UNIVERSITY OF OKLAHOMA

GRADUATE COLLAGE

NANOPORE DIRECT RNA SEQUENCING AND PROTEOMICS REVEAL VIRUS-INDUCED CHANGES IN HUMAN CELLS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements of the

Degree of

MASTER OF SCIENCE

BY

Dongyu Wang

Norman, Oklahoma

2024

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/341282

NANOPORE DIRECT RNA SEQUENCING AND PROTEOMICS REVEAL VIRUS-INDUCED CHANGES IN HUMAN CELLS

A THESIS APPROVED FOR THE

GALLOGLY COLLEGE OF ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Chongle Pan, Chair

Dr. Jizhong Zhou

Dr. Marmar Moussa

©Copyright by DONGYU WANG 2024

All Rights Reserved.

Acknowledgement

I would like to express my heartfelt gratitude to everyone who supported me throughout the journey of completing my Master's thesis in Data Science and Analytics. This accomplishment would not have been possible without the invaluable guidance, encouragement, and expertise of many individuals.

First and foremost, I extend my deepest appreciation to my thesis advisory committee, Dr. Chongle Pan, Dr. Jizhong Zhou, and Dr. Marmar Moussa, for their support and guidance throughout my research. Their insights and encouragement were instrumental in helping me shape and refine my work.

I would also like to express my sincere thanks to Dr. Susan Schroeder for her invaluable instruction and mentorship in guiding me through the various projects that formed the foundation of my master's research. Her dedication to my academic growth has left a lasting impact on my development as a researcher.

I am deeply grateful to Dr. Wenxin Wu, Dr. J. Leland Booth, Dr. Jordan P. Metcalf, and Dr. Guangpu Li for providing the data and computational resources essential to this research. Their contributions enabled me to explore and analyze data sets that added depth and real-world relevance to my work.

On a personal note, I owe a profound debt of gratitude to my family and friends for their unwavering love, patience, and encouragement. Their faith in my abilities and constant support provided me with the resilience needed to overcome the challenges of this journey.

Thank you all for your invaluable support.

Table of Contents

Acknowledgement IV

List of Tables VII

List of Figures VIII

Abstract XI

Chapter 1: Introduction 1

1.1 Abstract 1

1.2 Global Health Threat of Influenza Virus 3

1.3 Significance of RNA Modifications in Viral Infections 3

1.4 Technological Advances: Nanopore Direct RNA Sequencing and Proteomics 5

1.5 Virus-Induced ^{m6}A Modifications and Their Implications 6

1.6 Research Objectives and Potential Impact 7

Chapter 2: Nanopore Direct RNA Sequencing Reveals Virus-Induced Changes in the Transcriptional Landscape in Human Bronchial Epithelial Cells 8

2.1 Abstract 8

2.2 Introduction 9

2.3 Methods and materials 12

2.3.1 Cell culture 12

2.3.2 HBEC infection, Virus stock preparation 12

2.3.3 RNA preparation 13

2.3.4 Direct RNA nanopore sequencing data analysis 14

2.4 Results 15

2.4.1 Performance of Nanopore Native RNA Sequencing 15

2.4.2 Virus Binding and Host Response Without Viral Replication 19

2.4.2 ^{m6}A Methylation Patterns and Motif Usage in RNA Transcripts 20

2.4.3 Transcript Abundance and ^{m6}A Methylation as Independent Mechanisms 21

2.4.4 Impact of ^{m6}A Methylation on Alternative Splicing and Isoform Expression 23

2.4.5 Differential ^{m6}A Methylation and Regulatory Roles in Immune lncRNAs 23

2.4.6 Poly(A) Tail Length and Nonpolyadenylated RNA Expression Changes 24

2.4.7 Pseudouridylation and Noncoding RNA Modifications in Response to Influenza 26

2.5 Discussion 28

Chapter 3: Integrated Nanopore Direct RNA Sequencing and Proteomics Uncover Virus-Induced Alterations in Human A549 Cells 31

3.1 Abstract 31

3.2 Introduction 32
3.3 Materials and Methods 34
3.3.1 A549 Cell Cultivation 34
3.3.2 Sindbis Virus Infection 34
3.3.3 RNA Extraction 35
3.3.4 Nanopore Direct RNA Sequencing 36
3.3.5 Basecalling and Quality Control 37
3.3.6 Data Analysis for Nanopore Sequencing 37
3.4 Results 39
3.4.1 Transcriptomic and Proteomic Changes in A549 Cells Following Sindbis Virus Infection 39
3.4.2 Poly(A) Tail Length Differences in SINV-Infected A549 Cells 44
3.4.3 Sequence Motifs of ^{m6}A-Modified Sites in Control and SINV-Infected Conditions 46
3.4.4 Correlation Analysis of Transcript Abundance, Poly(A) Tail Length, ^{m6}A Modifications, and Protein Abundance 49
3.4.6 Functional Enrichment Analysis of Differentially Expressed Genes in SINV-Infected A549 Cells 56
3.5 Discussion 60
Chapter 4 Summary and Outlook 64
4.1 Summary 64
4.2 Outlook 65
Reference 68

List of Tables

Table 2.1 Transcript Isoform Diversity and m6A Modification in Infected and Uninfected Samples. 19
Table 3.1 Summary of Unique Transcripts, m6A Modifications, Coverage, and Unique Proteins in Control and Sindbis Virus-Infected Samples..... 42

List of Figures

Figure 2.1 Nanopore Native RNA Sequencing Performance. The alignment identity across different read lengths and a substitution matrix comparing base calls from nanopore sequencing to the GENCODE v45 reference transcriptome. Longer reads (>5,000 bases) achieve high alignment identity (~95%), while shorter reads (<1,000 bases) have lower identity (~85%). The substitution matrix compares known bases (x-axis) to nanopore-called bases (y-axis). Diagonal cells indicate correct calls, with accuracy ranging from 90% to 96% across bases, highest for guanine (G) and cytosine (C). Off-diagonal cells highlight substitution errors, especially A-to-G and U-to-C, occurring at 5% and 4%, respectively. Color intensity reflects confidence, with darker shades showing greater certainty in matches and mismatches. 17

Figure 2.2 The frequency of each polyadenylated RNA sequence motif containing ^{m6}A at the central position. GGACU was identified as the most common ^{m6}A modification site across all RNA types. Gray bars indicate uninfected control samples, while black bars represent infected samples..... 21

Figure 2.3 Differential gene expressions between uninfected and infected HBEC cells. 22

Figure 2.4 Poly(A) Tail Length Distribution (A) and Differential Gene Expression (B) in HBEC Cells Upon Influenza Infection...... 25

Figure 2.5 Nonpolyadenylated Noncoding RNA Distribution (A) and Differential Expression (B) in PR8-Infected HBEC Cells. Blue dots indicate statistically significant changes, with an adjusted p-value threshold of 0.05, corrected by the Benjamini-Hochberg method..... 26

Figure 2.6 Predicted centroid secondary structures of ENST00000629969.1 without (A) and with (B) ^{m6}A modification at nucleotide 492. Nucleotide colors indicate the probability of being paired or unpaired, with red representing a probability of 1 (fully paired) and purple representing a probability of 0 (unpaired). Structures were generated using Vienna RNAfold. The pink arrow in (B) marks the location of the ^{m6}A modification..... 27

Figure 3.1 Base Substitution Accuracy and Read Identity in Control and Sindbis Virus-Infected Samples. Base substitution accuracy (panels A, B, E, F, I, and J) and

alignment identity as a function of read length (panels C, D, G, H, K, and L) for control and infected samples. In the substitution matrices, rows represent the base called by nanopore sequencing, and columns represent the reference base. Diagonal cells show correct base calls with accuracy rates near 98% for guanine (G) and cytosine (C) across both conditions, while off-diagonal cells highlight substitution errors, particularly A-to-G and U-to-C. The read identity versus read length density plots indicate that alignment identity generally increases with read length. 40

Figure 3.2 RNA and protein expression changes in A549 cells after Sindbis virus (SINV) infection. 43

Figure 3.3 Poly(A) Tail Length Distribution Across RNA Types in Control and SINV-Infected A549 Cells. Bar plots with density distributions show the poly(A) tail lengths for different RNA types—lncRNA, protein-coding RNA, intronic RNA, and snoRNA—in control (blue) and SINV-infected (red) A549 cells. Each bar represents the mean poly(A) tail length, with error bars indicating the standard deviation. Density plots in the background provide the distribution of poly(A) tail lengths within each RNA type. 45

Figure 3.4 Sequence Motifs in ^{m6}A-Modified Sites for Different RNA Types in Control and SINV-Infected A549 Cells. Sequence logos display the enriched motifs surrounding ^{m6}A-modified sites in protein-coding RNA, lncRNA, and intronic RNA under control and SINV-infected conditions. Each logo represents the consensus sequence for ^{m6}A modification sites, with the height of each nucleotide reflecting its information content in bits. The large green "A" denotes the position of the ^{m6}A modification in each motif, indicating the central adenosine where methylation occurs..... 48

Figure 3.5 Correlation Analysis of Transcript Abundance, Poly(A) Tail Length, ^{m6}A Modifications, and Protein Abundance in SINV-Infected A549 Cells. Hexbin plots illustrate the correlations between transcript abundance, poly(A) tail length, ^{m6}A modification count, and protein abundance in SINV-infected A549 cells. Color intensity indicates the density of data points within each bin, with darker shades representing higher data density. Correlation coefficients (r) and p-values are shown for each plot to indicate the strength and significance of each relationship. 52

Figure 3.6 Differential Expression and Modifications in SINV-Infected A549 Cells. 55
Figure 3.7 Functional Enrichment Analysis of Differentially Expressed Genes in SINV-Infected A549 Cells. 59

Abstract

Influenza virus remains a pervasive global health threat due to its rapid mutation and immune evasion capabilities, resulting in frequent epidemics and sporadic pandemics. This thesis investigates the influence of N6-methyladenosine (^{m6}A) RNA modifications on host-pathogen interactions during influenza infection, emphasizing their role in regulating gene expression and host cellular processes. Leveraging nanopore direct RNA sequencing and proteomic analyses, this study offers a comprehensive view of the epitranscriptomic changes in influenza-infected human cells.

The research explores the global distribution of ^{m6}A modifications across host and viral RNA, focusing on their effects on RNA stability, splicing, and translation. Through proteomics integration, we assess how these epitranscriptomic modifications influence protein expression and cellular function, elucidating the contribution of ^{m6}A modifications and ^{m6}A-binding proteins to immune response regulation and viral replication control. The study identifies a dynamic role for influenza-induced ^{m6}A modifications, which can both support antiviral signaling and, conversely, facilitate viral replication. This dual functionality reveals an intricate balancing act between viral manipulation of host pathways and host adaptive responses aimed at containment.

Key findings underscore that influenza infection triggers specific ^{m6}A modifications, enhancing the stability and translation of immune-related transcripts, thereby bolstering antiviral defenses. Additionally, a conserved ^{m6}A consensus motif (GGACU) was observed, indicating a stable targeting mechanism for ^{m6}A methylation across host and viral RNA, a feature maintained even under infection-induced stress. This conservation suggests that while infection modulates ^{m6}A distribution and intensity,

it does not disrupt its fundamental targeting, allowing ^{m6}A regulation of host defenses to persist.

The research highlights the therapeutic potential of targeting ^{m6}A modifications, suggesting that precision interventions within this pathway could provide new strategies for antiviral therapies. By mapping the roles of ^{m6}A in virus-host interactions, this study contributes to the broader understanding of epitranscriptomic regulation in viral infections, with implications for developing antiviral therapies that harness RNA modification pathways.

Keywords: Influenza Virus, ^{m6}A RNA Modifications, Host-Pathogen Interactions, Nanopore Direct RNA Sequencing, Epitranscriptomics, Proteomics

Chapter 1: Introduction

1.1 Abstract

Influenza virus infections present a significant and ongoing threat to public health, characterized by seasonal epidemics and sporadic pandemics. The virus's rapid mutation and immune evasion capabilities exacerbate morbidity and mortality worldwide. This thesis investigates the role of N6-methyladenosine (^{m6}A) modifications in virus-host interactions, specifically focusing on their impact on gene expression and cellular processes during influenza infection. Employing nanopore direct RNA sequencing and proteomic analysis, this study provides a comprehensive examination of the dynamic landscape of ^{m6}A modifications in influenza-infected human cells.

The study aims to map the global ^{m6}A modification patterns in both host and viral RNA, assessing how these modifications influence RNA stability, splicing, and translation during infection. By integrating proteomics, we explore the downstream effects of ^{m6}A modifications on protein expression and function, linking these changes to antiviral defense mechanisms and viral replication strategies. Further, we investigate the role of ^{m6}A reader proteins in mediating these interactions, offering insight into their contribution to immune response modulation.

Our findings suggest that influenza-induced ^{m6}A modifications play a crucial role in orchestrating host responses, either enhancing or inhibiting pathways critical for viral replication and immune signaling. This research contributes to the understanding of epitranscriptomic regulation in viral infections and underscores the potential of targeting ^{m6}A modifications as a novel antiviral therapeutic approach. By elucidating these

molecular interactions, this thesis provides a foundation for the development of targeted antiviral strategies that disrupt virus-host interactions at the RNA modification level.

