Journal of Microbiology (2022) Vol. 60, No. 9, pp. 867–876
DOI 10.1007/s12275-022-2324-4
eISSN 1976-3794
pISSN 1225-8873

PROTOCOL

Analyzing viral epitranscriptomes using nanopore direct RNA sequencing[§]

Ari Hong[1,2†], Dongwan Kim[1,3†], V. Narry Kim[1,3], and Hyeshik Chang[1,2,3*]

[1]*Center for RNA Research, Institute for Basic Science (IBS), Seoul National University, Seoul 08826, Republic of Korea*
[2]*Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul 08826, Republic of Korea*
[3]*School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea*

(Received Jul 18, 2022 / Revised Aug 4, 2022 / Accepted Aug 5, 2022)

RNA modifications are a common occurrence across all domains of life. Several chemical modifications, including N^6-methyladenosine, have also been found in viral transcripts and viral RNA genomes. Some of the modifications increase the viral replication efficiency while also helping the virus to evade the host immune system. Nonetheless, there are numerous examples in which the host's RNA modification enzymes function as antiviral factors. Although established methods like MeRIP-seq and miCLIP can provide a transcriptome-wide overview of how viral RNA is modified, it is difficult to distinguish between the complex overlapping viral transcript isoforms using the short read-based techniques. Nanopore direct RNA sequencing (DRS) provides both long reads and direct signal readings, which may carry information about the modifications. Here, we describe a refined protocol for analyzing the RNA modifications in viral transcriptomes using nanopore technology.

Keywords: RNA virus, RNA modification, viral epitranscriptome, coronavirus, nanopore sequencing, direct RNA sequencing

Overview

The eukaryotic messenger RNAs often include chemically modified RNA bases in addition to the standard four ribonucleotides. N^6-methyladenosine (m^6A), which occurs as frequently as 2–3 in an average human mRNA molecule, is the most prevalent type of modification (Rottman *et al.*, 1974). The m^6A modification regulates a variety of cellular processes, including translation, RNA stability, splicing, and localization depending on the context (Zaccara *et al.*, 2019). Another common RNA modification is pseudouridine, which makes up about 0.3% of all uridines in mRNA (Li *et al.*, 2015). It is frequently associated with modulation of codon decoding and also has been utilized to avoid triggering the innate immune system initiated by cytoplasmic double-stranded RNAs in mRNA vaccines (Karikó *et al.*, 2005; Cerneckis *et al.*, 2022). Viral genomes constantly evolve in order to take advantage of new opportunities for host exploitation. RNA modifications are found in viral genomes and transcripts in several viral species, including HIV-1, HBV, SV40, Zika, and HCV (Baquero-Perez *et al.*, 2021). Many of these enhance the replication efficiency of the viruses, yet anti-viral RNA modifications also exist that may be triggered by the host defense mechanisms (Williams *et al.*, 2019; Courtney, 2021; Li and Rana, 2022).

Discovering, identifying, and quantifying RNA modifications is critical for understanding the dynamics and mechanisms of the viral epitranscriptome. Among the various approaches for measuring or detecting the RNA modifications (Table 1), the most fundamental tools for RNA modifications are small fragment analysis techniques, such as liquid chromatography-mass spectrometry (LC-MS) or nuclear magnetic resonance (NMR) (Wetzel and Limbach, 2016; Yoluç *et al.*, 2021). These methods allow for the measurement of the stoichiometry between modified and unmodified nucleotides. Thus, these can be used to identify novel RNA modifications in isolated RNA samples as well as to identify global RNA modification changes brought on by perturbation of writers or readers.

When combined with high-throughput RNA sequencing, specific antibody purification or chemical reactions allow for a survey of the nucleotide sequences nearby the modified bases. For instance, MeRIP-seq identifies m^6A-modified sites by sequencing RNA fragments that have been immunopurified with a modification-specific antibody (Meyer *et al.*, 2012), whereas miCLIP improves resolution and specificity by adding additional UV-crosslinking between the antibody and modified bases (Hussain *et al.*, 2013). Alternatively, bisulfite or N-cyclohexyl-N'-(2-morpholinoethyl)carbodiimide metho-p-toluenesulfonate (CMCT) treatments can reveal 5-methylcytidine (m^5C) or pseudouridine (ψ) through the differential introduction of chemical-induced reverse-transcription errors (Schaefer *et al.*, 2009; Carlile *et al.*, 2014; Lovejoy

†These authors contributed equally to this work.
*For correspondence. E-mail: hyeshik@snu.ac.kr
§Supplemental material for this article may be found at https://doi.org/10.1007/s12275-022-2324-4.

Table 1. RNA modification mapping techniques

	Modification type	Discriminative feature	Context for analysis	Advantages	Disadvantages
LC-MS Wetzel and Limbach (2016)	Many	Mass to charge ratio	< ~3–5 nt	-Quantitative -Can identify unknown modifications	-Low sensitivity -Requires extensive purification for mRNA
MeRIP-seq Meyer *et al.* (2012)	m^6A, or any with high quality Ab	Antibody affinity	Sites in gene-level	-Easy to adopt -Highly accessible	-Low resolution of sequence context -High risk of false positives
miCLIP Hussain *et al.* (2013)	m^6A, m^5C	Antibody affinity	Sites in gene-level	-High specificity -Single-nucleotide resolution	-Large amount input RNA -Complex and inefficient procedures
Bisulfite sequencing Schaefer *et al.* (2009)	m^5C, hm^5C	Resistance to deamination	Sites in gene-level	-Single-nucleotide resolution	-Extensive RNA degradation -Large amount input RNA
CMCT sequencing Carlile *et al.* (2014); Lovejoy *et al.* (2014); Schwartz *et al.* (2014)	ψ	Selective chemical modification	Sites in gene-level	-Single-nucleotide resolution	-Extensive RNA degradation -Large amount input RNA -Low sensitivity
Nanopore (differential error) Jenjaroenpun *et al.* (2021); Abebe *et al.* (2022a)	Many	Basecalling error	Sites in isoform-level	-Easy and fast preparation -Near single-nucleotide resolution	-Requires a good control -Low sensitivity for rarely modified sites
Nanopore (signal-level) Stoiber *et al.* (2017); Leger *et al.* (2021)	Many	Ionic conductance	Sites in each single molecule	-Easy and fast preparation -Near single-nucleotide resolution -Single-molecule association analysis	-Requires high-coverage control signals -Single-molecule accuracy is often not good enough -Often requires a custom analytic workflow

et al., 2014; Schwartz *et al.*, 2014). The sequences surrounding the modified sites offer additional details about where they are found within transcripts, suggesting sequence specificity, regulatory mechanisms, and biological functions. However, the contextual information of the RNA modification is limited to just adjacent regions due to the short fragment length of sequenced reads from these methods. As a result, this approach is ineffective for determining the relationships between RNA modifications and distal variations such as alternative splicing, poly(A) tails, or other modifications at different sites (Wiener and Schwartz, 2021). Furthermore, many viruses have compact genomes that produce multiple coterminal subgenomic transcripts that share a significant portion with other subgenomic transcripts (Depledge *et al.*, 2019). Because of this, short-read techniques using second-generation sequencing are particularly insufficient for examining viral epitranscriptomes.

