding complete influenza A virus genome. *Sci Rep* 8:14408. <https://doi.org/10.1038/s41598-018-32615-8>.
79. Drexler HL, Choquet K, Churchman LS. 2020. Splicing kinetics and coordination revealed by direct nascent RNA sequencing through nanopores. *Mol Cell* 77:985–998.e8. <https://doi.org/10.1016/j.molcel.2019.11.017>.
80. Greener JG, Kandathil SM, Moffat L, Jones DT. 2022. A guide to machine learning for biologists. *Nat Rev Mol Cell Biol* 23:40–16. <https://doi.org/10.1038/s41580-021-00407-0>.
81. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow P-M, Zietz M, Hoffman MM, Xie W, Rosen GL, Lengerich BJ, Israeli J, Lanchantin J, Woloszynek S, Carpenter AE, Shrikumar A, Xu J, Cofer EM, Lavender CA, Turaga SC, Alexandari AM, Lu Z, Harris DJ, DeCaprio D, Qi Y, Kundaje A, Peng Y, Wiley LK, Segler MHS, Boca SM, Swamidass SJ, Huang A, Gitter A, Greene CS. 2018. Opportunities and obstacles for deep learning in

- biology and medicine. *J R Soc Interface* 15:20170387. <https://doi.org/10.1098/rsif.2017.0387>.
82. Wick RR, Judd LM, Holt KE. 2019. Performance of neural network base-calling tools for Oxford Nanopore sequencing. *Genome Biol* 20:129. <https://doi.org/10.1186/s13059-019-1727-y>.
 83. Soneson C, Yao Y, Bratus-Neuenschwander A, Patrignani A, Robinson MD, Hussain S. 2019. A comprehensive examination of Nanopore native RNA sequencing for characterization of complex transcriptomes. *Nat Commun* 10:3359. <https://doi.org/10.1038/s41467-019-11272-z>.
 84. Delahaye C, Nicolas J. 2021. Sequencing DNA with nanopores: troubles and biases. *PLoS One* 16:e0257521. <https://doi.org/10.1371/journal.pone.0257521>.
 85. Furlan M, Delgado-Tejedor A, Mulrone L, Pelizzola M, Novoa EM, Leonardi T. 2021. Computational methods for RNA modification detection from nanopore direct RNA sequencing data. *RNA Biol* 18:31–40. <https://doi.org/10.1080/15476286.2021.1978215>.
 86. Liu H, Begik O, Novoa EM. 2021. EpiNano: detection of m6A RNA modifications using Oxford Nanopore direct RNA sequencing. *Methods Mol Biol* 2298:31–52. https://doi.org/10.1007/978-1-0716-1374-0_3.
 87. Wang Y, Wang H, Xi F, Wang H, Han X, Wei W, Zhang H, Zhang Q, Zheng Y, Zhu Q, Kohnen MV, Reddy ASN, Gu L. 2020. Profiling of circular RNA N6-methyladenosine in moso bamboo (*Phyllostachys edulis*) using nanopore-based direct RNA sequencing. *J Integr Plant Biol* 62:1823–1838. <https://doi.org/10.1111/jipb.13002>.
 88. Jenjaroenpun P, Wongsurawat T, Wadley TD, Wassenaar TM, Liu J, Dai Q, Wanchai V, Akel NS, Jamshidi-Parsian A, Franco AT, Boysen G, Jennings ML, Ussery DW, He C, Nookaew I. 2021. Decoding the epitranscriptional landscape from native RNA sequences. *Nucleic Acids Res* 49:e7. <https://doi.org/10.1093/nar/gkaa620>.
 89. Parker MT, Knop K, Sherwood AV, Schurch NJ, Mackinnon K, Gould PD, Hall AJ, Barton GJ, Simpson GG. 2020. Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m6A modification. *Elife* 9:e49658. <https://doi.org/10.7554/eLife.49658>.