Key words: Influenza Virus, RNA Modifications, Virus-Host Molecular Interactions, Nanopore Direct RNA Sequencing, Epitranscriptomics, Proteomics

1.2 Global Health Threat of Influenza Virus

Influenza virus infections represent a persistent global health challenge, with the potential to cause seasonal epidemics and unpredictable pandemics [1, 2]. Influenza is associated with significant morbidity and mortality worldwide, particularly affecting vulnerable populations, such as the elderly and immunocompromised individuals [3, 4]. Its rapid spread and capacity for genetic evolution exacerbate these public health threats, as new strains emerge that may evade immunity acquired from previous infections or vaccinations [5]. This ability to adapt and evade immune defenses poses ongoing challenges for both preventative and therapeutic interventions, underscoring the critical need for a deeper understanding of the molecular mechanisms underlying influenza pathogenesis [6, 7].

In this context, exploring virus-host molecular interactions is essential to developing more effective antiviral strategies [8]. By dissecting these interactions, researchers can identify potential molecular targets within host cells that could be modulated to prevent viral replication or enhance the immune response [9]. The present study aims to investigate these virus-host interactions at a detailed molecular level, focusing on the role of RNA modifications in modulating host responses to influenza virus infection. Such research offers a promising pathway toward novel antiviral strategies, particularly as traditional therapeutic approaches struggle to keep pace with the virus's rapid mutation rate [10-12].

1.3 Significance of RNA Modifications in Viral Infections

RNA modifications have emerged as crucial regulators in cellular processes, affecting various aspects of RNA function, including stability, splicing, localization, and

translation [13, 14]. These modifications enable cells to dynamically control gene expression in response to physiological and pathological conditions. Among the numerous RNA modifications, N6-methyladenosine (^{m6}A) has garnered significant attention due to its prevalence in eukaryotic mRNA and its regulatory roles in both cellular and viral contexts [15, 16].

The addition, removal, and interpretation of ^{m6}A marks on RNA are mediated by specific proteins: “writers” like METTL3 and METTL14 add the modification, “erasers” such as FTO and ALKBH5 remove it, and “readers” such as YTHDF1 and YTHDF2 recognize ^{m6}A-modified RNA and mediate downstream processes [17, 18]. In viral infections, ^{m6}A modifications serve as a sophisticated regulatory mechanism. For instance, viruses can hijack host ^{m6}A machinery to enhance their own replication by stabilizing viral RNA [19, 20]. Conversely, ^{m6}A modifications can act as a host defense mechanism by promoting the degradation of viral RNA, thereby restricting viral replication [21, 22]. This dual functionality of ^{m6}A in viral pathogenesis reflects the complexity of virus-host interactions and highlights the potential of targeting ^{m6}A regulatory pathways as a therapeutic strategy [23].

The current study aims to unravel these complex roles of ^{m6}A modifications in influenza infections, focusing on how ^{m6}A influences the molecular environment within host cells [19, 24]. Understanding this interplay can offer insights into the broader epitranscriptomic regulation mechanisms that viruses exploit, providing a foundation for new therapeutic approaches that disrupt these interactions [25-27].

1.4 Technological Advances: Nanopore Direct RNA Sequencing and Proteomics

The advent of nanopore direct RNA sequencing has revolutionized the study of RNA modifications by enabling the sequencing of full-length RNA molecules in their native state, without the need for cDNA conversion or amplification. Unlike traditional sequencing methods that often lose modification information, nanopore sequencing preserves epitranscriptomic data, allowing for the detection of modifications such as ^{m6}A at single-nucleotide resolution [17, 28]. This technology is particularly advantageous for studying virus-host interactions, as it allows simultaneous profiling of both host and viral RNA, offering a comprehensive view of transcriptional and epitranscriptomic changes during viral infection [12, 29, 30].

Nanopore sequencing's ability to detect RNA modifications and RNA processing events, such as alternative splicing and polyadenylation, is crucial for understanding how influenza viruses disrupt host cellular functions [31, 32]. By capturing real-time epitranscriptomic shifts, this technology provides insights into virus-induced changes in RNA structure, stability, and interactions [33, 34]. These insights are essential for deciphering the molecular landscape of infected cells, helping to identify key modifications that facilitate viral replication and immune evasion.

Proteomics serves as a complementary approach to nanopore sequencing, bridging the gap between transcriptomic changes and functional outcomes at the protein level. While transcriptomic analyses reveal changes in RNA expression and modifications, proteomics captures how these changes impact protein abundance, interactions, and activity [35, 36]. For example, ^{m6}A modifications can influence the translation efficiency of certain mRNAs, thereby affecting the levels of corresponding

proteins involved in immune responses or antiviral activities [37]. By integrating nanopore RNA sequencing with proteomics, this study aims to construct a comprehensive model of how influenza-induced ^{m6}A modifications affect cellular functions and contribute to the virus's manipulation of host defenses.

1.5 Virus-Induced ^{m6}A Modifications and Their Implications

This study hypothesizes that influenza-induced ^{m6}A modifications significantly influence the host's transcriptional and translational responses, thereby shaping the outcome of viral infection. Specifically, ^{m6}A modifications on host mRNA can alter RNA stability, splicing, and translation, affecting the expression of genes essential for immune responses, cell cycle regulation, and apoptosis [38, 39]. By manipulating these modifications, influenza viruses may create an environment conducive to viral replication, while simultaneously dampening the host immune response.

Evidence suggests that ^{m6}A modifications can modulate the splicing of viral RNA, enabling the production of alternatively spliced isoforms that may enhance viral replication or evade immune detection [40, 41]. Additionally, ^{m6}A modifications can regulate the expression of immune molecules such as cytokines, potentially attenuating the host's antiviral response [21]. By mapping ^{m6}A modifications in infected cells and analyzing their functional impact, this study seeks to identify the specific roles that ^{m6}A plays in modulating host defenses. These findings could inform future antiviral strategies aimed at targeting the ^{m6}A machinery to disrupt viral replication and enhance immune response.

1.6 Research Objectives and Potential Impact

The primary aim of this research is to uncover the role of ^{m6}A modifications in influenza virus-host interactions, with specific objectives focusing on mapping, functional analysis, and therapeutic implications. Utilizing nanopore direct RNA sequencing, the study will generate a comprehensive map of ^{m6}A modifications in both host and viral RNA, correlating these modifications with changes in gene expression and RNA stability during infection. By integrating RNA sequencing and proteomic data, the research will assess how ^{m6}A modifications impact RNA metabolism and translation efficiency, especially regarding antiviral responses and immune signaling. Additionally, it will investigate the function of ^{m6}A reader proteins in recognizing modified RNA and modulating host defenses through experimental knockdown and overexpression studies, assessing their role in antiviral responses and viral replication. Ultimately, this integrated approach combining transcriptomic and proteomic analyses aims to provide a holistic view of ^{m6}A modifications' influence on protein expression, interactions, and cellular pathways, potentially identifying novel therapeutic targets for antiviral interventions.

The anticipated outcomes of this research are twofold: First, it will advance our understanding of the epitranscriptomic regulation of virus-host interactions, revealing how ^{m6}A modifications orchestrate host responses during influenza infection. Second, it may identify novel targets for therapeutic intervention, as targeting ^{m6}A-regulating proteins or modifying ^{m6}A deposition on specific transcripts could offer new avenues for controlling viral infections and enhancing antiviral therapies.

Chapter 2: Nanopore Direct RNA Sequencing Reveals Virus-Induced Changes in the Transcriptional Landscape in Human Bronchial Epithelial Cells

2.1 Abstract

Using direct RNA nanopore sequencing, we reveal dynamic changes in gene expression, polyadenylation, splicing, ^{m6}A methylation, and pseudouridylation in primary human bronchial epithelial cells following influenza virus exposure. This study details the epitranscriptomic landscape associated with the host immune response. Beyond polyadenylated noncoding RNA, we also purified and sequenced nonpolyadenylated noncoding RNA, identifying significant shifts in expression, N6-methyladenosine (^{m6}A) methylation, and pseudouridylation (Ψ) in these unique RNA types. Notably, two lincRNAs recently linked to immune response, Chaserr and LEADR, showed heightened methylation in response to influenza exposure. Furthermore, certain H/ACA snoRNAs responsible for guiding pseudouridylation exhibited decreased expression, correlating with reduced pseudouridylation in two novel lncRNAs. These findings underscore previously uncharacterized epitranscriptomic changes in epithelial gene networks, offering novel insights into host cellular responses to influenza virus through advanced nanopore sequencing technology.

Key words: Direct RNA Sequencing, Gene Expression, ^{m6}A Methylation, Pseudouridylation (Ψ), Influenza Virus Exposure, Host Immune Response

2.2 Introduction

Seasonal and pandemic influenza pose substantial health risks globally. According to the Centers for Disease Control, influenza has led to 9 to 41 million illnesses, 140,000 to 710,000 hospitalizations, and 12,000 to 52,000 deaths annually between 2010 and 2020 [42, 43]. Viruses are adept at manipulating host cellular resources, and changes in ribonucleic acid (RNA) serve as genetic instructions that modulate host cell responses to viral infections [44]. Modifications in nucleotides can impact RNA splicing, folding, and interactions with proteins, fine-tuning gene regulation in response to viral exposure [45]. The epitranscriptome comprises all nucleotide modifications that respond to cellular stresses and developmental cues [46]. This study explores transcriptomic and epitranscriptomic changes in primary human bronchial epithelial cells (HBECs) exposed to influenza, leveraging direct RNA sequencing from Oxford Nanopore Technology to detect changes in RNA abundance, nucleotide modifications, splicing isoforms, and poly(A) tail length. We also identify RNA methylation changes in both coding and noncoding RNA. Direct RNA nanopore sequencing (DRS), a long-read single-molecule sequencing technology, enables measurement of nucleotide modifications, splicing isoforms, poly(A) tail length, and differential expression within a single assay. This method thus provides a high-resolution view of the dynamic changes in viral and host RNA during influenza A virus (IAV) infection, enhancing our understanding of host immune responses.

RNA adenosine methylation is a key regulatory mechanism in the transcriptional response to viral infection [47]. The most prevalent RNA modification, N6-methyladenosine (^{m6}A), varies in response to numerous viral infections [21]. ^{m6}A

modifications influence RNA splicing, translational efficiency, folding, RNA-protein interactions, and transcript stability, modulating host responses across various gene networks. In response to IAV infection, the host activates a strong interferon (IFN) immune response mediated by IRF7 in bronchial and alveolar epithelial cells [48]. Adenosine methylation negatively regulates the type I interferon response through ^{m6}A modifications of IFNβ mRNA, enhancing its stability and thus prolonging IFN-β production, leading to a robust antiviral response [49]. In herpes simplex virus 1 infections, ^{m6}A methylation prolongs the half-lives of immune response genes like IFIH1, TNFIP3, and IFIT1, while reducing IFIT2 mRNA stability [50]. In Kaposi's sarcoma-associated herpesvirus, ^{m6}A methylation in the 3'UTR of IL-6 transcripts recruits YTHDC2, preventing viral SOX-induced degradation and extending transcript stability [51]. Similarly, ^{m6}A methylation in adenovirus regulates splicing efficiency of late-stage viral genes [17, 52]. In RSV vaccine development, RSV RNA lacking ^{m6}A induces higher IFN responses by increasing RIG-I binding, expression, and signaling [53]. These ^{m6}A modifications either stabilize or destabilize RNA, as seen in STAT1 ^{m6}A modifications that increase transcript stability and PTX3 ^{m6}A methylation in the 3'UTR, which accelerates degradation and represses M2 activation [54]. Studies have documented that ^{m6}A levels globally shift in response to viral infection [55]. Our HBEC data, derived through direct RNA nanopore sequencing, provides full-length, single-molecule transcript reads that reveal epitranscriptomic changes in immune response gene networks.

Traditional studies on RNA methylations during influenza infections have utilized mass spectrometry or variations of Illumina sequencing with ^{m6}A-specific antibodies to

isolate modified RNA fragments [12, 56]. In contrast, direct RNA nanopore sequencing enables long-read detection of RNA methylations in full-length transcripts [57]. This technology uses a motor protein in a nanopore embedded within a synthetic lipid bilayer, allowing single DNA or RNA molecules to traverse the membrane. The resulting voltage in picoamperes, measured over time, varies with the nucleotide sequence in the pore, providing insight into sequence modifications [58, 59]. ^{m6}A modifications affect signal intensity, duration, and base-calling error frequency [60]. Tools like m6Anet, employing a multiple-instance learning framework, predict ^{m6}A modifications within transcripts based on nanopore data, achieving superior accuracy over other methods [61].