Third-generation sequencing technology is a promising platform to address the issues. Pacific Biosciences' single-molecule real-time (SMRT) sequencing and Oxford Nanopore Technologies' nanopore sequencing both have the ability to report modified bases along with full-length RNA sequences at the single-molecule level. Nanopore sequencing tracks changes in ionic conductance as the native RNA molecule passes through a small pore protein, in contrast to other sequencing technologies that synthesize complementary strands for sequencing. Thus, this is considered to have the potential to detect a wide range of RNA modifications in addition to the full-length RNA sequences (Garalde *et al.*, 2018; Workman *et al.*, 2019). This property makes nanopore direct RNA sequencing (DRS) a favorable method for studying post-transcriptional virus-host interactions.

In this article, we present an optimized protocol for examining the epitranscriptome of a variety of viral species. It should be noted that there is still no single best method for DRS analysis to find modified RNA bases that works in all circumstances (Abebe *et al.*, 2022b). We describe the approaches to try first, but other strategies might also be needed depending on the type of RNA modification, frequency, sequence contexts, and other variables. For more information on this issue, see a comprehensive review of various methods for viral epitranscriptome analysis using DRS (Abebe *et al.*, 2022b).

Applications

In 2019, Viehweger *et al.* (2019) published the first use of DRS for analyzing RNA modifications in the viral transcriptome. They employed DRS to sequence the human coronavirus 229E and sorted the full-length reads as per the subgenomic RNAs that they came from. By comparing the ion current signals to Tombo's m^5C model (Stoiber *et al.*, 2017), they were able to identify dozens of sites that are uniformly methylated irrespective of the sgRNA isoform. Unsurprisingly, studies seeking m^6A in viral transcriptomes were quickly followed. Price *et al.* (2020) looked at the dynamics of m^6A modifications throughout the adenovirus life cycle using DRS. By using DRUMMER (Abebe *et al.*, 2022a) to compare the signals from wild-type control and METTL3-knockout samples, they were able to spot the signal changes brought on by the depletion of the methyltransferase enzyme. Using the DRACH-motif as a guide, they were able to narrow down the modifications to a single-nucleotide level of statistical significance. With isoform-level analyses enabled by full-length reads, the researchers could conclude that METTL3's m^6A modification specifically affects how late viral genes are expressed. In a 2021 study, Srinivas *et al.* (2021) discovered that a human herpes simplex virus 1 (HSV-1) protein, ICP27, redistributes m^6A factors from the

nucleus in fibroblasts. The authors used DRS to examine whether the methylation of host messenger RNAs is impacted by the expression of viral proteins. The data show that expression of the viral protein significantly lowers modification-induced basecall errors, thus suggesting the viral protein's role in releasing m^6A methylation factors from the nucleus.

With the start of the COVID-19 pandemic, SARS-CoV-2 has quickly risen as the most studied virus for research using nanopore DRS. Early in 2020, Kim *et al.* (2020b) used both short-read sequencing and nanopore DRS to map the overall transcriptome architecture and epitranscriptomic features, which include RNA modifications and poly(A) tails, of SARS-CoV-2 subgenomic RNAs. This study used *in vitro* transcribed (IVT) controls, which contain only the standard four bases and no modified bases. They used Tombo to estimate m^5C methylation status, but the measured methylation rates between the viral RNAs and IVT controls were almost identical, reflecting the apparently high false positive rate. By comparing viral RNAs to IVT controls, they found 41 modified sites, many of which are differentially methylated based on the length of subgenomic RNAs. As the pandemic persisted, more research into specific modifications followed. Burgess *et al.* (2021) identified m^6A modifications in SARS-CoV-2 and another human coronavirus OC43 by comparing control and METTL3 inhibitor treated cells. Chang *et al.* (2021) conducted a time-course DRS analysis throughout the SARS-CoV-2 infection timeline. The RNA modifications were found to remain unchanged throughout the infection. Two more studies confirmed that the SARS-CoV-2 RNAs' m^6A modification is controlled by the host factors METTL3 and FTO and that the extent of methylation varies among the virus variants (Campos *et al.*, 2021; Zhang *et al.*, 2021). Finally, a different approach led to the discovery of five new pseudouridine sites in the SARS-CoV-2 genome (Fleming *et al.*, 2021). The study employed dwell-time analysis, which made it possible to find delayed translocation as well as a changed current signal around the modified sites.

Methods

Viral RNA extraction and sequencing

Preparing RNA extracts of virus-infected cells: virus-infected cells are prepared using a cell line permissive for extensive viral growth. A high viral RNA fraction among total RNA is frequently preferred to achieve the required read coverage for viral transcripts with low expression. Vero and its derivatives are generally suitable for many mammalian viruses (Ammerman *et al.*, 2008). Still, researchers should acknowledge that some cultured host cell types lack a specific component of the host machinery, resulting in inefficient or non-specific viral RNA modification. Infections with a high multiplicity of infection (MOI) generate more sequencing reads; however, in this case, the condition may diverge too far from the natural transcriptome due to insufficient host machinery for viral RNA processing or maturation. The highest purity and quality of RNA, essential for use in DRS, is typically produced when using cultured cells and TRIzol or comparable reagents. DNase I is applied to the extracted RNA to remove any DNA contamination that may interfere with the subsequent enzymatic reactions in library preparation.

***In vitro* transcription**: it is highly recommended to use a negative control library that accurately reflects the nucleotide sequences of viral RNA to mitigate the high false positive rate in model-based classifiers without controls. Given that lengthy amplicons and templates are problematic in both PCR and in vitro transcription, it is preferable to split the genomic sequence into 2 kb tiles with small overlaps. Resquiggling is required for many types of signal analysis, but it frequently results in alignment errors at either end of the signals. Small tile overlaps can help solve such issues and make up for reads that have less coverage near their 5′ ends. DRS reads molecules in the sequencing library from 3′ end to 5′ end, and at a largely constant rate, the sequencing of a molecule terminates even in the middle of an RNA molecule. As a result, for large tiles, the read coverage near the 5′ ends of each tile

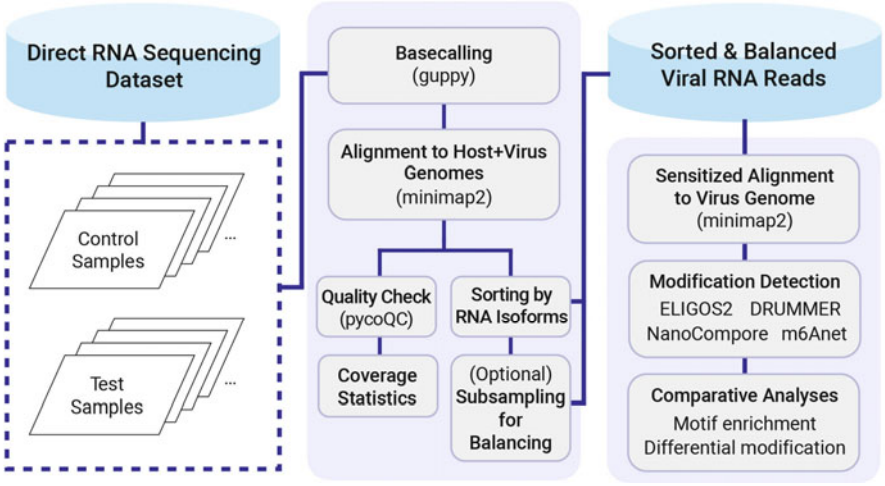

Fig. 1. Outline of modification analysis workflow. Following the sequencing of the viral transcriptome, the reads are subjected to primary analysis, which involves basecalling, alignment to both viral and host genomes, and quality checks. After that, reads are grouped according to the viral RNA isoforms and subsampled for easier handling and more accurate statistics. The subsequent modification analysis consists of realigning the viral genome to increase alignment sensitivity, detecting modifications, and then performing comparative analyses to interpret the data biologically.

