Nanopore Direct RNA Sequencing Reveals Virus-Induced Changes in ^{m6}A Methylation Patterns in Human Cells

10th Annual Research Symposium, 2023 in Stillwater

Dongyu Wang[1] (graduate student), Lee Booth[2], Wenxin Wu[2], Jordan Metcalf[2], Zhimin Liang[3], Guangpu Li[3], Chongle Pan[1], Susan Schroeder[1],

[1]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma, USA,
[2]Section of Pulmonary, Critical Care, and Sleep Medicine, Department of Medicine, University of Oklahoma School of Medicine, Oklahoma City, Oklahoma, USA
[3]Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA
[4]Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma, USA

Introduction:

- RNA modifications are fundamental in regulating several cellular processes and play a critical role in cancer, cell-fate determination, development, and diseases.
- Nanopore direct RNA sequencing provides a powerful method for characterizing the changing transcriptional landscape of viruses and their host.
- Modified nucleotides will display a different signal intensity compared to unmodified nucleotides in direct RNA-Seq, allowing the computational identification of modified sites for each individual RNA molecule.
- ML-based neural network models can take in signal intensity and sequence features to identify potential ^{m6}A sites from direct RNA-Seq data.
- **Hypothesis**: ^{m6}A methylations alter RNA folding and thus potential binding partners in virus-host interactions
- **Proposed approach:** Correlate the ^{m6}A modification with alternative splicing, polyA tail length, RNA folding, gene expression network, and RNA function in virus response.

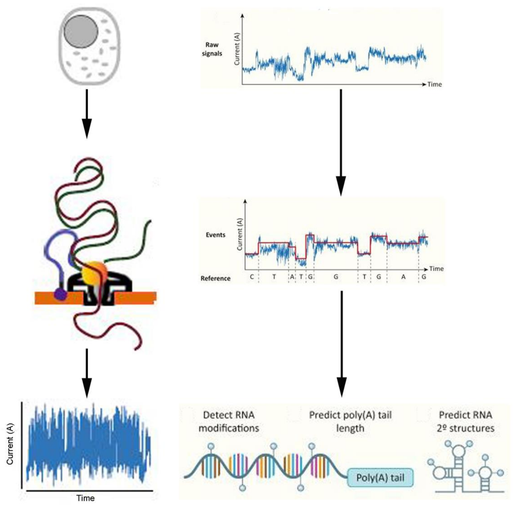

Figure 1. Overview of Nanopore analysis and applications (Figure adapted from Workman et al. 2019 reference 1)

Workflow & Datasets:

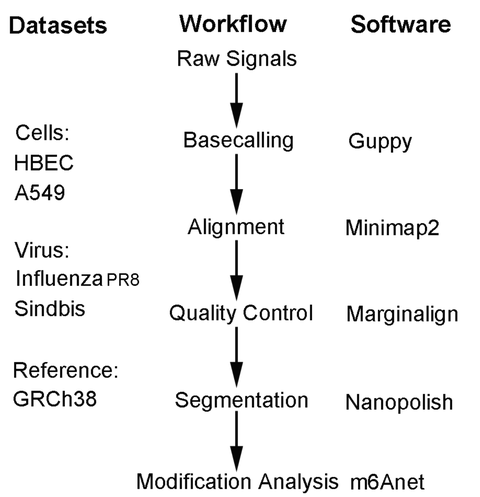

Figure 2. Datasets, software and workflow used in this study

Results:

The identification of differentially methylated RNA sites during viral infection provides new insights into the mechanisms of viral pathogenesis and host defense against viral infections. Our results show that probable ^{m6}A patterns vary with the type of human cell and the virus.

Quality Control: The average length of RNA reads was above 6kb, with high quality data evidenced by **an average identity** and **read accuracy** above **0.95** and **95%**, respectively

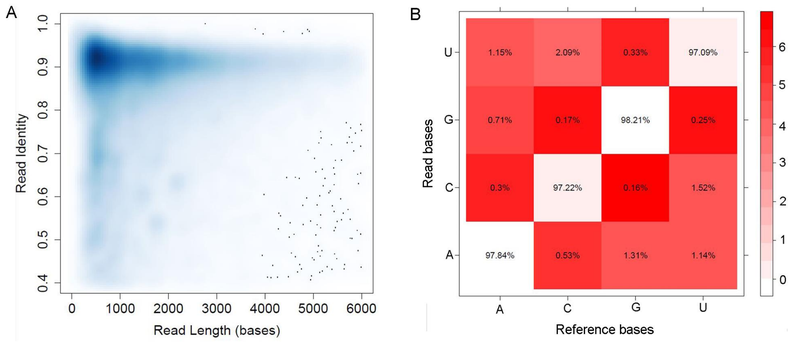

Figure 3. Performance statistics for nanopore native RNA sequencing. A, Alignment identity versus read length for native RNA reads. B, Substitution matrix for native RNA reads.

Poly(A) length: Viral infection changes the overall number of RNA transcripts with long poly(A) tails

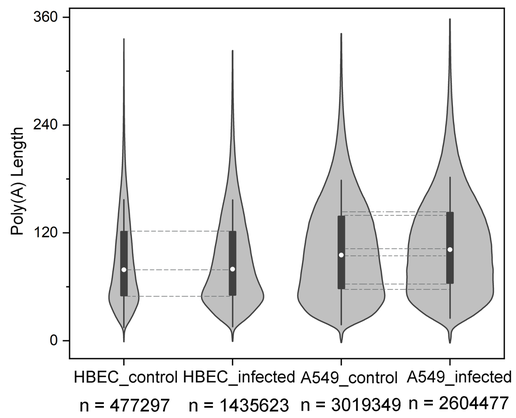

Figure 4. Distribution of poly(A) tail lengths

Modified RNA: The number of ^{m6}A sites ranged from 206,128 to 632,965 across all four datasets.