Nanopore sequencing is increasingly applied for influenza RNA sequencing, with Oxford Nanopore Technology's ARTIC protocols identifying novel RNA viruses and influenza RNA sequences [62]. Initial observations of ^{m6}A in influenza viral mRNA were made using high-performance liquid chromatography and RNA digestion [52, 63]. Recent nanopore sequencing studies have also used specific adapters to detect viral RNA in cell cultures [64, 65]. Since its application to influenza research, advancements in basecalling software, such as Guppy and Dorado, and ^{m6}A prediction via machine learning tools like m6Anet, have enhanced the capabilities of nanopore DRS. Both host and viral RNA undergo methylation changes in response to viral infections [57, 61, 66]. This study focuses on host immune responses to influenza exposure, examining local and global ^{m6}A methylation variations in coding and noncoding RNA within primary human airway epithelial cells—a model highly relevant to human respiratory infections.

2.3 Methods and materials

2.3.1 Cell culture

Human bronchial epithelial cells (HBECs) were collected from three clinical samples donated by healthy adult non-smokers of varying ages, genders, and ethnic backgrounds. All sample collections followed protocols approved by the Institutional Review Board of the University of Oklahoma Health Sciences Center (IRB #2197). HBECs were obtained through bronchoscopy with bronchial brushing, targeting three to four distinct third- or fourth-order bronchi. Cells were rinsed from the brush into 10 ml of sterile saline, achieving a total count of 5 × 10^6 to 1 × 10^7 viable cells as assessed using hemocytometer counts and trypan blue exclusion.

For isolation, cells were resuspended at 5 × 10^5 cells/ml in complete Bronchial Epithelial Cell Growth Medium (BEGM; Lonza Group Ltd.) and seeded into collagen-coated tissue culture plates (Bio-Coat, BD Biosciences) at a density of 1 × 10^5 cells/cm². Cultures were incubated at 37°C in a 5% CO_2 atmosphere. After 24 hours, cells were washed with HBSS to remove non-adherent cells, and fresh complete BEGM was added. Once cultures approached confluency (typically 7–10 days), the monolayers were detached using 1× Accutase solution and subcultured at a 1:5 ratio. Following each passage, cultures reached confluency within 4–5 days. When splitting the cultures, freezer stocks were prepared in a solution of 80% BEGM, 10% fetal bovine serum, and 10% DMSO and stored in liquid nitrogen vapor at -190°C.

2.3.2 HBEC infection, Virus stock preparation

HBECs were infected with influenza following established protocols. The H1N1 influenza A virus strain, A/PR/34/8 (PR8), was propagated in Madin-Darby canine

kidney (MDCK) cells. Viral propagation occurred in MDCK cells using DMEM/F12 medium supplemented with ITS+ (BD Biosciences, Franklin Lakes, NJ) and trypsin, with the virus harvested 72 hours post-infection and titrated by plaque assay in MDCK cells. The virus was then stored in aliquots at −80 °C. No detectable endotoxin was present in the final viral preparations, as verified by the limulus amebocyte lysate assay (Cambrex, Walkersville, MD) with a sensitivity threshold of 0.1 EU/ml, equivalent to approximately 20 pg/ml LPS. HBEC cultures were infected at a multiplicity of infection (MOI) of 1, and total cellular RNA was collected 24 hours post-infection.

2.3.3 RNA preparation

Total RNA from cells was extracted using Trizol with Zymo columns, incorporating modifications described below. RNA quantity was measured with a Nanodrop spectrometer, and quality was assessed on an Agilent TapeStation. Sample N24 underwent additional ethanol precipitation to remove residual Trizol reagent, using sodium acetate (0.1 volume, pH 5.2) and 2 volumes of ethanol, with incubation at -80 °C, followed by centrifugation at 4 °C and a 70% ethanol wash at 4 °C.

For polyadenylated RNA isolation, 4-50 µg of total cellular RNA was processed with the Lexogen polyA RNA selection kit and eluted into a 12 µL volume. RNA quality and rRNA depletion were verified on an Agilent TapeStation. Nonpolyadenylated RNA enriched for noncoding RNA (ncRNA) was further isolated by purifying the RNA from step 9 of the polyA purification protocol using the Lexogen RiboCop kit, thereby depleting ribosomal RNA. Subsequently, the ncRNA was in vitro polyadenylated with New England Biolabs polyA polymerase from *E. coli* (product # M0276S) through a 90-

minute incubation at 37 °C, followed by another round of purification using the polyA selection kit.

Libraries for nanopore sequencing were prepared according to Oxford Nanopore Technologies' RNA-SQK002 protocols, including the reverse transcription step, with an input range of 180 ng to 2 µg RNA. Sequencing was performed on Oxford Nanopore R9.4 MinION flow cells using the Oxford MinION platform (MinKnow version 22.12.7).

2.3.4 Direct RNA nanopore sequencing data analysis

High-quality basecalling was performed using Oxford Nanopore's Guppy software integrated within Dorado version 0.3.2, applying a minimum quality score threshold of 7 for passing reads [66, 67]. RNA sequences were aligned to the human transcriptome (GENCODE v45) using the Minimap2 algorithm [68]. Analysis of alternative splicing was conducted with FLAIR v1.4 (Full-Length Alternative Isoform Analysis of RNA) [69].

For ^{m6}A methylation analysis, m6Anet software was utilized to identify probable ^{m6}A methylation sites within the RNA sequences, requiring a minimum of 30 copies per transcript for analysis [70]. m6Anet outputs included transcript ID, nucleotide position, ^{m6}A probability, and the percentage of transcripts exhibiting ^{m6}A at each nucleotide site. Only sites with an ^{m6}A probability of 0.90 or higher were considered in subsequent analyses. NanoSPA software was employed to detect both ^{m6}A and pseudouridine modifications, while differential expression analysis was carried out using DESeq2 [71].

2.4 Results

2.4.1 Performance of Nanopore Native RNA Sequencing

The performance of nanopore sequencing on native RNA was evaluated by examining alignment identity across various read lengths and assessing base substitution accuracy against the GENCODE v45 reference transcriptome (Figure 2.1). The analysis of alignment identity revealed a strong positive correlation between read length and alignment fidelity. For reads longer than 5,000 bases, alignment identity reached an average of 95%, indicating that longer sequences contribute to enhanced basecalling accuracy due to improved contextual resolution. In contrast, shorter reads—particularly those under 1,000 bases—displayed a reduced alignment identity, averaging around 85%, which may be attributable to limited sequence context that heightens sensitivity to sequencing noise. These findings underscore the influence of read length on nanopore sequencing accuracy, aligning with previous studies and suggesting that maximizing read length could be beneficial in optimizing basecalling performance in native RNA applications.

The substitution matrix further delineates base-specific accuracy in nanopore sequencing by comparing each known base in the GENCODE reference with the corresponding nanopore-called base. Diagonal cells within the matrix represent correct base calls, with accuracy values ranging from 90% to 96% across bases, indicating robust basecalling for most nucleotide types. The highest accuracy was observed for guanine (G) and cytosine (C) bases, which each achieved correct identification rates of approximately 96%. Adenine (A) and uracil (U) displayed slightly lower accuracy,

averaging around 92% and 90%, respectively, suggesting minor but notable variability in basecalling fidelity across nucleotide types.

Off-diagonal cells indicate substitution errors, with red shading marking the frequency of these errors. The most common substitutions observed were A to G and U to C, occurring with a frequency of 5% and 4%, respectively, suggesting that certain nucleotide transitions are more prone to error in nanopore sequencing. This pattern aligns with known difficulties in distinguishing between purine and pyrimidine bases due to their structural differences in native RNA. The intensity of color in each cell reflects the negative natural logarithm of the probability of basecall matches or mismatches, with darker cells indicating higher certainty in observed outcomes. This color coding reveals a clear distribution, with high-confidence correct basecalls along the diagonal and systematic substitution errors in specific off-diagonal cells.

In summary, these results illustrate the high overall accuracy of nanopore native RNA sequencing, particularly for longer reads, and identify nucleotide-specific substitution patterns that may guide further refinements in sequencing protocols. The accuracy of nanopore basecalling for native RNA is evident, with overall high alignment identity and consistent substitution trends, which provide a foundation for optimizing sequencing accuracy in transcriptomic applications.

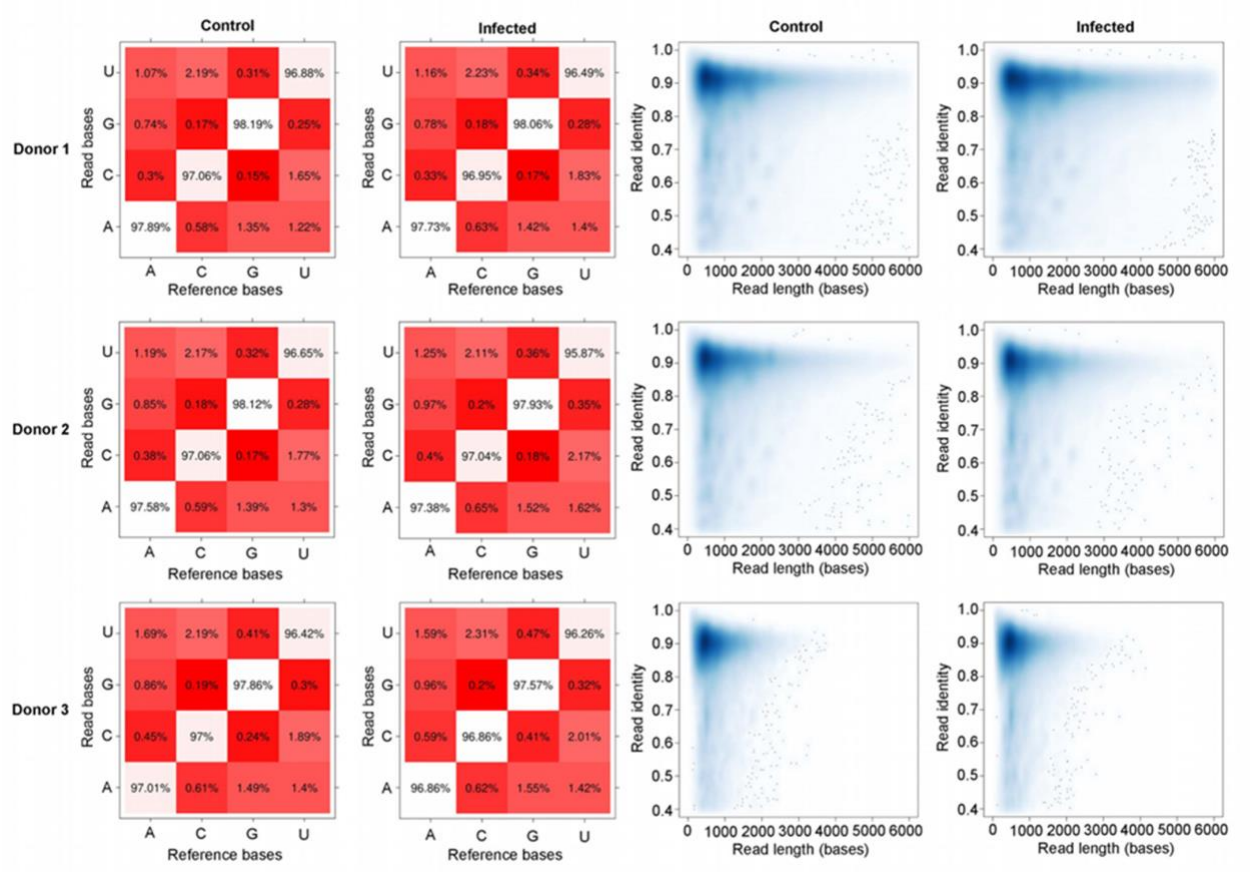

Figure 2.1 Nanopore Native RNA Sequencing Performance. The alignment identity across different read lengths and a substitution matrix comparing base calls from nanopore sequencing to the GENCODE v45 reference transcriptome. Longer reads (>5,000 bases) achieve high alignment identity (~95%), while shorter reads (<1,000 bases) have lower identity (~85%). The substitution matrix compares known bases (x-axis) to nanopore-called bases (y-axis). Diagonal cells indicate correct calls, with accuracy ranging from 90% to 96% across bases, highest for guanine (G) and cytosine (C). Off-diagonal cells highlight substitution errors, especially A-to-G and U-to-C, occurring at 5% and 4%, respectively. Color intensity reflects confidence, with darker shades showing greater certainty in matches and mismatches.

The analysis of transcript isoform diversity and ^{m6}A modifications between infected and uninfected samples reveals significant differences in transcript complexity

and ^{m6}A methylation patterns. In uninfected samples, 2,729 transcripts were identified with more than two isoforms, indicating a high level of transcript diversity in the absence of infection. In contrast, infected samples displayed a lower count, with 1,473 transcripts exhibiting more than two isoforms, suggesting that infection may reduce transcript isoform diversity or alter splicing patterns (Table 2.1).

The presence of ^{m6}A modifications also differed between the two sample conditions. Among the infected transcripts, 562 were identified with ^{m6}A modifications, while the uninfected samples had a lower count of 395 ^{m6}A-modified transcripts. This increase in ^{m6}A modification in the infected condition implies that infection may enhance ^{m6}A methylation activity, potentially as a regulatory response to infection (Table 2.1).