is often insufficient. In addition, for difficult-to-PCR sequences, such as those with high GC content, tiling frequency and primer positions may need to be changed. Some IVT libraries might need several iterations of design-IVT-sequencing because some problems are challenging to detect otherwise. The targeting sensitivity of the modification detection may be considered when determining the target coverage of IVT control. The coverage of 2000X in the shallowest region should deliver good enough performance while still being feasible with reasonable effort.

Nanopore direct RNA sequencing: the standard direct RNA sequencing kit (SQK-RNA002, Oxford Nanopore Technologies) employs a 3′ adapter ligation that depends on poly(A) tails without any other nucleotides close to the 3′ end. Viral transcripts lacking poly(A) tails are not picked up during library preparation, making them invisible in the sequencing data. *In vitro* polyadenylation should come before library preparation for such viruses. It is known that the poly(A) tails of many viral transcripts contain mixed tails, such as A/G or U tails (Kim *et al.*, 2020a). It is necessary to perform a prior polyadenylation because mixed tails in the adapter ligation pose a challenge. Without a special adaptation for the mixed tails, data interpretation should be cautious due to their uneven distribution across the cellular RNAs and preference for a subset of cytoplasmic RNAs with long poly(A) tails. One million to two million reads are typically produced by a single DRS run. Because most viruses have small genomes, a single run usually produces more reads than is necessary. However, the sequencing reads of viral origin may not be sufficient for analysis depending on the characteristics of the viral life cycle. In this sort of situation, one should either pool data from multiple sequencing runs or attempt to enrich viral RNAs using an antisense oligo pull-down method.

Post-processing and data analysis

The primary analysis includes basecalling, sequence alignment, quality control, transcript sorting, and optional subsampling (Fig. 1). Comparisons between samples under two or more conditions or the IVT control are then made using one of the RNA modification detection programs. Following the discovery of modifications, potential biological associations are investigated in line with the researcher's own hypothesis.

Basecalling and sequence alignment: it is advised to perform basecall again using the high-accuracy model if the sequencing run was configured with fast basecalling or no basecalling. Minimap2 is commonly used for sequence alignment of nanopore reads (Li, 2018). The alignment programs are typically tuned against non-canonical junctions that are not likely produced by eukaryotic RNA splicing mechanisms by default. For viruses that produce transcripts as discontinuous copies of the genome, extensive parameter tuning is needed. Additionally, sequence alignment for viral genomes can be improved upon over the conventional sensitivity settings, taking into account the high error rate of both DRS and viral replication. To enable more sensitive alignment against the viral genome, align the sequences to both the host and viral genomes, then align the subset of reads that were aligned to the viral genome with more sensitive options. Most viral genomes can have minimap2's k-mer size and minimizer window size reduced to -k6 and -w2, respectively, without sacrificing scalability.

Subsampling reads for RNA modification analysis: most RNA modification detection methods look for differences between two populations at each transcript position. In the comparisons, uneven sample sizes for each transcript position in the control group lead to unevenly applied statistical power across the viral genome. The biological interpretation becomes challenging as a result. Since coverage continuously decreases from the 3′ side to the 5′ side in every tile of IVT controls, choosing the same number of full-length reads that span the 5′ most positions of each tile should work (Fig. 2). Subsampling is also occasionally required for biological RNA reads, such as wild-type or knock-down sample reads. Because the distribution of viral transcript expression levels is extremely varied, transcripts with excessive abundance, such as nucleocapsid mRNAs in *Nidovirales*, occasionally consume a tremendous amount of computational time and memory and cause numerical data type overflows in analysis software. To reduce the computational workload, consider random subsampling of reads from transcripts with a high abundance for RNA modification detection.

RNA modification detection: modification detection can be performed at either the basecalled sequence or the ionic current signal levels. In most cases, it is practical to begin the analysis with an algorithm that uses differential error rates at the site level. The most popular programs for this approach are ELIGOS2 and DRUMMER (Jenjaroenpun *et al.*, 2021; Abebe *et al.*, 2022a). With this method, it is possible to compile a list of potential transcripts that are frequently or differently modified and analyze any correlations at the transcript level, such as those between modification rate and translation activity. Signal-level analyses are necessary for more complex analyses, such as single-molecule level analysis, enhancing sensitivity, and associations between regulatory elements within a molecule. The majority of analysis software in this approach employs "resquiggled" signals, which are scaled signals that alleviate uneven translocation speed through the nanopore by stretching and constricting them. Nanopolish and Tombo are the two most popular applications for resquiggling (Stoiber *et al.*, 2017). In addition to ionic current

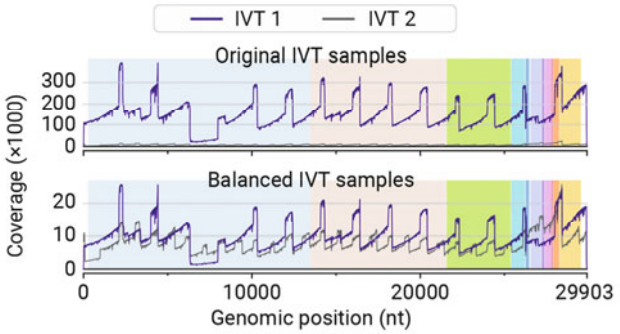

Fig. 2. Balancing IVT control coverage. (A) Raw read coverage before downsampling for balanced read depth distribution. Two IVT control libraries were constructed with different tiling widths, 2,000 and 1,000, respectively. The spannings of known SARS-CoV-2 open reading frames are represented by background shades. (B) Read coverage in balanced subsamples of the IVT control libraries.

level changes, dwell time can be considered when comparing samples. It is known that certain RNA modifications cause a significant slowdown in the translocation of the modified base near its -20 position (Stephenson *et al.*, 2022).