 90. Piechotta M, Wang Q, Altmüller J, Dieterich C. 2021. RNA modification mapping with JACUSA2. *bioRxiv* <https://doi.org/10.1101/2021.07.02.450888>.
 91. Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods* 12:733–735. <https://doi.org/10.1038/nmeth.3444>.
 92. Stoiber M, Quick J, Egan R, Lee JE, Celniker S, Neely RK, Loman N, Pennacchio LA, Brown J. 2017. De novo identification of DNA modifications enabled by genome-guided nanopore signal processing. *bioRxiv* <https://doi.org/10.1101/094672>.
 93. Liu S, Zhu A, He C, Chen M. 2020. REPIC: a database for exploring the N6-methyladenosine methylome. *Genome Biol* 21:100. <https://doi.org/10.1186/s13059-020-02012-4>.
 94. Zhang R, Wang P, Ma X, Wu Y, Luo C, Qiu L, Zeshan B, Yang Z, Zhou Y, Wang X. 2021. Nanopore-based direct RNA-sequencing reveals a high-resolution transcriptional landscape of porcine reproductive and respiratory syndrome virus. *Viruses* 13:2531. <https://doi.org/10.3390/v13122531>.
 95. Zhang S, Li R, Zhang L, Chen S, Xie M, Yang L, Xia Y, Foyer CH, Zhao Z, Lam H-M. 2020. New insights into Arabidopsis transcriptome complexity revealed by direct sequencing of native RNAs. *Nucleic Acids Res* 48:7700–7711. <https://doi.org/10.1093/nar/gkaa588>.
 96. Leger A, Amaral PP, Pandolfini L, Capitanchik C, Capraro F, Miano V, Migliori V, Toolan-Kerr P, Sideri T, Enright AJ, Tzelepis K, van Werven FJ, Luscombe NM, Barbieri I, Ule J, Fitzgerald T, Birney E, Leonardi T, Kouzarides T. 2021. RNA modifications detection by comparative Nanopore direct RNA sequencing. *Nat Commun* 12:7198. <https://doi.org/10.1038/s41467-021-27393-3>.
 97. Pratanwanich PN, Yao F, Chen Y, Koh CWQ, Wan YK, Hendra C, Poon P, Goh YT, Yap PML, Chooi JY, Chng WJ, Ng SB, Thiery A, Goh WSS, Göke J. 2021. Identification of differential RNA modifications from nanopore direct RNA sequencing with xPore. *Nat Biotechnol* 39:1394–1402. <https://doi.org/10.1038/s41587-021-00949-w>.
 98. Begik O, Lucas MC, Prysycz LP, Ramirez JM, Medina R, Milenkovic I, Cruciani S, Liu H, Vieira HGS, Sas-Chen A, Mattick JS, Schwartz S, Novoa EM. 2021. Quantitative profiling of pseudouridylation dynamics in native RNAs with nanopore sequencing. *Nat Biotechnol* 39:1278–1291. <https://doi.org/10.1038/s41587-021-00915-6>.
 99. Parker MT, Barton GJ, Simpson GG. 2021. Yanocomp: robust prediction of m6A modifications in individual nanopore direct RNA reads. *bioRxiv* <https://doi.org/10.1101/2021.06.15.448494>.
 100. Lorenz DA, Sathe S, Einstein JM, Yeo GW. 2020. Direct RNA sequencing enables m6A detection in endogenous transcript isoforms at base-specific resolution. *RNA* 26:19–28. <https://doi.org/10.1261/rna.072785.119>.
 101. Gao Y, Liu X, Wu B, Wang H, Xi F, Kohnen MV, Reddy ASN, Gu L. 2021. Quantitative profiling of N6-methyladenosine at single-base resolution in stem-differentiating xylem of *Populus trichocarpa* using Nanopore direct RNA sequencing. *Genome Biol* 22:22. <https://doi.org/10.1186/s13059-020-02241-7>.