Finally, the intersection of ^{m6}A modifications and isoform diversity shows that only a small subset of transcripts, 7 in infected samples and 10 in uninfected samples, exhibit both ^{m6}A modification and the presence of multiple isoforms. This suggests that while both isoform diversity and ^{m6}A modification are present in the transcriptome, their co-occurrence within the same transcript is relatively rare, particularly in the context of infection. These findings provide insights into how infection may impact transcript processing and ^{m6}A methylation, which are both essential aspects of gene expression regulation (Table 2.1).

Table 2.1 Transcript Isoform Diversity and m6A Modification in Infected and Uninfected Samples.

Samples	Transcripts with more than two isoforms	^{m6}A modified transcripts	^{m6}A and isoforms
Infected	1,473	562	7
Uninfected	2,729	395	10

2.4.2 Virus Binding and Host Response Without Viral Replication

This study employed moderate multiplicity of infection (MOI) and submerged cell culture methods to examine the specific host responses to influenza virus binding without initiating viral replication. By isolating the effects of virus exposure, these conditions facilitate the analysis of cellular responses triggered by the virus binding to the cell surface. This design avoids confounding influences from active viral replication, which can independently affect cellular transcriptional responses. Prior research, including single-cell transcriptomic studies, has demonstrated that even in low MOI conditions, many cells remain uninfected but still exhibit distinct transcriptional responses to viral exposure, reflecting a "bystander effect." These bystander responses reveal that viral binding alone can modulate host cell activity without necessitating viral entry or replication. Unlike single-cell methods that currently lack the resolution to capture nucleotide modifications, our use of nanopore direct RNA sequencing (DRS) provides a unique advantage. DRS allows us to identify and quantify modifications at the single-nucleotide level, offering a high-resolution view into the host response to virus binding, especially in the form of RNA modifications critical for immune signaling and regulation.

2.4.2 ^{m6}A Methylation Patterns and Motif Usage in RNA Transcripts

Figure 2.2 highlights the distribution of ^{m6}A methylation across various RNA types isolated from HBECs, comparing influenza-exposed cells to unexposed controls. Our analysis revealed a clear preference in protein-coding transcripts for ^{m6}A methylation at the GGACA motif, particularly targeting the central adenine. This modification pattern aligns with known ^{m6}A sites and reflects a broader methylation landscape that modulates RNA stability, translation, and localization. Interestingly, ^{m6}A in protein-coding RNAs is often enriched near stop codons and within 3' untranslated regions (UTRs), which may suggest its role in translation efficiency and mRNA decay, contributing to dynamic gene expression changes.

In contrast, noncoding RNAs such as lncRNA and intronic RNA exhibited distinct, selective ^{m6}A motif usage. lncRNAs, for instance, showed strong methylation at GGACA and AGACU motifs, with an increase in AGACU methylation specifically in response to influenza exposure. These motifs may play a role in regulating lncRNA interactions with other RNA molecules or proteins, potentially affecting gene networks involved in antiviral responses. Intronic RNAs, however, displayed no significant shifts in ^{m6}A motif preference between exposed and unexposed conditions, indicating a more stable methylation pattern that may not directly respond to viral exposure. This differential motif usage suggests that ^{m6}A modifications are context-specific, driven by both the type of RNA and external stimuli like viral infection. The unique shift in ^{m6}A motif preference in response to influenza exposure points to a nuanced, cell-specific ^{m6}A regulatory landscape that may facilitate adaptive changes in cellular function during infection.

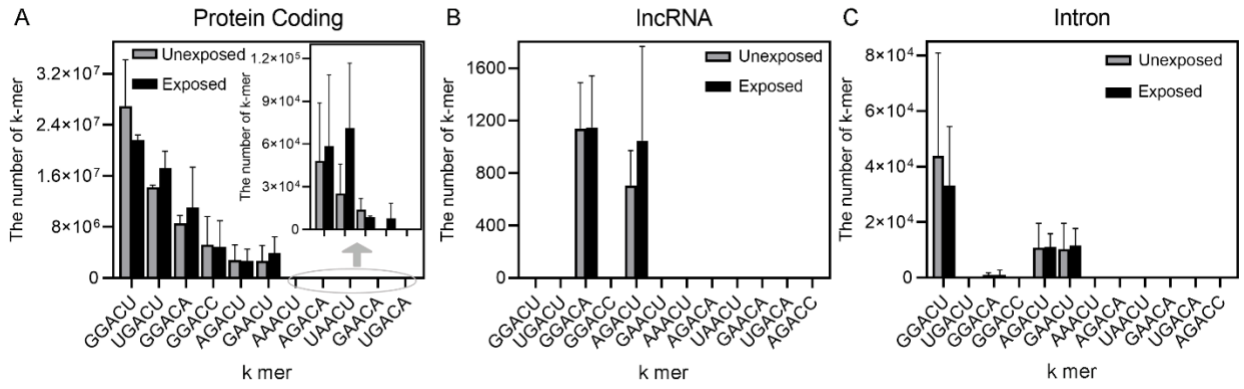

Figure 2.2 The frequency of each polyadenylated RNA sequence motif containing ^{m6}A at the central position. GGACU was identified as the most common ^{m6}A modification site across all RNA types. Gray bars indicate uninfected control samples, while black bars represent infected samples.

2.4.3 Transcript Abundance and ^{m6}A Methylation as Independent Mechanisms

As illustrated in Figure 2.3, most transcripts with altered ^{m6}A levels did not exhibit corresponding changes in abundance, suggesting that ^{m6}A methylation and transcriptional regulation operate through separate mechanisms. This observation implies that ^{m6}A modifications influence gene expression at the post-transcriptional level, potentially by modulating mRNA stability, export, or translation. Studies have shown that ^{m6}A modifications can act as signals for RNA-binding proteins like YTHDF1, which enhance translation efficiency, or YTHDF2, which promotes mRNA degradation.

A few transcripts did show concurrent changes in both ^{m6}A methylation and abundance, providing insights into gene-specific responses. For instance, AHNAK2, a transcript linked to calcium signaling, exhibited increased expression alongside reduced ^{m6}A methylation. This could suggest a regulatory mechanism in which ^{m6}A removal stabilizes AHNAK2 mRNA, thereby promoting its translation during infection. Similarly,

SEMA4B and HSPA5 both showed reduced ^{m6}A levels and decreased abundance. SEMA4B is implicated in cellular proliferation pathways, including tumor progression in lung cancer, while HSPA5, a heat shock protein, supports viral protein stability during the influenza lifecycle by protecting viral RNA-protein interactions. The reduction in HSPA5's ^{m6}A levels could indicate an attempt by the host cell to modulate protein interactions in response to influenza. These results align with findings in other infectious contexts, such as Pseudomonas aeruginosa and HSV-1, where ^{m6}A methylation acts as a flexible, dynamic regulatory mechanism that responds to cellular and environmental changes.

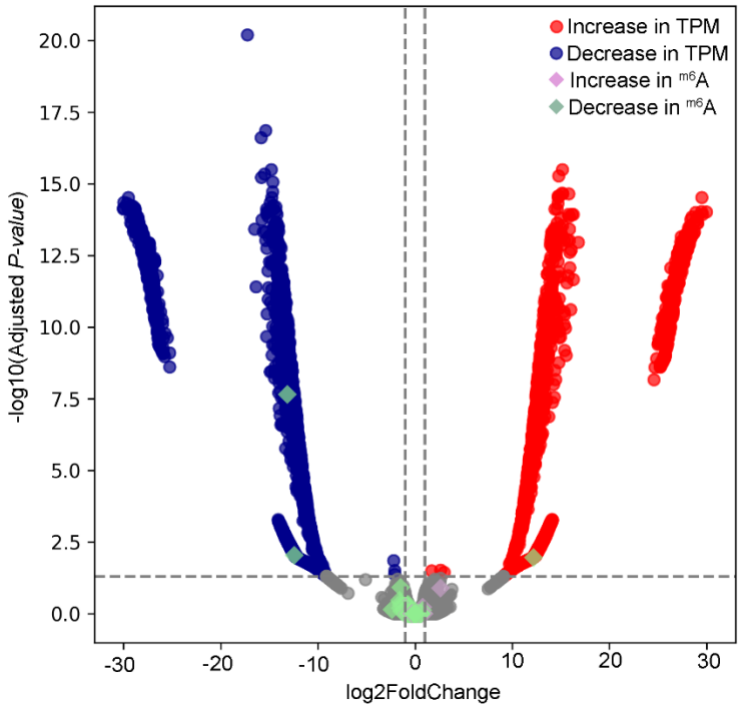

Figure 2.3 Differential gene expressions between uninfected and infected HBEC cells.

2.4.4 Impact of ^{m6}A Methylation on Alternative Splicing and Isoform Expression

Our results revealed minimal correlation between ^{m6}A methylation and alternative splicing changes, consistent with previous studies that examined ^{m6}A's influence on splicing in influenza-infected A549 cells. However, the transcript for Calmodulin 2 (CALM2), a gene with six exons and nine protein isoforms involved in calcium signaling, showed both ^{m6}A methylation and isoform expression changes. CALM2 is known for dynamically regulating cellular responses to stress and disease through its isoforms, and this dual regulation suggests that ^{m6}A methylation may selectively influence splicing patterns in specific genes.

Long-read nanopore sequencing provided additional resolution, enabling the precise identification of isoforms in response to influenza exposure. ISG12, a key interferon-stimulated gene with antiviral properties, showed significant upregulation with a 7.83 log2 fold change. ISG12's eight isoforms display complex alternative splicing, with each isoform contributing distinct functionalities to the immune response. For example, specific isoforms of ISG12 may bind different partners or target different subcellular locations, facilitating a more tailored response to infection. The ability to resolve isoform-specific responses through long-read sequencing highlights the role of alternative splicing as an adaptable mechanism that enriches immune responses.

2.4.5 Differential ^{m6}A Methylation and Regulatory Roles in Immune lncRNAs

Two immune-regulatory lncRNAs, Linc01578 (Chaserr) and Linc00510 (LEADR/MIR205HG), exhibited differential ^{m6}A methylation in response to influenza exposure. Linc01578, methylated at 11 sites, is known to regulate Chromodomain Helicase DNA Binding Protein 2 (CHD2) and participates in a feedback loop with NFκB,

a pathway crucial for immune response and oncogenesis. This feedback loop may help fine-tune immune responses, balancing antiviral activity and cell survival.

Linc00510, methylated at six sites, has dual functionality: it acts as a pri-miRNA for miR-205 and as a regulator of IRF1 transcription factor activity via Alu base pairing. ^{m6}A modifications in these lncRNAs could impact RNA secondary structure, potentially altering molecular interactions with transcription factors, RNA-binding proteins, or other regulatory molecules. In particular, ^{m6}A on Linc00510 may disrupt IRF1 interactions, modulating interferon pathways central to antiviral defenses. These findings suggest that ^{m6}A serves as a molecular switch in lncRNAs, influencing their structure, stability, and function in immune response regulation.

2.4.6 Poly(A) Tail Length and Nonpolyadenylated RNA Expression Changes

Influenza exposure was associated with a reduction in poly(A) tail length across various transcripts (Figure 2.4). Shortened poly(A) tails generally result in decreased mRNA stability and reduced translational efficiency, promoting a faster turnover of mRNA that is adaptive for rapid changes in gene expression. Among the 410 immune-related genes affected, key regulators like NFKB1, IRF3, IL6, and STAT1 exhibited changes in poly(A) tail length. The alteration of poly(A) tails in these transcripts could modulate the intensity and timing of immune responses, enhancing the cell's ability to respond swiftly to infection.

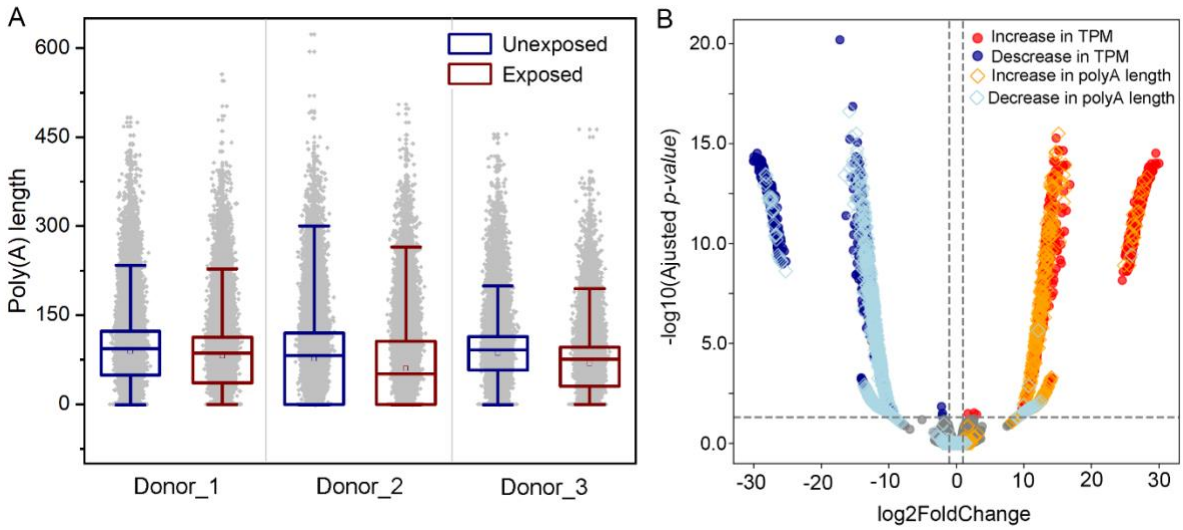

Figure 2.4 Poly(A) Tail Length Distribution (A) and Differential Gene Expression (B) in HBEC Cells Upon Influenza Infection.