Materials

Reagents

- TRIzol (Invitrogen, 15596018)
- Chloroform (Amresco, 0757-500ML)
- Isopropanol (Merck, 1.09634.1011)
- Ethanol (Fisher, A9951)
- Dithiothreitol (DTT) (Sigma-Aldrich, 43819-5F)
- Qubit RNA High Sensitivity Assay Kit (Thermo Fisher Scientific, Q32852)
- Recombinant DNase I (RNase-free) (TaKaRa, 2270A)
- RNeasy MinElute Cleanup Kit (Qiagen, 74204)
- SuperScript IV Reverse Transcriptase (Invitrogen, 18090200)
- Q5 High-Fidelity 2X Master Mix (New England Biolabs, M0492L)
- 10X Loading Buffer (TaKaRa, 9157)
- Agarose (Seakem, 50004)
- Gel Extraction Kit (Labopass, CMG0112)
- MEGAscript T7 Transcription Kit (Invitrogen, AMB13345)
- TAE buffer (Dyne Bio, CBT3021)
- Gelgreen dye (Biotium, 41005)
- TURBO DNase (Ambion, AM2239)
- Oligo Clean & Concentrator (Zymo Research, D4061, optional)
- 1 kb Plus DNA Ladder (Invitrogen, 10787026)
- SUPERase·In RNase Inhibitor (Invitrogen, AM2696)
- Deionized Purified Water (obtained in house from Milli-Q Ultrapure Water System)
- DNA LoBind Microcentrifuge 1.5 ml tubes (Eppendorf, 022431021)
- Direct RNA Sequencing Kit (Oxford Nanopore Technologies, SQK-RNA002)
- MinION Flow Cell (R9.4.1) (Oxford Nanopore Technologies, FLO-MIN106D)

Equipment

- Thermocycler
- NanoDrop (Thermo Fisher Scientific)
- Safe Imager 2.0 Blue-Light Transilluminator (Thermo Fisher Scientific, G6600)
- MinION Mk1b or Mk1c Device (Oxford Nanopore Technologies, MIN-101B)
- Computer running Linux or an equivalent environment that is capable of running the software listed below. It is recommended to install a GPU card because basecalling with CPUs alone takes several days rather than one or two hours. At the time of writing, Guppy, the ONT production basecaller, only supports NVIDIA GPUs.

Computer software

- Guppy 6.1.2 (Oxford Nanopore Technologies)
- Minimap2 2.17 (Li, 2018)
- Bedtools 2.30.0 (Quinlan, 2014)
- Seqtk 1.3 (https://github.com/lh3/seqtk)
- Samtools 1.7 (Li *et al.*, 2009)
- PycoQC 2.5.0.19 (Leger and Leonardi, 2019)
- ELIGOS2 2.0.1 (Jenjaroenpun *et al.*, 2021)
- Nanopolish 0.13.2 (https://github.com/jts/nanopolish)
- Picard Toolkit 2.27.4 (https://github.com/broadinstitute/picard)
- m6Anet 1.1.0 (Hendra *et al.*, 2021)

Protocols

Construction of sequencing libraries

A. RNA purification

1. This step begins with cell pellets in 1 ml TRIzol in a 1.5 ml tube. Pellets should be lysed completely.
2. Incubate the tube at 50°C for 15 min, occasionally inverting the tube.
3. Add 200 µl chloroform.
   ▸ **NOTE**: perform this step in a fume hood.
4. Vortex vigorously.
   ▸ **NOTE**: hold the tube cap tight to avoid leakage.
5. Incubate the tube at 25°C for 3 min.
6. Vortex vigorously.
   ▸ **NOTE**: hold the tube cap tight to avoid leakage.
7. Centrifuge the tube at 4°C for 10 min.
   ▸ **NOTE**: close the lid to avoid aerosols.
8. Carefully move the tube to the rack to avoid mixing two separate phases.
9. Set a P200 pipette to 150 µl, move all supernatants carefully (~600 µl) into a new tube. Discard the tube with the lower phase.
   ▸ **NOTE**: aqueous phase solution is prone to leakage.
10. Add 600 µl isopropanol.
11. Vortex vigorously.
    ▸ **NOTE**: hold the tube cap tight to avoid leakage.
12. Incubate the tube at 25°C for 10 min, occasionally vortexing or inverting the tube.
13. Centrifuge the tube at 4°C for 30 min.
14. Check if there is a white pellet. Discard the supernatant carefully.
    ▸ **NOTE**: always check if the supernatant does not contain the pellet.
15. Centrifuge the tube at 4°C for 10 sec.
16. Discard the supernatant carefully with a P20 pipette.
17. Add 1 ml pre-chilled 75% ethanol and vortex.
18. Centrifuge the tube at 4°C for 10 min.
19. Discard the supernatant carefully and repeat steps 17–18.
20. Centrifuge the tube at 4°C for 10 sec.
21. Discard the supernatant carefully with a P20 pipette.
    ▸ **NOTE**: remove ethanol as much as possible.
22. Open the tube cap and air-dry for 3 min. You can see a transparent pellet turn white.
23. Add 30 µl water and dissolve the pellet by pipetting.

▸ **NOTE**: the amount of water varies according to the initial cell amount.

24. Measure the RNA concentration by using the Qubit RNA HS Assay Kit or NanoDrop.

B. DNase I treatment and purification

If the amount of RNA is low (< 10 μg) and the purity is good, DNase I treatment and purification step may be skipped to avoid loss of RNA.

25. Mix RNA (up to 45 μg), water up to 87 μl, 10 μl 10X DNase I buffer, 2 μl DNase I, and 1 μl SUPERase-In RNase Inhibitor in 1.5 ml tube by pipetting.
    ▸ **NOTE**: do not vortex. DNase I is sensitive to physical denaturation.
26. Incubate the tube at 37°C for > 30 min.
27. Add 350 μl RLT buffer and mix by pipetting.
28. Add 250 μl of 100% ethanol and mix by pipetting.
29. Move all solutions to a RNeasy MinElute spin column.
30. Centrifuge at 25°C for 1 min.
31. Discard the flow-through.
32. Add 500 μl of RPE buffer to the column.
33. Centrifuge at 25°C for 1 min.
34. Discard the flow-through.
35. Add 500 μl 80% ethanol into the column.
36. Centrifuge at 25°C for 1 min.
37. Discard the flow-through.
38. With a P20 pipette, remove residual ethanol from the ring around the membrane of the column and from the ring around the cap.
39. Centrifuge at 25°C for 2 min.
40. Move the column to a new 1.5 ml tube.
41. Open the tube cap and air-dry at 25°C for 3 min.
42. Add > 20 μl water directly to the membrane.
    ▸ **NOTE**: be sure to put water into the membrane, not into the ring.
43. Incubate at 25°C for 1 min.
44. Centrifuge at 25°C for 1 min.
45. Check the eluate and measure the RNA concentration by using Qubit RNA HS Assay Kit or NanoDrop.

C. Reverse transcription

• Use thermocycler

46. Prepare 200 μl PCR tube and mix RNA (200 ng–1 μg), water up to 11 μl, 1 μl of each 10 μM virus-specific RT primer and 1 μl of 10 mM dNTPs.
    ▸ **NOTE**: required amount of RNA varies depending on the sample. As viral replication is faster and the cytopathic effect is milder, a lesser amount of RNA would be needed.
47. Incubate at 65°C for 10 min and 4°C until the next step.
48. Add 4 μl of 5X SuperScript IV Buffer, 1 μl of 0.1 M dithiothreitol (DTT), 1 μl of SUPERase-In RNase Inhibitor (or RNaseOUT Recombinant RNase Inhibitor, Invitrogen), and 1 μl of SuperScript IV Reverse Transcriptase.
    ▸ **NOTE**: use fresh DTT.
49. Briefly spin down and mix by pipetting.
50. Incubate at 50°C for > 30 min.
51. Incubate at 80°C for 10 min and store at 4°C until the next step.