 102. Hendra C, Pratanwanich PN, Wan YK, Goh WSS, Thiery A, Göke J. 2021. Detection of m6A from direct RNA sequencing using a multiple instance learning framework. *bioRxiv* <https://doi.org/10.1101/2021.09.20.461055>.
 103. Koh CWQ, Goh YT, Goh WSS. 2019. Atlas of quantitative single-base-resolution N6-methyl-adenine methylomes. *Nat Commun* 10:5636. <https://doi.org/10.1038/s41467-019-13561-z>.
 104. Qin H, Ou L, Gao J, Chen L, Wang J-W, Hao P, Li X. 2022. DENA: training an authentic neural network model using Nanopore sequencing data of Arabidopsis transcripts for detection and quantification of N6-methyladenosine on RNA. *Genome Biol* 23:25. <https://doi.org/10.1186/s13059-021-02598-3>.
 105. Hassan D, Acevedo D, Daulatabad SV, Mir Q, Janga SC. 2021. Penguin: a tool for predicting pseudouridine sites in direct RNA nanopore sequencing data. *bioRxiv* <https://doi.org/10.1101/2021.03.31.437901>.
 106. Huang S, Zhang W, Katanski CD, Dersh D, Dai Q, Lolans K, Yewdell J, Eren AM, Pan T. 2021. Interferon inducible pseudouridine modification in human mRNA by quantitative nanopore profiling. *Genome Biol* 22:330. <https://doi.org/10.1186/s13059-021-02557-y>.
 107. Abebe JS, Price AM, Hayer KE, Mohr I, Weitzman MD, Wilson AC, Depledge DP. 2022. DRUMMER—rapid detection of RNA modifications through comparative nanopore sequencing. *Bioinformatics* 2022:btac274. <https://doi.org/10.1093/bioinformatics/btac274>.
 108. Viehweger A, Krautwurst S, Lamkiewicz K, Madhugiri R, Ziebuhr J, Hölzer M, Marz M. 2019. Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. *Genome Res* 29:1545–1554. <https://doi.org/10.1101/gr.247064.118>.
 109. Kim D, Lee J-Y, Yang J-S, Kim JW, Kim VN, Chang H. 2020. The architecture of SARS-CoV-2 transcriptome. *Cell* 181:914–921.e10. <https://doi.org/10.1016/j.cell.2020.04.011>.
 110. Burgess HM, Depledge DP, Thompson L, Srinivas KP, Grande RC, Vink EI, Abebe JS, Blackaby WP, Hendrick A, Albertella MR, Kouzarides T, Stapleford KA, Wilson AC, Mohr I. 2021. Targeting the m6A RNA modification pathway blocks SARS-CoV-2 and HCoV-OC43 replication. *Genes Dev* 35:1005–1019. <https://doi.org/10.1101/gad.348320.121>.
 111. Campos JHC, Maricato JT, Braconi CT, Antoneli F, Janini LMR, Briones MRS. 2021. Direct RNA sequencing reveals SARS-CoV-2 m6A sites and possible differential DRACH motif methylation among variants. *Viruses* 13:2108. <https://doi.org/10.3390/v13112108>.
 112. Zhang X, Hao H, Ma L, Zhang Y, Hu X, Chen Z, Liu D, Yuan J, Hu Z, Guan W. 2021. Methyltransferase-like 3 modulates severe acute respiratory syndrome coronavirus-2 RNA N6-methyladenosine modification and replication. *mBio* 12:e01067-21. <https://doi.org/10.1128/mBio.01067-21>.
 113. Chang JJ-Y, Rawlinson D, Pitt ME, Tairaoa G, Gleeson J, Zhou C, Mordant FL, De Paoli-Iseppi R, Cally L, Purcell DFJ, Stinear TP, Londrigan SL, Clark MB, Williamson DA, Subbarao K, Coin LJM. 2021. Transcriptional and epitranscriptional dynamics of SARS-CoV-2 during cellular infection. *Cell Rep* 35:109108. <https://doi.org/10.1016/j.celrep.2021.109108>.