Further analysis of nonpolyadenylated RNAs was performed after in vitro polyadenylation and bioinformatic filtering (Figure 2.5). The distribution of nonpolyadenylated RNA types remained consistent across influenza-exposed and unexposed samples (Figure 6), with 7SL RNA as the most abundant transcript. Given that unbound 7SL RNA can serve as a substrate for RIG-I, activating IFN responses, its abundance may enhance antiviral defenses. The stability in nonpolyadenylated RNA profiles, despite changes in poly(A) tail length in other RNAs, suggests that influenza selectively modulates polyadenylation dynamics rather than broadly altering the expression of nonpolyadenylated RNA.

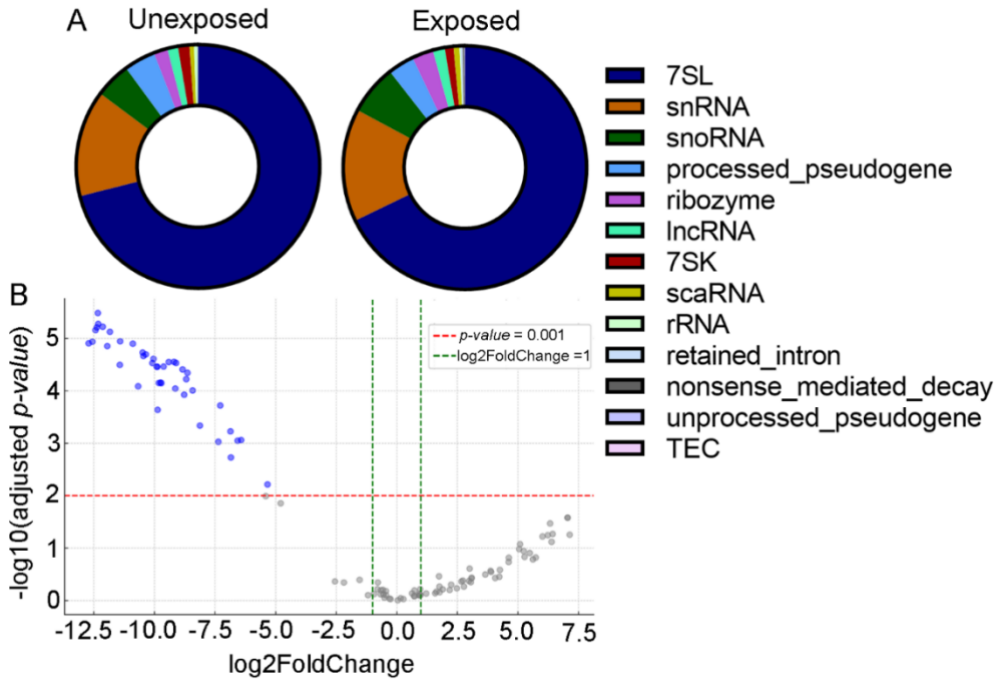

Figure 2.5 Nonpolyadenylated Noncoding RNA Distribution (A) and Differential Expression (B) in PR8-Infected HBEC Cells. Blue dots indicate statistically significant changes, with an adjusted p-value threshold of 0.05, corrected by the Benjamini-Hochberg method.

2.4.7 Pseudouridylation and Noncoding RNA Modifications in Response to Influenza

Consistent with the downregulation of H/ACA snoRNAs, our study found decreased pseudouridylation in specific nonpolyadenylated ncRNAs upon influenza exposure. NanoSPA analysis detected known pseudouridylation sites on 5.8S rRNA, supporting our conservative 90% probability threshold. linc00273 and ENST00000625598.1, two lncRNAs with known pseudouridylation sites, were highly modified in unexposed cells but lost these modifications upon influenza exposure.

Linc00273 is associated with macrophage activation and lung cancer metastasis (64, 65), with pseudouridylation potentially stabilizing its structure for regulatory interactions. The pseudouridylation decrease upon influenza exposure may disrupt these interactions, modifying the transcript's functional role in immune regulation. Additionally, ENST00000629969.1, an antisense lncRNA, exhibited an ^{m6}A modification that destabilized its secondary structure, altering helix stability and disrupting junction formation (Figure 2.6). This structural destabilization could impact RNA-protein interactions, highlighting how ^{m6}A and pseudouridylation modifications cooperatively influence RNA stability and function, enabling precise regulatory responses to viral infections.

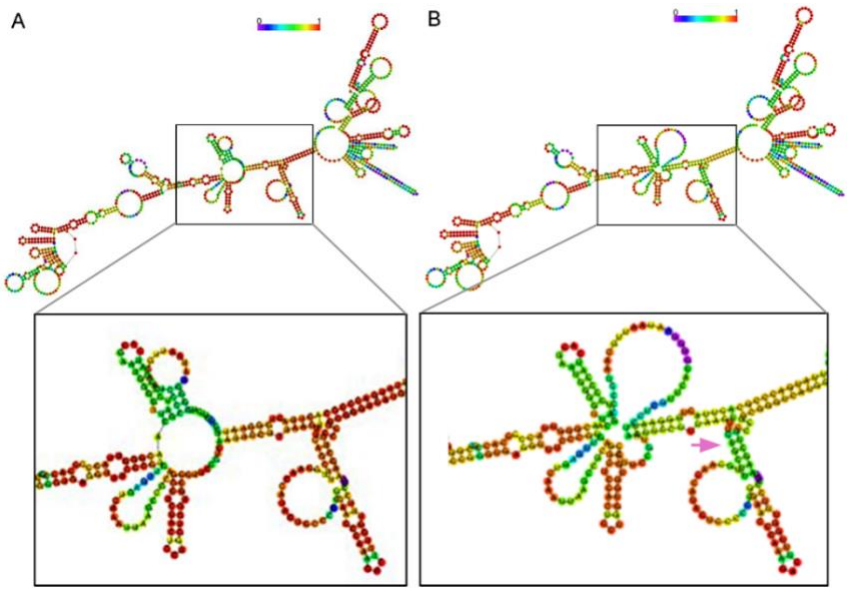

Figure 2.6 Predicted centroid secondary structures of ENST00000629969.1 without (A) and with (B) ^{m6}A modification at nucleotide 492. Nucleotide colors indicate the probability of being paired or unpaired, with red representing a probability of 1 (fully paired) and purple representing a probability of 0 (unpaired). Structures were generated using Vienna RNAfold. The pink arrow in (B) marks the location of the ^{m6}A modification.

2.5 Discussion

This study provides a detailed examination of the transcriptional and epitranscriptomic responses in human bronchial epithelial cells to influenza virus exposure, leveraging direct RNA nanopore sequencing to reveal new layers of complexity in host-virus interactions. The findings underscore the dynamic changes in RNA modifications, specifically ^{m6}A methylation and pseudouridylation, which may play essential roles in modulating the host's immune response and regulating gene expression under viral infection conditions.

The observed reduction in transcript isoform diversity in infected cells suggests that viral exposure may impact RNA processing mechanisms, potentially as part of a host response strategy to limit alternative splicing flexibility and streamline gene expression profiles toward an antiviral state. The decrease in isoform variety, accompanied by an increase in ^{m6}A modifications, highlights a compensatory adjustment where ^{m6}A may provide an alternative mechanism to control RNA stability and translation without reliance on diverse isoforms. This aligns with previous research on ^{m6}A's role in stabilizing immune response transcripts, where methylation is used to fine-tune immune gene expression in response to viral infections [72, 73].

Interestingly, ^{m6}A modification was significantly elevated in infected samples, particularly in immune-related long noncoding RNAs (lncRNAs) such as Chaser and LEADR. This selective methylation suggests a possible function for ^{m6}A in modulating immune lncRNAs, which may influence gene networks crucial for antiviral defense. Such differential methylation in response to viral exposure highlights the adaptive capacity of the epitranscriptome, as methylation can affect RNA structure and protein

interactions, potentially altering lncRNA-mediated regulatory pathways. This finding contributes to the understanding of how RNA modifications in noncoding RNAs support immune function, with implications for therapeutic targeting of immune lncRNAs [74, 75].

The interplay between ^{m6}A methylation and pseudouridylation is another notable aspect of this study. The downregulation of specific H/ACA snoRNAs and the associated decrease in pseudouridylation levels in certain lncRNAs suggest that influenza infection may selectively modulate pseudouridylation pathways. This modification has known effects on RNA structure and stability, and its decrease upon infection may represent a regulatory mechanism to destabilize certain transcripts, potentially preventing overactivation of immune responses or redirecting cellular resources toward antiviral defenses [76, 77]. Together, these observations indicate that the host epitranscriptome dynamically adjusts both ^{m6}A and pseudouridylation levels as part of a coordinated response to viral exposure.

Moreover, the low overlap of ^{m6}A modification with transcripts that possess multiple isoforms indicates that ^{m6}A and alternative splicing may serve largely independent regulatory roles in the context of infection. This suggests that while ^{m6}A primarily influences RNA stability and interaction with translation machinery, alternative splicing likely operates separately, perhaps allowing for the generation of specific isoforms essential for infection response without the interference of methylation [24, 78, 79]. Such compartmentalized regulatory mechanisms could facilitate a more precise and adaptable transcriptional response to pathogens, maintaining cellular function and optimizing immune activation.

This study, by applying direct RNA nanopore sequencing, illustrates the power of single-molecule, long-read technology to elucidate modifications across full-length transcripts, offering a comprehensive view of the RNA modifications that govern host response. This approach has not only confirmed the presence of ^{m6}A and pseudouridylation modifications in response to influenza exposure but also provided a high-resolution mapping of their distribution and co-occurrence within the transcriptome. Such insights highlight the value of nanopore sequencing for studying viral-host interactions, paving the way for further research on the role of RNA modifications in immune regulation and their potential as therapeutic targets for modulating host responses to viral infections.

In conclusion, this study expands the understanding of how influenza infection alters the host epitranscriptome, particularly through ^{m6}A methylation and pseudouridylation. These modifications appear to be critical regulators of immune-related transcripts, contributing to an adaptive and efficient immune response. Future research could explore how these modifications interact with other regulatory layers, such as RNA-binding proteins and translation factors, to achieve a coordinated response to viral infections, providing potential pathways for therapeutic intervention in respiratory diseases.

Chapter 3: Integrated Nanopore Direct RNA Sequencing and Proteomics Uncover Virus-Induced Alterations in Human A549 Cells

3.1 Abstract

Sindbis virus (SINV), an arthropod-borne alphavirus, poses a significant health risk by inducing febrile illness, arthritis, and rash in humans. This study investigates the host cell response to SINV infection using A549 human lung epithelial cells, with a focus on transcriptomic, epitranscriptomic, and proteomic changes. Employing nanopore direct RNA sequencing and proteomic analysis, we examined the landscape of ^{m6}A RNA modifications, alternative splicing, poly(A) tail dynamics, and protein expression to uncover mechanisms underlying host-pathogen interactions. Our findings reveal that SINV infection triggers a selective upregulation of immune-related transcripts and proteins, with a concurrent repression of non-essential cellular maintenance functions, suggesting a strategic reallocation of host resources to bolster immune defenses. We observed an increase in ^{m6}A modifications in immune-related transcripts, supporting enhanced RNA stability and translation, alongside a reduction in poly(A) tail lengths in lncRNAs and protein-coding RNAs, likely facilitating the rapid turnover of these transcripts. Conservation of the GGACU ^{m6}A motif across infected and control samples suggests that SINV does not disrupt host ^{m6}A machinery, allowing the host to retain essential post-transcriptional controls even during infection. Together, these results highlight a sophisticated, multi-layered host response to SINV infection, underscoring the regulatory roles of ^{m6}A methylation and poly(A) tail modulation in antiviral defenses. This study provides insights into the complex regulatory networks involved in alphavirus infections and suggests potential molecular targets for therapeutic intervention.

3.2 Introduction

Sindbis virus (SINV), a positive-sense single-stranded RNA virus within the *Togaviridae* family and *Alphavirus* genus, is a globally distributed arthropod-borne pathogen primarily transmitted by *Culex* and occasionally *Aedes* mosquitoes [80-82]. SINV is associated with outbreaks of febrile illness, arthralgia, and rash in humans, and in some cases, persistent arthritis, underscoring its potential impact on public health across endemic regions of northern Europe, Africa, and parts of Asia and Australia [83]. In host cells, SINV is known for its rapid induction of cytopathic effects (CPE), often within 24 hours of infection, mediated primarily through its non-structural proteins, nsP2 and nsP3 [84]. These proteins facilitate viral replication and contribute to transcriptional and translational shutoff in the host cell, thereby triggering cellular death and limiting antiviral defenses. As a model system, SINV has been extensively utilized to investigate alphavirus biology, especially mechanisms underlying host-virus interactions and the modulation of cellular processes during infection [80, 85].