D. PCR and agarose gel purification for in vitro transcription templa

• Use thermocycler

52. Prepare 200 μl PCR tube and mix cDNA (e.g. 0.5 μl), water up to 12.5 μl, 1 μl of 10 μM forward PCR primer, 1 μl of 10 μM reverse PCR primer, and 15 μl of Q5 High-Fidelity 2X Master Mix by pipetting.
    ▸ **NOTE**: the required amount of cDNA varies depending on the sample.
53. Run PCR program: 98°C for 30 sec, (98°C for 20 sec, 60°C for 30 sec, 72°C for 1 min 30 sec) × 25–30 cycles, 72°C for 1 min, and hold at 16°C.
    ▸ **NOTE**: annealing temperature, the number of PCR cycles, and extension time vary depending on the GC content of the primer, the amount of cDNA used, and amplicon size, respectively.
54. After the run, add 3.4 μl of 10X loading buffer.
55. Prepare 1% agarose gel by using a wide well enough to load ~35 μl PCR product, if applicable.
56. Load ~35 μl PCR product and 10–20 μl 1 kb plus ladder.
57. Run the gel at 130 V for 25 min using the 1X TAE buffer.
58. After the run, stain the gel with Gelgreen dye in a fresh 1X TAE buffer for 10 min.
    ▸ **NOTE**: use a fresh 1X TAE buffer.
59. Weigh the empty 1.5 ml tube.
60. Put the gel on Safe Imager 2.0 Blue-Light Transilluminator, cut the gel with a razor and move the gel into a new 1.5 ml tube.
    ▸ **NOTE**: DO NOT use UV illuminator to avoid DNA damage.
61. Weigh the gel slice.
62. Add 3X volume of GB buffer from Gel extraction kit.
63. Vortex for 5 sec and incubate at 50°C for 10 min with occasional vortexing.
    ▸ **NOTE**: be sure that the gel is perfectly melted by checking that there is no haze. If the solution is orange, add 3 M sodium acetate following the instructions.
64. Load the solution onto a spin column.
65. Centrifuge at 25°C for 30 sec and discard the flow-through.
66. Load a 750 μl NW buffer on a spin column, centrifuge at 25°C for 30 sec and discard the flow-through.
67. Repeat step 66 once.
68. With a P20 pipette, remove the residual buffer at the ring around the membrane of the column and at the ring around the cap.
69. Centrifuge at 25°C for 2 min.
70. Move the column to a new 1.5 ml tube.
71. Open the tube cap and air-dry at 25°C for 3 min.
72. Add > 20 μl water directly to the membrane.
    ▸ **NOTE**: be sure to put water into the membrane, not into the ring.
73. Incubate at 25°C for 1 min.
74. Centrifuge at 25°C for 1 min.
75. Check the eluate and measure the DNA concentration by NanoDrop.

E. *In vitro* transcription and RNA purification

• Use thermocycler

76. Prepare 200 μl PCR tube and mix IVT template DNA 5 μl (e.g. 50 ng), each NTP solution 2 μl, 10X Reaction buffer

2 µl, 10X Enzyme mix 2 µl, water 2 µl, 1 µl of SUPERase-In RNase Inhibitor by pipetting.
77. Incubate at 37°C overnight.
78. Add 1 µl of TURBO DNase and mix by pipetting.
79. Incubate at 37°C for > 15 min.
▸ **NOTE**: if IVT RNA is longer than 200 nt, recommend using RNeasy MinElute spin column to efficiently remove remaining NTPs during in vitro transcription. Otherwise, use Zymo Oligo Clean & Concentrator column to preserve smaller IVT RNAs although it cannot completely remove remaining NTPs, which affect the measurement of RNA concentration.
80. Add water 79 µl.
81. Add 350 µl RLT buffer and mix by pipetting.
82. Add 250 µl of 100% ethanol and mix by pipetting.
83. Move all solutions to a RNeasy MinElute spin column.
84. Centrifuge at 25°C for 1 min.
85. Discard the flow-through.
86. Add 500 µl of RPE buffer to the column.
87. Centrifuge at 25°C for 1 min.
88. Discard the flow-through.
89. Add 500 µl 80% ethanol into the column.
90. Centrifuge at 25°C for 1 min.
91. Discard the flow-through.
92. With a P20 pipette, remove residual ethanol from the ring around the membrane of the column and from the ring around the cap.
93. Centrifuge at 25°C for 2 min.
94. Move the column to a new 1.5 ml tube.
95. Open the tube cap and air-dry at 25°C for 3 min.
96. Add > 20 µl water directly to the membrane.
▸ **NOTE**: be sure to put water into the membrane, not into the ring.
97. Incubate at 25°C for 1 min.
98. Centrifuge at 25°C for 1 min.
99. Check the eluate and measure the RNA concentration by using Qubit RNA HS Assay Kit or NanoDrop.

F. Nanopore direct RNA sequencing

Follow the manufacturer's instructions with a few changes listed below (Direct RNA sequencing Kit, Oxford Nanopore Technologies). We used these procedures for SARS-CoV-2-infected Vero cells (Kim *et al.*, 2020b), but other virus or host cells may require a modification due to the different RNA length, poly(A) tail length distribution, and also non-coding RNA composition.

100. For nanopore sequencing on non-infected and SARS-CoV-2-infected Vero cells, each 4 µg of DNase I-treated total RNA in 8 µl was used for library preparation with minor adaptations.
a. Add 1 µl of SUPERase-In RNase inhibitor (20 U/µl) to both adapter ligation steps.
b. SuperScript IV Reverse Transcriptase was adopted instead of SuperScript III.
c. The reaction time of reverse transcription was lengthened by 2 h.
101. For nanopore sequencing on SARS-CoV-2 IVT RNA fragments, mix an equal amount of IVT RNA fragments for a total of 2 µg.
a. Add 1 µl of SUPERase-In RNase inhibitor (20 U/µl) to both adapter ligation steps.
b. SuperScript IV Reverse Transcriptase was adopted instead of SuperScript III for 30 min.

Data analysis

A. Preparing the reference sequences

Choose reference genome assemblies for the host cell line and the virus. Genomes from GENCODE or ENSEMBL are preferred because they provide a variety of preprocessed files and annotation files with extensive annotation. Because minimap2 only accepts a single file as a reference, the host genome sequences and viral genome must be combined into a single FASTA file. Pre-built index files can be made to avoid having to build index files from scratch each time a sample is processed.

1. Merge host and virus genome sequences into a single combined FASTA file.
$ cat host-genome.fasta virus-genome.fasta > hostvirus-genomes.fasta
2. Create minimap2 indices for each FASTA file, both combined and virus-only.
$ minimap2 -k 13 -w 6 -d hostvirus-genomes.mmidx hostvirus-genomes.fasta
$ minimap2 -k 6 -w 2 -d virus-genome.mmidx virus-genome.fasta

B. Basecalling and quality check

The supplementary code contains an example analysis pipeline and notebooks (see Supplementary data 1). For the procedure in more detail, refer to the command lines and Jupyter notebooks in the repository.