- Viral infection promotes ^{m6}A methylation of RNA compared to their host cells.
- The number of ^{m6}A varies in different types of RNA, including protein coding, introns, antisense, and lincRNA.

Table1. The overall number of modified and unmodified RNA reads

Samples	Category	Intron	Protein coding	Processed RNA	LincRNA	Antisense
HBEC_control	Modified	5404	199369	0	989	366
	Unmodified	15472	414429	13216	3103	2994
HBEC_infected	Modified	6756	206256	1164	468	141
	Unmodified	14626	457437	10012	3141	2854
A549_control	Modified	19301	559516	4783	2118	1077
	Unmodified	48825	1264085	29073	9873	8296
A549_infected	Modified	20849	602329	6043	2573	1171
	Unmodified	52115	1299676	31222	11448	10225

References:

[1] Workman, R. E., et. al, Nanopore native RNA sequencing of a human poly(A) transcriptome. (2019), Nature Methods volume 16, pages 1297-1305.
[2] Hendra, C., et. Al., Detection of m6A from direct RNA sequencing using a multiple instance learning framework.(2022), Nature Methods volume 19, pages 1590-1598.

Conserved motifs: ^{m6}A conserved motifs vary across different cell types.

- In the **HBEC cells**, the motif **GGACU** was dominant in all types of detected RNA.
- **A549 cells** showed that the **GGACU** motif was dominant in protein-coding RNA, introns, lincRNA, and antisense RNA.
- In processed transcripts in uninfected **A549 cells,** the dominant ^{m6}A motif was **GGACA**, while the number of conserved ^{m6}A motif **GGACU** and **GGACA** in infected cells increased.

HBEC

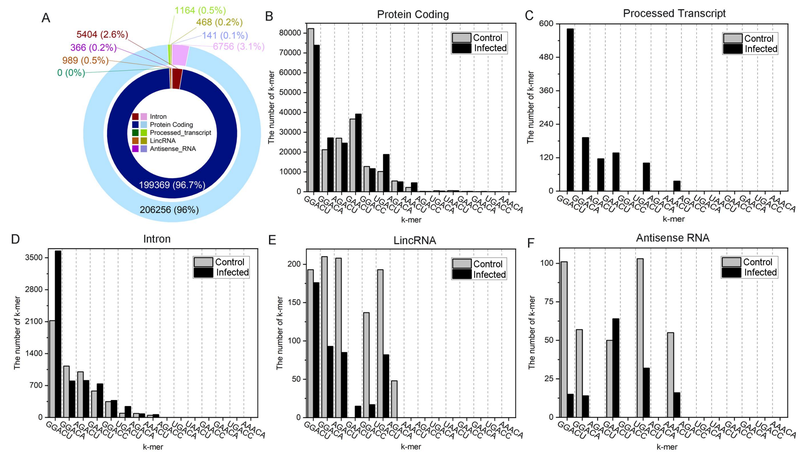

Figure 5. Summary of conserved ^{m6}A motifs in HBEC cells

A549

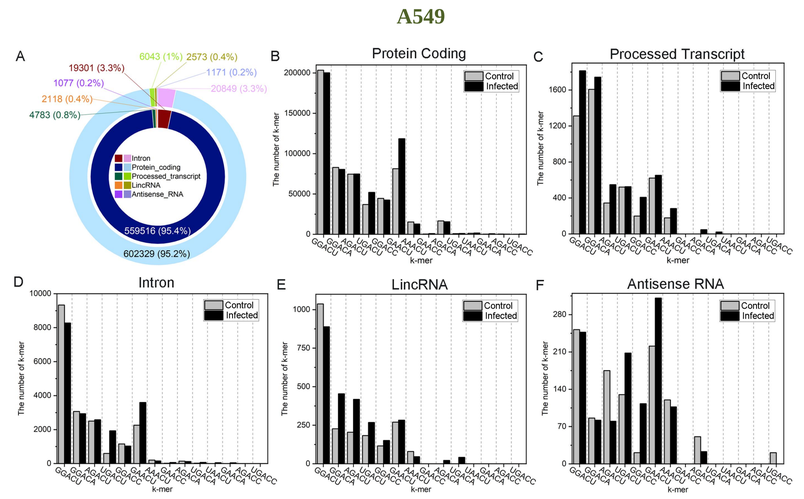

Figure 6. Summary of conserved ^{m6}A motifs in A549 cells

Conclusions & Perspective:

This study using Nanopore direct RNA sequencing revealed differences in ^{m6}A RNA methylation in virus-host interactions. The findings suggest that viral infection promotes ^{m6}A methylation of RNA in host cells, and ^{m6}A patterns vary with the type of human cell and virus. Future research will test the hypothesis that ^{m6}A methylations alter RNA folding and binding partners in virus-host interactions. Future research will investigate the correlations of ^{m6}A modification with alternative splicing, RNA folding, gene expression network, and RNA function in virus response, providing further insights into the mechanisms of virus-host interactions.

Acknowledgment:

This work is supported by grants from Data Institute for Societal Challenges at University of Oklahoma, Presbyterian Health Foundation, Oklahoma Center for Respiratory and Infectious Diseases NIH Center of Biomedical Research Excellence. This work used the computational resources at Oklahoma Supercomputing Center for Education and Research, which is funded by the National Science Foundation.