Host cells respond to viral infection with complex regulatory changes at both transcriptional and post-transcriptional levels [86-88]. Among these regulatory mechanisms, N6-methyladenosine (^{m6}A) has emerged as a critical post-transcriptional RNA modification that modulates mRNA stability, splicing, and translation, and is implicated in various immune responses [89, 90]. Recent studies highlight the role of ^{m6}A in host-pathogen interactions, where it can influence both viral replication and host immune regulation [91, 92]. Viral RNAs, including those of SINV, often carry ^{m6}A modifications that enhance RNA stability and facilitate viral protein production. In contrast, ^{m6}A modifications on host immune transcripts can either bolster or attenuate

the antiviral response, depending on the context, by influencing transcript stability, translation efficiency, or RNA-protein interactions. Understanding the precise dynamics of ^{m6}A during SINV infection is thus essential for elucidating the regulatory mechanisms that underpin host-pathogen interactions [93, 94].

Traditional RNA sequencing approaches, including mass spectrometry and short-read sequencing, offer limited insight into the full complexity of RNA modifications, polyadenylation, and isoform diversity [95, 96]. Oxford Nanopore Technology's direct RNA sequencing (DRS) method has transformed this landscape by enabling full-length transcript analysis in a single assay, including the direct identification of ^{m6}A modifications, poly(A) tail length, and alternative splicing. This high-resolution technique provides unprecedented access to both global and localized RNA modification landscapes, facilitating a more comprehensive understanding of the transcriptional and epitranscriptomic changes induced by viral infection [97-99].

In this study, we employ an integrative approach combining nanopore direct RNA sequencing and proteomic analysis to investigate the transcriptomic and epitranscriptomic response of A549 human lung epithelial cells to SINV infection. Our objectives are to map the landscape of ^{m6}A modifications, alternative splicing, and poly(A) tail changes in both host and viral transcripts and to correlate these findings with proteomic data to elucidate the post-transcriptional and post-translational modifications induced by SINV. This study aims to provide insights into how RNA modifications contribute to the host's immune response and to viral replication strategies, revealing potential molecular targets for antiviral intervention and advancing our understanding of the complex regulatory networks involved in alphavirus infections.

3.3 Materials and Methods

3.3.1 A549 Cell Cultivation

A549 human lung carcinoma epithelial cells were cultured as a model for respiratory viral infection studies. Cells were obtained from [source or supplier], verified for mycoplasma contamination, and maintained under standard culture conditions. Cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (Pen-Strep) to prevent bacterial contamination. Cultures were incubated at 37°C in a humidified atmosphere with 5% CO_2. A549 cells were passaged at 70–80% confluency using 0.25% trypsin-EDTA solution and reseeded at a 1:5 ratio.

To ensure consistency across experimental conditions, cells were seeded at a density of approximately 1×10^5 cells/cm² in 6-well plates 24 hours prior to infection or RNA extraction procedures. Media were refreshed every 48 hours until cells reached the desired confluency for infection studies or RNA extraction.

3.3.2 Sindbis Virus Infection

Sindbis virus (SINV) was propagated in Madin-Darby Canine Kidney (MDCK) cells and stored in aliquots at -80°C. For infection, A549 cells were seeded in 6-well plates and allowed to reach 70–80% confluency. SINV stocks were thawed on ice and diluted in serum-free DMEM to achieve a multiplicity of infection (MOI) of 1.0. Prior to infection, the culture medium was removed from A549 cells, and cells were washed once with phosphate-buffered saline (PBS) to remove residual FBS that could interfere with viral binding.

The diluted SINV solution was then added to the cells, ensuring even distribution across each well. Plates were incubated at 37°C with 5% CO_2 for 1 hour to allow viral adsorption. Following this incubation, the inoculum was removed, cells were washed with PBS to remove unbound virus, and fresh DMEM containing 2% FBS was added to each well. Infected cells were then incubated at 37°C with 5% CO_2 for 24 hours.

3.3.3 RNA Extraction

At 24 hours post-infection, total RNA was extracted from both infected and uninfected A549 cells using the TRIzol reagent (Thermo Fisher Scientific), following the manufacturer's protocol with minor modifications for optimized RNA purity and yield. Cells were washed with cold PBS and lysed directly in 1 mL of TRIzol reagent per well. Lysates were transferred to RNase-free tubes and incubated at room temperature for 5 minutes to allow complete dissociation of nucleoprotein complexes.

Following lysis, 200 µL of chloroform (Sigma-Aldrich) was added per 1 mL of TRIzol, and samples were vigorously shaken by hand for 15 seconds, then incubated at room temperature for 2–3 minutes. Phase separation was achieved by centrifugation at 12,000 × g for 15 minutes at 4°C. The aqueous phase containing RNA was carefully transferred to a new RNase-free tube, and RNA was precipitated by adding 500 µL of isopropanol (Sigma-Aldrich) and incubating for 10 minutes at room temperature.

The RNA pellet was collected by centrifugation at 12,000 × g for 10 minutes at 4°C, washed with 1 mL of 75% ethanol, and air-dried briefly. Pellets were resuspended in RNase-free water and incubated for 10 minutes at 55–60°C to facilitate complete dissolution. RNA purity and concentration were measured using a NanoDrop

spectrophotometer (Thermo Fisher Scientific), with an A260/A280 ratio of ~2.0 indicating high-quality RNA. Integrity was further assessed on an Agilent Bioanalyzer (Agilent Technologies) before proceeding to downstream applications.

RNA was stored at -80°C until used for nanopore sequencing or additional analyses.

3.3.4 Nanopore Direct RNA Sequencing

To prepare libraries for nanopore sequencing, total RNA isolated from A549 cells was first processed to enrich for polyadenylated and non-polyadenylated RNA. RNA libraries were prepared using the Oxford Nanopore Technologies RNA-SQK002 kit, following the manufacturer's protocol with modifications to optimize yield and quality. A total of 180 ng to 2 µg of RNA was used as input per reaction, depending on the sample's RNA concentration and quality, as measured by NanoDrop and Agilent TapeStation. The library preparation protocol included an initial reverse transcription step to stabilize RNA and enhance sequencing efficiency, which involved annealing the oligo(dT) primer and synthesizing complementary DNA (cDNA) using reverse transcriptase. The resulting RNA-cDNA hybrids were purified and prepared for sequencing adapter ligation.

Adapters provided in the RNA-SQK002 kit were ligated to the RNA-cDNA hybrids, allowing RNA to be directly sequenced while minimizing sample degradation. Following ligation, the libraries were purified to remove unligated adapters and other contaminants. Libraries were then loaded onto Oxford Nanopore R9.4 MinION flow cells, and sequencing was carried out on the MinION platform, controlled by MinKnow

software version 22.12.7. Sequencing runs were monitored to ensure data quality, targeting an output of several million reads per sample.

3.3.5 Basecalling and Quality Control

High-quality basecalling was performed using Oxford Nanopore's Guppy software, which was integrated within Dorado version 0.3.2 [66, 100]. Reads were filtered by applying a minimum quality score threshold of 7, ensuring that only high-confidence reads were retained for downstream analysis. Basecalled reads were assessed for quality metrics, such as read length distribution and mean quality scores, to confirm sequencing performance across all samples.

3.3.6 Data Analysis for Nanopore Sequencing

The basecalled RNA sequences were aligned to the human transcriptome (GENCODE v45) using the Minimap2 algorithm with optimized parameters for long-read RNA data [68]. Minimap2 settings included specific options for spliced alignment to accurately map full-length transcripts and capture alternative splicing events.

Alternative splicing analysis was conducted using FLAIR v1.4 (Full-Length Alternative Isoform Analysis of RNA) to identify novel splicing events and quantify known isoforms [101]. This tool was employed to generate high-confidence isoform annotations, comparing infected and uninfected samples to reveal infection-induced splicing changes.

For methylation analysis, we utilized m6Anet software to identify probable ^{m6}A methylation sites within the RNA sequences [61]. Analysis was restricted to sites with at least 30 read coverage per transcript, ensuring statistical reliability. Only ^{m6}A sites with a

probability score of 0.90 or higher were considered in further analysis, highlighting the most confident methylation sites. The NanoSPA software was additionally used to detect both ^{m6}A and pseudouridine modifications, allowing for comprehensive profiling of key RNA modifications [102].

Differential expression analysis was conducted using DESeq2 to compare infected versus uninfected samples. DESeq2 enabled normalization and statistical testing across conditions, identifying significant shifts in transcript abundance and modification patterns in response to SINV infection [103].

3.4 Results

3.4.1 Transcriptomic and Proteomic Changes in A549 Cells Following Sindbis Virus Infection

Nanopore sequencing and proteomic analysis revealed significant transcriptomic and proteomic alterations in A549 cells following infection with Sindbis virus (SINV). Substitution matrices and read identity versus read length plots confirmed high basecalling accuracy across control and infected conditions (Figure 3.1). Correct basecalls are prominently represented along the diagonal in the substitution matrices, with guanine (G) and cytosine (C) exhibiting nearly 98% accuracy across all samples, while adenine (A) and uracil (U) demonstrated slightly lower accuracy, around 96–97%. Off-diagonal substitution errors, notably A-to-G and U-to-C, were observed in both conditions, potentially indicating sequencing biases or RNA modifications, such as ^{m6}A methylation, that impact base recognition. Read identity plots showed that alignment identity generally increased with read length, with reads longer than 2,000 bases achieving alignment identities above 90%. Shorter reads displayed more variability, with identity scores ranging from 40% to 90%, indicating that longer reads provide more contextual information, essential for accurate transcriptome profiling.

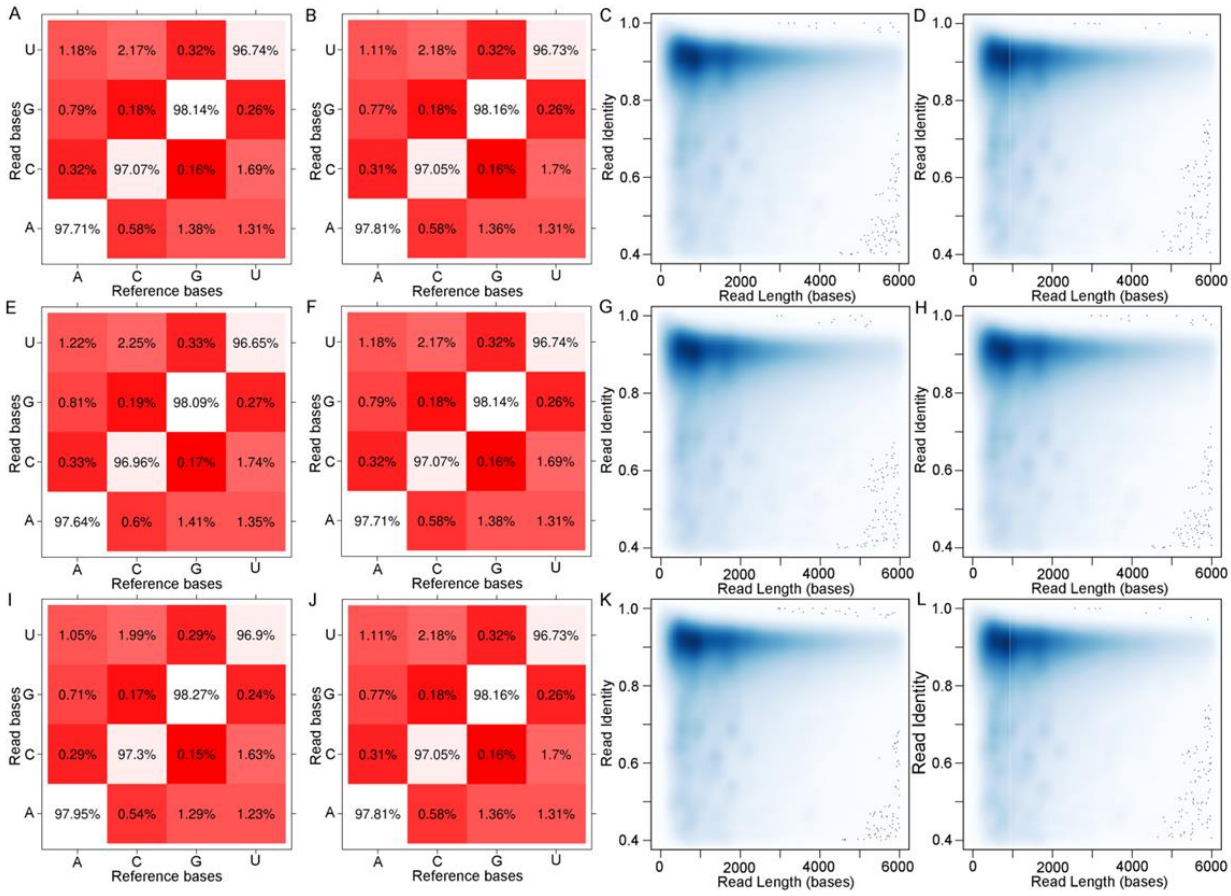

Figure 3.1 Base Substitution Accuracy and Read Identity in Control and Sindbis Virus-Infected Samples. Base substitution accuracy (panels A, B, E, F, I, and J) and alignment identity as a function of read length (panels C, D, G, H, K, and L) for control and infected samples. In the substitution matrices, rows represent the base called by nanopore sequencing, and columns represent the reference base. Diagonal cells show correct base calls with accuracy rates near 98% for guanine (G) and cytosine (C) across both conditions, while off-diagonal cells highlight substitution errors, particularly A-to-G and U-to-C. The read identity versus read length density plots indicate that alignment identity generally increases with read length.