1. Transfer the FAST5 files to a computer for data analysis from the sequencer or host computer. Use both "pass" and "fail" reads because re-analysis may improve the basecall quality.
2. Run basecalling using guppy with the high accuracy RNA model.
$ guppy_basecaller -c rna_r9.4.1_70bps_hac.cfg --fast5_out --calib_detect --min_qscore 5 --num_callers 4 -i original-fast5 -s guppy-basecalls -x cuda:all
The input FAST5 files should be located in a directory named "original-fast5" and outputs are produced in "guppy-basecalls."
▸ **NOTE**: on computers without an NVIDIA GPU, the "-x cuda:all" command line option should be removed.
3. Merge the FASTQ sequences from the output into a single compressed FASTQ file.
$ zcat guppy-basecalls/pass/*.fastq | gzip -c - > pass-all.fastq.gz
4. Align sequences to the combined reference index.
$ minimap2 -a --MD -a -x splice -N 32 -un -t 20 hostvirus-genomes.mmidx pass-all.fastq.gz | samtools sort -o pass-hostvirus.bam
5. Run pycoQC to generate a quality check report. Inspect each report carefully to ensure that the quality of each sequencing run is even and contains sufficient viral sequence reads.
$ pycoQC -f guppy-basecalls/pass/sequencing_summary.txt -a pass-hostvirus.bam -o pycoQC.html
▸ **NOTE**: verify the distribution of reads by length and

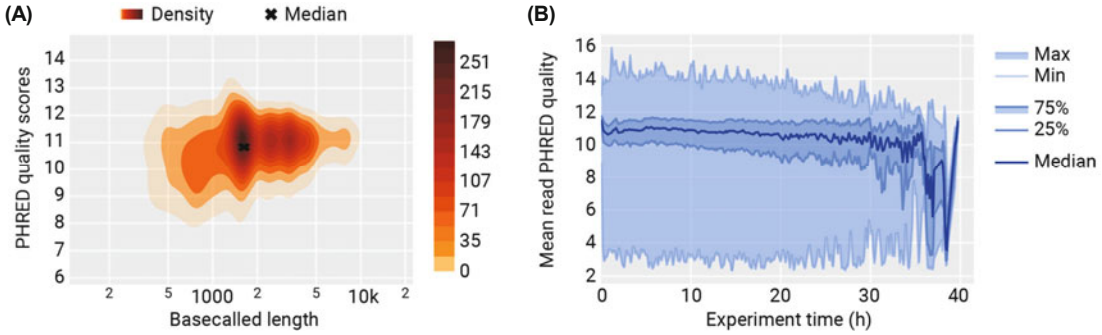

Fig. 3. Quality check plots from pycoQC. (A) Distribution of reads based on the basecalled length (x-axis) and the average basecall quality score (y-axis). As shown in, a typical DRS run has a peak density of ~1.3k nt (the length of the calibration control used in the DRS kit) in length and ~11 in phred score. In viruses with an aggressive growth rate, such as SARS-CoV-2, the peaks formed by abundant RNAs are also evident, such as the S (8k), ORF3a (4.6k), and M (3.4k) peaks seen here. (B) Average read quality changes over the course of the sequencing run. It might be necessary to limit the time window in which sequenced reads are to be used in order to avoid a significant decline in sequencing quality having a negative impact on the analysis of modifications.

quality as well as any changes in these parameters over time (Fig. 3). Those library quality factors must be consistent across all samples, within the expected range of distributions. The majority of techniques for DRS-based RNA modification detection are particularly sensitive to the overall basecall quality shift between samples.

C. Processing viral sequence reads

RNA modification analysis can be done for either all reads of viral origin combined or separately for each viral transcript. In order to compare differentially modified sites based on viral transcript, the viral reads must be divided into separate folders containing sorted FASTQ and FAST5 files. The majority of the time, read coverage greater than 2000X has little to no bearing on the accuracy or sensitivity of the analysis; if necessary, one can randomly subsample highly abundant transcripts for modification analysis.

1. Extract viral RNA reads to a separate file. The command that follows presumes that the viral genome sequence has the ID "NC_045512.2."
   $ samtools view -F4 pass-hostvirus.bam | grep NC_045512.2 | cut -f1 | sort | uniq > pass-hostvirus.virus-id.txt
   $ seqtk subseq pass-all.fastq.gz pass-hostvirus.virus-id.txt | gzip -c - > pass-virus.fastq.gz
2. Align the viral reads to the viral genome using more sensitive settings. To allow the virus-specific sequence recombination mechanisms in the alignments, additional parameter changes might be required.
   $ minimap2 –-splice -g 30000 -G 30000 -A1 -B2 -O 2,24 -E1, 0 -C0 -z 400,200 –-no-end-flt -F 40000 -N 350 –-splice-flank=no –-max-chain-skip=40 -un –-MD -a -p 0.7 -t 20 virus-genome.mmidx pass-virus.fastq.gz | samtools sort -o pass-virus.bam; samtools index pass-virus.bam
   ▸ **NOTE**: add these options to give accepted recombination sites some preference in locating long deletions: –-junc-bed viral-junction.bed –-junc-bonus=50
3. Optionally, downsample reads for transcripts with an unusually high abundance. This procedure requires some programming that is specific to the structure of the viral transcriptome. See rules from "extract_junctions" to "get_balanced_fastq" in "Snakefile" in the Supplementary data 1 code for an example.

D. RNA modification analysis using ELIGOS2

1. Run ELIGOS2 with the sorted and index bam file as the input file.
   $ eligos2 pair_diff_mod -tbam pass-virus.bam -cbam pass-control.bam -reg viral-annotation.bed -ref viral-genome.fasta -t 5 --pval 1 --oddR 0 --esb 0 -o eligos2-output/
   ▸ **NOTE**: select the sample with the greater expected modification rate as the "test" sample and the other as the "control" sample.
   ▸ **NOTE**: the ELIGOS2 default threshold values are too restrictive. Try relaxing the threshold values for p-value, odd ratio, and error at specific base (ESB) values, and then filter in subsequent steps following inspection of global patterns.
2. Perform further visualization and inspection of RNA modification distributions across the viral genome in an exploratory data analysis environment, such as JupyterLab or RStudio. See the Supplementary data 1 code for an example.

E. RNA modification analysis using m6Anet

Ionic current signals are supported by m6Anet for single-molecule level measurements for RNA modification detection. Try using different programs as well because the outcomes of programs built on this methodology frequently do not coincide.

1. Prepare a resquiggled signal database using nanopolish.
   $ nanopolish index -s guppy-basecalls/sequencing_summary.txt -d guppy-basecalls/workspace/ pass-virus.fastq.gz
   $ nanopolish eventalign -t 5 --signal-index -r pass-virus.fastq.gz -b pass-virus.bam -g viral-genome.fasta --summary=nanopolish-summary.txt --scale-events -n > nanopolish-eventalign.txt
2. Run m6Anet with the nanopolish eventalign outputs.
   $ m6anet-dataprep -i nanopolish-eventalign.txt -o m6anet-dataprep-pdir/
   $ m6anet-run_inference --input_dir m6anet-dataprep-dir/ --out_dir m6anet-outputdir/ --n_process 10
   ▸ **NOTE**: because m6Anet assumes the user inputted the transcriptome alignment, the minus strand is not evaluated. The procedure should be modified accordingly if the gene is expressed from both strands of the viral genome.