 114. Fleming AM, Mathewson NJ, Howpay Manage SA, Burrows CJ. 2021. Nanopore dwell time analysis permits sequencing and conformational assignment of pseudouridine in SARS-CoV-2. *ACS Cent Sci* 7:1707–1717. <https://doi.org/10.1021/acscentsci.1c00788>.
 115. Yankova E, Blackaby W, Albertella M, Rak J, De Braekeleer E, Tsagkogeorga G, Pilka ES, Aspris D, Leggate D, Hendrick AG, Webster NA, Andrews B, Fosbeary R, Guest P, Irigoyen N, Eleftheriou M, Gozdecka M, Dias JML, Bannister AJ, Vick B, Jeremias I, Vassiliou GS, Rausch O, Tzelepis K, Kouzarides T. 2021. Small-molecule inhibition of METTL3 as a strategy against myeloid leukaemia. *Nature* 593:597–601. <https://doi.org/10.1038/s41586-021-03536-w>.
 116. Srinivas KP, Depledge DP, Abebe JS, Rice SA, Mohr I, Wilson AC. 2021. Widespread remodeling of the m6A RNA-modification landscape by a viral regulator of RNA processing and export. *Proc Natl Acad Sci U S A* 118:e2104805118. <https://doi.org/10.1073/pnas.2104805118>.
 117. Kovaka S, Zimin AV, Perteau GM, Razaghi R, Salzberg SL, Perteau M. 2019. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol* 20:278. <https://doi.org/10.1186/s13059-019-1910-1>.
 118. Tang AD, Soulette CM, van Baren MJ, Hart K, Hrabetova E, Wu CJ, Brooks AN. 2020. Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. *Nat Commun* 11:1438. <https://doi.org/10.1038/s41467-020-15171-6>.

119. Braspenning SE, Sadaoka T, Breuer J, Verjans GMGM, Ouwendijk WJD, Depledge DP. 2020. Decoding the architecture of the varicella-zoster virus transcriptome. *mBio* 11:e01568-20. <https://doi.org/10.1128/mBio.01568-20>.
120. Whisnant AW, Jürges CS, Hennig T, Wyler E, Prusty B, Rutkowski AJ, L'hernault A, Djakovic L, Göbel M, Döring K, Menegatti J, Antrobus R, Matheson NJ, Künzig FWH, Mastrobuoni G, Bielow C, Kempa S, Liang C, Dandekar T, Zimmer R, Landthaler M, Grässer F, Lehner PJ, Friedel CC, Erhard F, Dölken L. 2020. Integrative functional genomics decodes herpes simplex virus 1. *Nat Commun* 11:2038. <https://doi.org/10.1038/s41467-020-15992-5>.
121. Braspenning SE, Verjans GMGM, Mehraban T, Messaoudi I, Depledge DP, Ouwendijk WJD. 2021. The architecture of the simian varicella virus transcriptome. *PLoS Pathog* 17:e1010084. <https://doi.org/10.1371/journal.ppat.1010084>.
122. Finkel Y, Schmiedel D, Tai-Schmiedel J, Nachshon A, Winkler R, Dobesova M, Schwartz M, Mandelboim O, Stern-Ginossar N. 2020. Comprehensive annotations of human herpesvirus 6A and 6B genomes reveal novel and conserved genomic features. *Elife* 9:e50960. <https://doi.org/10.7554/eLife.50960>.
123. Price AM, Hayer KE, Depledge DP, Wilson AC, Weitzman MD. 2019. Novel splicing and open reading frames revealed by long-read direct RNA sequencing of adenovirus transcripts. *bioRxiv*. <https://doi.org/10.1101/2019.12.13.876037>.
124. Ferguson JM, Smith MA. 2019. SquiggleKit: a toolkit for manipulating nanopore signal data. *Bioinformatics* 35:5372–5373. <https://doi.org/10.1093/bioinformatics/btz586>.