Infected samples consistently exhibited a higher number of unique transcripts than controls, with counts ranging from 2,163,091 to 2,319,885, compared to the

broader range observed in control samples (1,442,161 to 2,650,104). This pattern suggests a focused transcriptional response in SINV-infected cells, likely due to activation of specific pathways related to immune response and antiviral defense. An increase in ^{m6}A-modified transcripts was also observed in infected samples, with counts between 3,957 and 5,458 compared to 2,565 to 4,292 in controls. Infected Rep_3 showed the highest level of ^{m6}A modification, with 5,458 ^{m6}A-modified transcripts, suggesting that SINV infection may induce ^{m6}A methylation as a regulatory mechanism to stabilize or modulate the translation of immune-related transcripts (Table 3.1).

Proteomic analysis showed a similar trend, with infected samples displaying an increased number of unique proteins (Table 3.1). Control replicates had unique protein counts ranging from 1,573 to 1,637, while infected replicates ranged from 1,602 to 1,793, with infected Rep_1 exhibiting the highest number. This increase likely corresponds to the activation of proteins related to immune defense, stress responses, and other pathways critical for antiviral activity. Together, these transcriptomic and proteomic findings indicate that SINV infection not only activates transcription of immune-related genes but also translates these changes into increased protein levels, reflecting a coordinated response at both the transcriptional and translational levels.

Table 3.1 Summary of Unique Transcripts, m6A Modifications, Coverage, and Unique Proteins in Control and Sindbis Virus-Infected Samples.

Methods	Categories	Control			Infected		
		Rep_1	Rep_2	Rep_3	Rep_1	Rep_2	Rep_3
Nanopore	Unique transcripts	2163091	1442161	2650104	2319885	2163091	2319885
	^{m6}A	3911	2565	4292	4310	3957	5458
	Coverage	16	22	12	18	18	16
Proteomics	Unique proteins	1637	1573	1595	1793	1623	1602

The differential expression analysis illustrates a clear polarization of gene expression changes in response to SINV infection (Figure 3.2). Upregulated RNAs (red circles) cluster significantly, likely representing genes associated with immune activation and antiviral defense, as the host mobilizes resources to combat viral replication. On the left side, downregulated RNAs (blue circles) are prominently clustered, suggesting a suppression of cellular processes that may be non-essential or even counterproductive during infection, perhaps as a strategic adaptation by the host to conserve resources or inhibit viral exploitation.

Proteins with significant expression changes align with these RNA-level alterations (Figure 3.1). Upregulated proteins and downregulated proteins show a correlation with RNA expression changes, indicating that transcriptional modifications induced by SINV infection are often reflected at the protein level. This consistency underscores the functional impact of gene expression changes, where transcriptional activation of immune pathways and repression of other functions are efficiently translated into protein-level adjustments to support an effective antiviral response.

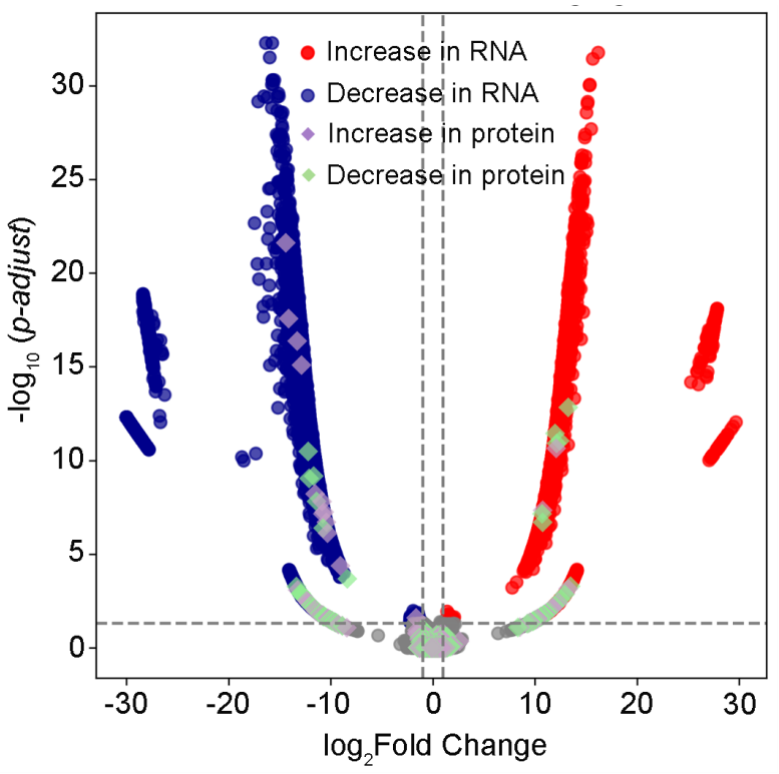

Figure 3.2 RNA and protein expression changes in A549 cells after Sindbis virus (SINV) infection.

In summary, SINV infection induces a coordinated molecular response in A549 cells, characterized by heightened transcriptional activity, increased ^{m6}A methylation, and elevated protein expression. The high accuracy of nanopore sequencing supports these findings, highlighting transcriptional and post-transcriptional modifications as integral components of the host's antiviral defense. These transcriptomic and proteomic shifts provide a comprehensive view of the cellular adaptations occurring in response to viral infection, offering insights into regulatory mechanisms that may be targeted for therapeutic intervention.

3.4.2 Poly(A) Tail Length Differences in SINV-Infected A549 Cells

Analysis of poly(A) tail lengths revealed significant differences in specific RNA categories between control and Sindbis virus (SINV)-infected A549 cells. Poly(A) tail lengths were measured for four RNA types: long non-coding RNA (lncRNA), protein-coding RNA, intronic RNA, and small nucleolar RNA (snoRNA) (Figure 3.3).

For both lncRNAs and protein-coding RNAs, infected samples showed a significant reduction in poly(A) tail length compared to controls. In lncRNAs, the mean poly(A) tail length decreased from approximately 240 bases in control cells to around 120 bases in infected cells, with a highly significant p-value (< 0.0001). Similarly, protein-coding RNAs exhibited a decrease in poly(A) tail length from about 180 bases in controls to 100 bases in infected samples, also with a p-value of < 0.0001. These reductions in poly(A) tail length may suggest a viral influence on RNA stability or translational control mechanisms specific to these RNA types, potentially facilitating rapid turnover of host transcripts to favor viral replication.

In contrast, intronic RNAs did not show a statistically significant change in poly(A) tail length between control and infected samples (p-value = 0.18). This stability in poly(A) length suggests that intronic RNAs might not be directly impacted by viral mechanisms that modify tail length, or their turnover may not be as tightly regulated in response to SINV infection.

For snoRNAs, no significant difference in poly(A) tail length was observed between the two conditions (p-value = 0.85), indicating that SINV infection does not appear to affect the stability or processing of snoRNAs, which typically have short or

absent poly(A) tails and are involved in RNA modification rather than coding or regulatory functions.

Overall, these findings suggest that SINV infection selectively alters poly(A) tail length in specific RNA types, particularly lncRNAs and protein-coding RNAs, likely reflecting a targeted viral strategy to modulate host gene expression at the post-transcriptional level. The reduction in poly(A) tail length in these categories may facilitate more rapid degradation, enabling the virus to efficiently manipulate the host cellular environment to support viral replication.

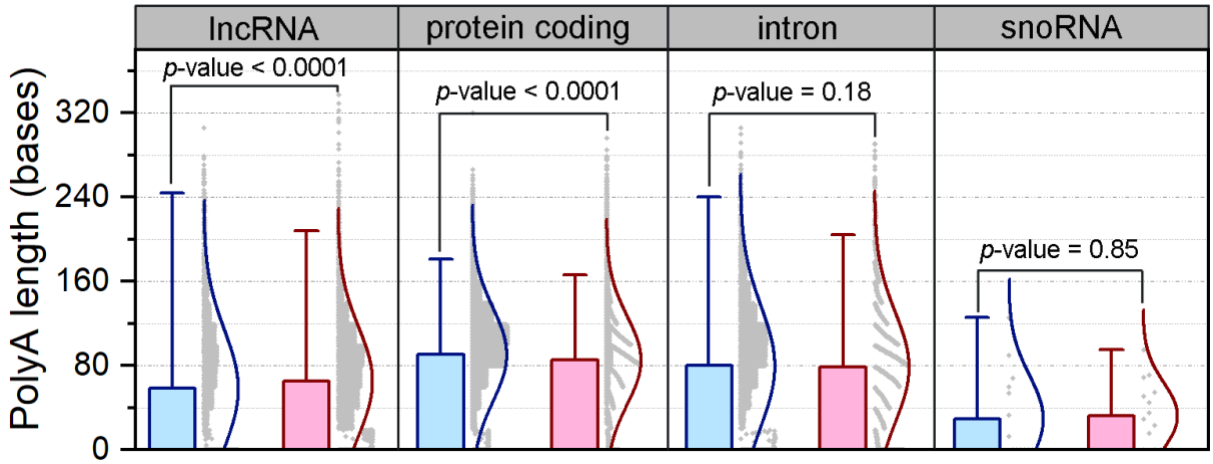

Figure 3.3 Poly(A) Tail Length Distribution Across RNA Types in Control and SINV-Infected A549 Cells. Bar plots with density distributions show the poly(A) tail lengths for different RNA types—lncRNA, protein-coding RNA, intronic RNA, and snoRNA—in control (blue) and SINV-infected (red) A549 cells. Each bar represents the mean poly(A) tail length, with error bars indicating the standard deviation. Density plots in the background provide the distribution of poly(A) tail lengths within each RNA type.

3.4.3 Sequence Motifs of ^{m6}A-Modified Sites in Control and SINV-Infected Conditions

The analysis of sequence motifs surrounding ^{m6}A-modified sites in different RNA types—protein-coding RNA, lncRNA, and intronic RNA—revealed a conserved pattern across both control and SINV-infected A549 cells. Sequence logos for each RNA category show a strong enrichment of the GGACU motif, with a prominent ^{m6}A-modified adenosine (A) in the central position, marked by its large representation in the sequence logos. This motif was consistently observed in both control and infected conditions, indicating that SINV infection does not alter the sequence specificity of ^{m6}A modification (Figure 3.4).

In protein-coding RNA, the GGACU motif appears with high information content, suggesting a strong preference for ^{m6}A modifications at this site. The presence of this conserved motif in both conditions implies that ^{m6}A methylation machinery targets similar sites regardless of infection status. The stability of this motif in protein-coding RNA aligns with the role of ^{m6}A in regulating the stability, splicing, and translation of key transcripts involved in cellular response mechanisms.

A similar motif pattern was observed in lncRNA, where GGACU remained the primary sequence context for ^{m6}A modification in both control and infected samples. This conservation suggests that ^{m6}A modifications in lncRNA are likely important for maintaining specific regulatory functions, potentially influencing lncRNA interactions and stability in a manner independent of viral infection.

For intronic RNA, the same GGACU motif was again prominent across both conditions, though with slightly lower information content compared to protein-coding RNA and lncRNA. This pattern indicates a conserved preference for ^{m6}A sites within intronic regions, suggesting potential roles in pre-mRNA processing or in modulating splicing decisions. The presence of ^{m6}A modifications in introns may facilitate rapid adjustments to gene expression in response to cellular or viral cues, although the data suggest that SINV infection does not specifically alter these sequence preferences.

Overall, the conservation of the GGACU motif across RNA types and conditions implies that SINV infection does not disrupt the ^{m6}A methylation machinery's sequence targeting. This conserved sequence specificity indicates that while SINV infection may affect the frequency or distribution of ^{m6}A modifications, the underlying sequence context for ^{m6}A remains stable, pointing to a fundamental role for this modification in regulating RNA functions both in normal and infected states.