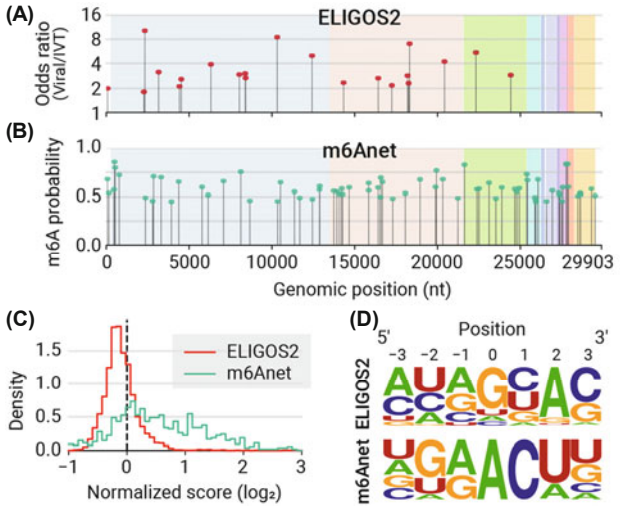

Fig. 4. RNA modifications detected using ELIGOS2 or m6Anet. (A) Predicted frequently modified sites detected by ELIGOS2. Red dots indicate sites with an error of specific bases (ESB) odds ratio of 1.8 or higher. Each ORF of SARS-CoV-2 is represented by a different shade. (B) Predicted most probable m^6A sites detected by m6Anet. Green dots represent sites with an m^6A probability of at least 0.44. (C) Distribution of $\log_2$ ratios between the scores in the viral RNA and the IVT control at each position. The percentage error of specific bases (%ESB) in ELIGOS2 and the m^6A probability in m^6Anet are used to calculate the normalized scores. (D) Sequence motif around the detected modified sites using ELIGOS2 and m6Anet. It should be noted that in ELIGOS2, the detected modifications are not always of a single type, resulting in a mixture of different modifications.

3. Compare the outcomes from ELIGOS2 with the outcomes from m6Anet by performing additional statistical analysis and visualization. Explore the data and analyze the data based on hypotheses about the mechanisms and interactions of host factors that affect methylation.

Expected Results

Once RNA modification detection is complete, the reported sites can be analyzed in a number of ways to determine their reliability, whether or not they are subject to systematic bias, and whether they are associated with any particular biological mechanism. Statistics at the site level highlight potential heavily modified sites and their global distribution throughout the viral genome (Fig. 4A–B). This suggests modifications at particular locations within the viral genome, clustering of the modification sites, or general frequencies of the modifications. It should be noted that these types of statistics may not capture conditionally regulated modification that maintains low equilibrial stoichiometry at a site level. Such circumstances call for in-depth analysis, which frequently involves read sorting by RNA isoform, feature quantification (such as poly[A] length or local folding energy), and statistical association screening.

The various analytic approaches report different sites and scores, but each program frequently has a unique ability to detect particular types of modifications in terms of modification type and sequence context (Fig. 4C–D). In our previous study with SARS-CoV-2, ELIGOS2 found a relatively small number of sites with high rates of modification of any kind, whereas m6Anet found that m^6A modification may be more widespread across the SARS-CoV-2 genome. Therefore, it should be easier to obtain potent statistical power in the result interpretations by preparing multiple conditions of viral RNA samples, such as latent and lytic phases.

Still in its very early stages, nanopore DRS anticipates years of continued growth and development. In late 2022, it is expected that new nanopore proteins with better homopolymer resolution will be made available for DRS. Analytical algorithms used in RNA modification are still underutilizing the information present in the squiggles. The accuracy and adaptability would be greatly enhanced if approaches similar to those used in state-of-art DNA basecallers were applied. Long and direct reads from DRS are the ideal features for high-throughput analysis of viral genomes and transcripts. Despite its limitations (low accuracy and a high rate of false positives), DRS-based transcriptome mapping is a cost-effective first step when analyzing a viral epitranscriptome.

Acknowledgements

This research was supported by the Institute for Basic Science funding from the Ministry of Science and ICT of Korea (IBS-R008-D1); the National Research Foundation of Korea (NRF) grants funded by the Korean government (2019R1-A6A1A10073437 to H.C. and 2021R1C1C1010758 to A.H. and H.C.).

Conflict of Interest

The authors have no conflict of interest to report.

Data Availability

Source codes for the pipeline and interactive notebooks are available from https://github.com/ChangLabSNU/viral-epitranscriptome-pipeline and Supplementary data 1.

References

Abebe, J.S., Price, A.M., Hayer, K.E., Mohr, I., Weitzman, M.D., Wilson, A.C., and Depledge, D.P. 2022a. DRUMMER-Rapid detection of RNA modifications through comparative nanopore sequencing. *Bioinformatics* **38**, 3113–3115.

Abebe, J.S., Verstraten, R., and Depledge, D.P. 2022b. Nanopore-based detection of viral RNA modifications. *mBio* **13**, e0370221.

Ammerman, N.C., Beier-Sexton, M., and Azad, A.F. 2008. Growth and maintenance of Vero cell lines. *Curr. Protoc. Microbiol.* **11**, A.4E.1–A.4E.7.

Baquero-Perez, B., Geers, D., and Díez, J. 2021. From A to m^6A: the emerging viral epitranscriptome. *Viruses* **13**, 1049.

Burgess, H.M., Depledge, D.P., Thompson, L., Srinivas, K.P., Grande, R.C., Vink, E.I., Abebe, J.S., Blackaby, W.P., Hendrick, A., Albertella, M.R., *et al.* 2021. Targeting the m^6A RNA modification pathway blocks SARS-CoV-2 and HCoV-OC43 replication. *Genes Dev.* **35**, 1005–1019.

Campos, J.H.C., Maricato, J.T., Braconi, C.T., Antoneli, F., Janini, L.M.R., and Briones, M.R.S. 2021. Direct RNA sequencing reveals SARS-CoV-2 m6A sites and possible differential DRACH motif methylation among variants. *Viruses* **13**, 2108.

Carlile, T.M., Rojas-Duran, M.F., Zinshteyn, B., Shin, H., Bartoli, K.M., and Gilbert, W.V. 2014. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* **515**, 143–146.

Cerneckis, J., Cui, Q., He, C., Yi, C., and Shi, Y. 2022. Decoding pseudouridine: an emerging target for therapeutic development. *Trends Pharmacol. Sci.* **43**, 522–535.

Chang, J.J.Y., Rawlinson, D., Pitt, M.E., Taiaroa, G., Gleeson, J., Zhou, C., Mordant, F.L., De Paoli-Iseppi, R., Caly, L., Purcell, D.F.J., *et al.* 2021. Transcriptional and epi-transcriptional dynamics of SARS-CoV-2 during cellular infection. *Cell Rep.* **35**, 109108.

Courtney, D.G. 2021. Post-transcriptional regulation of viral RNA through epitranscriptional modification. *Cells* **10**, 1129.

Depledge, D.P., Mohr, I., and Wilson, A.C. 2019. Going the distance: optimizing RNA-Seq strategies for transcriptomic analysis of complex viral genomes. *J. Virol.* **93**, e01342-18.

Fleming, A.M., Mathewson, N.J., Howpay Manage, S.A., and Burrows, C.J. 2021. Nanopore dwell time analysis permits sequencing and conformational assignment of pseudouridine in SARS-CoV-2. *ACS Cent. Sci.* **7**, 1707–1717.

Garalde, D.R., Snell, E.A., Jachimowicz, D., Sipos, B., Lloyd, J.H., Bruce, M., Pantic, N., Admassu, T., James, P., Warland, A., *et al.* 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods* **15**, 201–206.

Hendra, C., Pratanwanich, P.N., Wan, Y.K., Goh, W.S.S., Thiery, A., and Göke, J. 2021. Detection of m6A from direct RNA sequencing using a Multiple Instance Learning framework. *bioRxiv*. doi: https://doi.org/10.1101/2021.09.20.461055.

Hussain, S., Sajini, A.A., Blanco, S., Dietmann, S., Lombard, P., Sugimoto, Y., Paramor, M., Gleeson, J.G., Odom, D.T., Ule, J., and Frye, M. 2013. NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Rep.* **4**, 255–261.

Jenjaroenpun, P., Wongsurawat, T., Wadley, T.D., Wassenaar, T.M., Liu, J., Dai, Q., Wanchai, V., Akel, N.S., Jamshidi-Parsian, A., Franco, A.T., *et al.* 2021. Decoding the epitranscriptional landscape from native RNA sequences. *Nucleic Acids Res.* **49**, e7.

Karikó, K., Buckstein, M., Ni, H., and 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.

Kim, D., Lee, Y.S., Jung, S.J., Yeo, J., Seo, J.J., Lee, Y.Y., Lim, J., Chang, H., Song, J., Yang, J., *et al.* 2020a. Viral hijacking of the TENT4-ZCCHC14 complex protects viral RNAs via mixed tailing. *Nat. Struct. Mol. Biol.* **27**, 581–588.

Kim, D., Lee, J.Y., Yang, J.S., Kim, J.W., Kim, V.N., and Chang, H. 2020b. The architecture of SARS-CoV-2 transcriptome. *Cell* **181**, 914–921.

Leger, A., Amaral, P.P., Pandolfini, L., Capitanchik, C., Capraro, F., Miano, V., Migliori, V., Toolan-Kerr, P., Sideri, T., Enright, A.J., *et al.* 2021. RNA modifications detection by comparative Nanopore direct RNA sequencing. *Nat. Commun.* **12**, 7198.

Leger, A. and Leonardi, T. 2019. pycoQC, interactive quality control for Oxford Nanopore Sequencing. *J. Open Source Softw.* **4**, 1236.

Li, H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

Li, N. and Rana, T.M. 2022. Regulation of antiviral innate immunity by chemical modification of viral RNA. *Wiley Interdiscip. Rev. RNA*, e1720.

Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., and Yi, C. 2015. Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* **11**, 592–597.

Lovejoy, A.F., Riordan, D.P., and Brown, P.O. 2014. Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*. *PLoS ONE* **9**, e110799.

Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. 2012. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–1646.

Price, A.M., Hayer, K.E., McIntyre, A.B.R., Gokhale, N.S., Abebe, J.S., Della Fera, A.N., Mason, C.E., Horner, S.M., Wilson, A.C., Depledge, D.P., *et al.* 2020. Direct RNA sequencing reveals m^6A modifications on adenovirus RNA are necessary for efficient splicing. *Nat. Commun.* **11**, 6016.

Quinlan, A.R. 2014. BEDTools: The Swiss-Army tool for genome feature analysis. *Curr. Protoc. Bioinformatics* **47**, 11.12.1–11.12.34.

Rottman, F., Shatkin, A.J., and Perry, R.P. 1974. Sequences containing methylated nucleotides at the 5' termini of messenger RNAs: possible implications for processing. *Cell* **3**, 197–199.

Schaefer, M., Pollex, T., Hanna, K., and Lyko, F. 2009. RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Res.* **37**, e12.

Schwartz, S., Bernstein, D.A., Mumbach, M.R., Jovanovic, M., Herbst, R.H., León-Ricardo, B.X., Engreitz, J.M., Guttman, M., Satija, R., Lander, E.S., *et al.* 2014. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **159**, 148–162.

Srinivas, K.P., Depledge, D.P., Abebe, J.S., Rice, S.A., Mohr, I., and Wilson, A.C. 2021. Widespread remodeling of the m^6A RNA-modification landscape by a viral regulator of RNA processing and export. *Proc. Natl. Acad. Sci. USA* **118**, e2104805118.

Stephenson, W., Razaghi, R., Busan, S., Weeks, K.M., Timp, W., and Smibert, P. 2022. Direct detection of RNA modifications and structure using single-molecule nanopore sequencing. *Cell Genom.* **2**, 100097.

Stoiber, M., Quick, J., Egan, R., Lee, J.E., Celniker, S., Neely, R.K., Loman, N., Pennacchio, L.A., and Brown, J. 2017. *De novo* identification of DNA modifications enabled by genome-guided nanopore signal processing. *bioRxiv*. doi: https://doi.org/10.1101/094672.

Viehweger, A., Krautwurst, S., Lamkiewicz, K., Madhugiri, R., Ziebuhr, J., Hölzer, M., and 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.

Wetzel, C. and Limbach, P.A. 2016. Mass spectrometry of modified RNAs: recent developments. *Analyst* **141**, 16–23.

Wiener, D. and Schwartz, S. 2021. The epitranscriptome beyond m^6A. *Nat. Rev. Genet.* **22**, 119–131.

Williams, G.D., Gokhale, N.S., and Horner, S.M. 2019. Regulation of viral infection by the RNA modification *N6*-methyladenosine. *Annu. Rev. Virol.* **6**, 235–253.

Workman, R.E., Tang, A.D., Tang, P.S., Jain, M., Tyson, J.R., Razaghi, R., Zuzarte, P.C., Gilpatrick, T., Payne, A., Quick, J., *et al.* 2019. Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nat. Methods* **16**, 1297–1305.

Yoluç, Y., Ammann, G., Barraud, P., Jora, M., Limbach, P.A., Motorin, Y., Marchand, V., Tisné, C., Borland, K., and Kellner, S. 2021. Instrumental analysis of RNA modifications. *Crit. Rev. Biochem. Mol. Biol.* **56**, 178–204.

Zaccara, S., Ries, R.J., and Jaffrey, S.R. 2019. Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 608–624.

Zhang, X., Hao, H., Ma, L., Zhang, Y., Hu, X., Chen, Z., Liu, D., Yuan, J., Hu, Z., and Guan, W. 2021. Methyltransferase-like 3 modulates severe acute respiratory syndrome coronavirus-2 RNA N6-methyladenosine modification and replication. *mBio* **12**, e0106721.