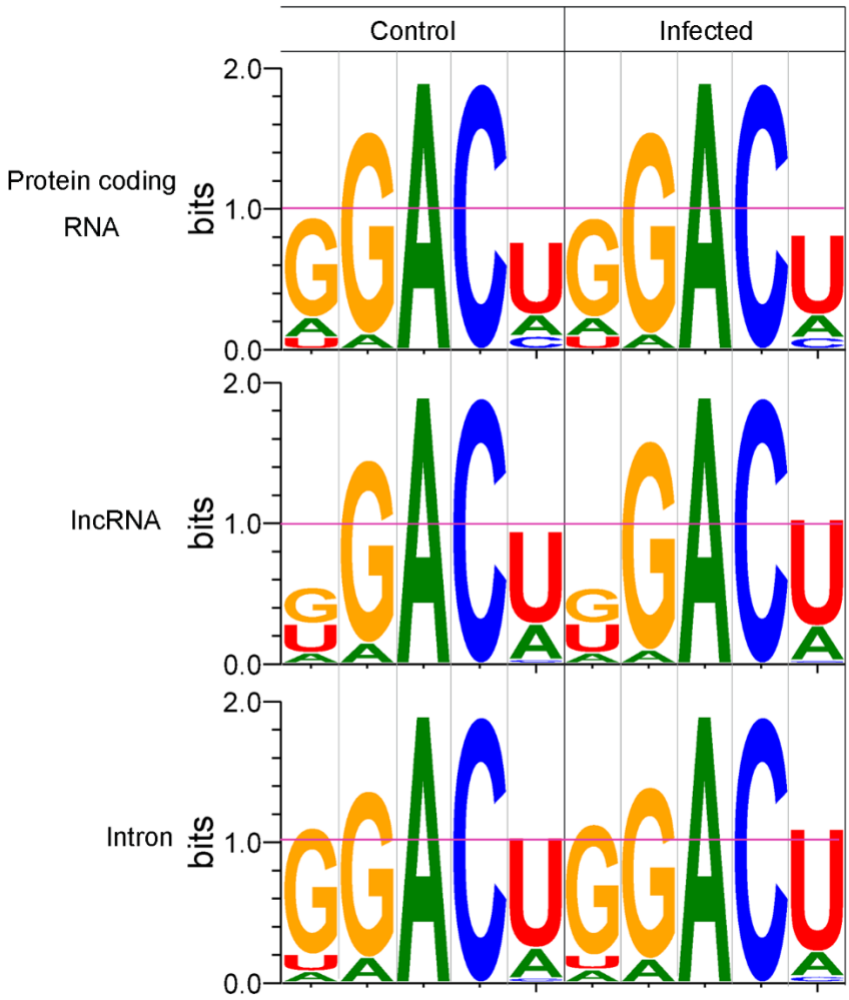

Figure 3.4 Sequence Motifs in ^{m6}A-Modified Sites for Different RNA Types in Control and SINV-Infected A549 Cells. Sequence logos display the enriched motifs surrounding ^{m6}A-modified sites in protein-coding RNA, lncRNA, and intronic RNA under control and SINV-infected conditions. Each logo represents the consensus sequence for ^{m6}A modification sites, with the height of each nucleotide reflecting its information content in bits. The large green "A" denotes the position of the ^{m6}A modification in each motif, indicating the central adenosine where methylation occurs.

3.4.4 Correlation Analysis of Transcript Abundance, Poly(A) Tail Length, ^{m6}A Modifications, and Protein Abundance

Correlation analyses using hexbin plots revealed weak positive relationships between transcript abundance, poly(A) tail length, ^{m6}A modification frequency, and protein abundance in SINV-infected A549 cells. Each hexbin plot displays the density of data points as a function of color intensity, with correlation coefficients (r) and p-values indicating the statistical strength and significance of each observed relationship (Figure 3.5).

A weak positive correlation was observed between transcript abundance and poly(A) tail length ($r = 0.13$, $p < 0.0001$; Figure A). This finding suggests that transcripts with higher abundance may have slightly longer poly(A) tails on average, though the association is not strong. The poly(A) tail length, often associated with mRNA stability and translational efficiency, may contribute incrementally to the stability of highly abundant transcripts, potentially enhancing their persistence in the cellular environment. However, the modest correlation implies that poly(A) tail length alone does not strongly dictate transcript abundance in SINV-infected cells.

The relationship between transcript abundance and the number of ^{m6}A modifications also demonstrated a weak positive correlation ($r = 0.10$, $p < 0.0001$; Figure B). This suggests that more abundant transcripts tend to possess slightly higher numbers of ^{m6}A modifications, although the effect is limited. Given ^{m6}A's known roles in regulating mRNA stability, splicing, and translation, this slight positive association may reflect a regulatory mechanism wherein ^{m6}A modifications help maintain or enhance the stability and functionality of abundant transcripts. Despite this tendency, the weak

correlation indicates that ^{m6}A modification frequency alone does not serve as a major determinant of transcript abundance.

In exploring the link between transcript and protein abundance, a weak positive correlation ($r = 0.10$, $p < 0.0001$; Figure C) was observed, suggesting that higher-abundance transcripts are somewhat more likely to correspond to higher protein levels. This modest association aligns with the expectation that transcriptional levels influence protein production. However, the limited strength of this correlation underscores the complexity of translational control in SINV-infected cells, where additional factors beyond transcript abundance—such as ^{m6}A modifications and poly(A) tail length—likely contribute to protein expression levels.

Further analysis of protein abundance relative to poly(A) tail length revealed a very weak positive correlation ($r = 0.07$, $p < 0.0001$; Figure D), suggesting that proteins translated from transcripts with longer poly(A) tails may exhibit slightly higher abundance. This finding aligns with the concept that longer poly(A) tails can enhance mRNA stability and translation, potentially contributing to increased protein production. However, given the weak correlation, poly(A) tail length alone does not appear to be a strong predictor of protein abundance, highlighting the influence of other regulatory factors.

Lastly, the correlation between protein abundance and ^{m6}A modification frequency was also weakly positive ($r = 0.06$, $p < 0.0001$; Figure E), indicating that proteins produced from ^{m6}A-modified transcripts may have slightly higher abundance. ^{m6}A modifications are known to facilitate RNA stability and translation by modulating RNA structure and interactions with RNA-binding proteins. The observed weak

correlation suggests that while ^{m6}A contributes to the regulation of protein levels, it operates within a broader landscape of translational controls, limiting its direct impact on protein abundance.

In summary, these analyses reveal only modest positive correlations between transcript abundance, poly(A) tail length, ^{m6}A modifications, and protein abundance in SINV-infected A549 cells. These findings suggest that while poly(A) tail length and ^{m6}A modifications may enhance RNA stability and translational efficiency, their direct impact on protein abundance remains limited. The weak correlations reflect a complex regulatory environment, where numerous post-transcriptional and translational mechanisms interact to fine-tune protein output beyond transcript abundance alone. This complexity likely enables the cell to dynamically adapt to viral infection by modulating RNA and protein stability in a nuanced manner.

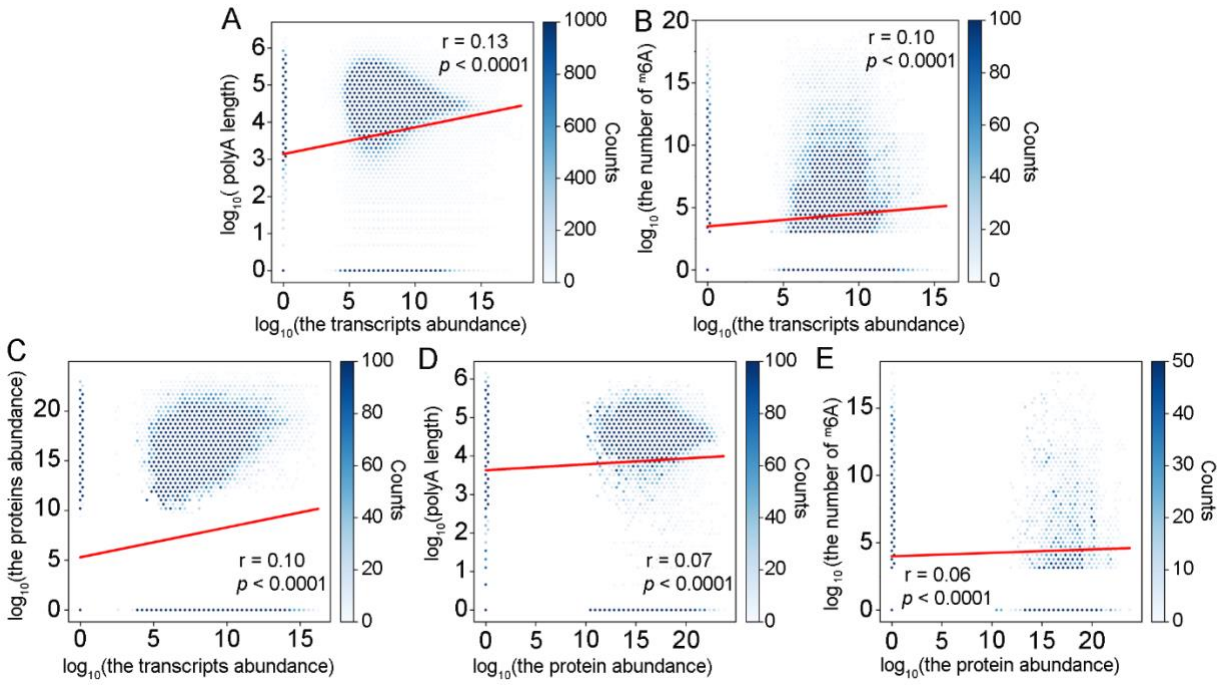

Figure 3.5 Correlation Analysis of Transcript Abundance, Poly(A) Tail Length, ^{m6}A Modifications, and Protein Abundance in SINV-Infected A549 Cells. Hexbin plots illustrate the correlations between transcript abundance, poly(A) tail length, ^{m6}A modification count, and protein abundance in SINV-infected A549 cells. Color intensity indicates the density of data points within each bin, with darker shades representing higher data density. Correlation coefficients (r) and p-values are shown for each plot to indicate the strength and significance of each relationship.

3.4.5 Differential Expression and Post-Transcriptional Modifications in SINV-Infected A549 Cells

The volcano plots demonstrate extensive alterations in transcript abundance, protein abundance, isoform diversity, poly(A) tail length, and ^{m6}A modification frequency in A549 cells upon Sindbis virus (SINV) infection (Figure 3.6). These molecular changes indicate a complex and adaptive response by the host, with significant adjustments across multiple regulatory layers.

Changes in transcript abundance (Figure 3.6A) reflect a widespread transcriptional activation and suppression pattern. The increase in transcript levels, represented by red data points on the right side, suggests upregulation of genes involved in immune and antiviral responses. This upregulation likely supports cellular defense mechanisms against the virus by activating pathways aimed at inhibiting viral replication and managing cellular stress. In contrast, the blue points on the left side indicate downregulated transcripts, likely corresponding to pathways suppressed to conserve cellular resources or to prevent viral exploitation. Protein abundance changes, though fewer in number than transcript-level changes, follow a similar trend, suggesting

that transcriptional modifications are selectively translated into protein-level adjustments, with additional regulatory mechanisms influencing which transcripts lead to protein production.

SINV infection impacts isoform diversity and poly(A) tail length as well (Figure 3.6B). Increased poly(A) tail lengths, indicated by yellow data points, suggest enhanced stability and translational efficiency for certain transcripts, potentially enabling prolonged expression of key proteins necessary for an effective antiviral response. On the other hand, decreased poly(A) tail lengths for other transcripts may facilitate rapid turnover, aligning with the need to dynamically adjust mRNA stability in response to viral infection. This modulation of poly(A) tail length underscores the cell's strategy to stabilize specific transcripts while allowing others to degrade quickly, balancing resource allocation and functional needs during infection.

Differential ^{m6}A modification patterns further characterize the host's response to SINV (Figure 3.6C). Increased ^{m6}A modifications indicate selective stabilization or enhanced translation of specific RNAs, which may include immune or stress-response genes critical to the host's defense. In contrast, decreased ^{m6}A modifications on other transcripts suggest an intentional downregulation at the post-transcriptional level. This bidirectional adjustment of ^{m6}A methylation highlights the cell's ability to finely regulate RNA stability and translation in response to viral infection, supporting an efficient and targeted antiviral response.

Additional analyses (Figure 3.6 D, E and F) show specific patterns in isoform diversity, protein abundance, and ^{m6}A modification changes across various transcripts and proteins. Increased isoform diversity in Figure 3.6D suggests that alternative

splicing is actively modulated, potentially allowing the host to produce a wider array of protein isoforms tailored to different functional demands under viral stress. Such isoform diversity contributes to a versatile proteome, equipping the cell to respond to various challenges posed by SINV infection. Protein abundance and ^{m6}A modifications across multiple transcripts and proteins, as shown in Figure 3.6E and 3.6F, illustrate a broad, coordinated adjustment in response to infection.

In summary, these analyses reveal a multi-layered response to SINV infection in A549 cells, characterized by transcriptional upregulation, isoform diversification, modulation of poly(A) tail length, and selective ^{m6}A methylation changes. The cell employs a combination of transcriptional and post-transcriptional mechanisms, including RNA stability control and alternative splicing, to optimize its response to viral infection. The interplay among these regulatory layers underscores the complexity of the cellular response, with post-transcriptional modifications such as poly(A) tail length and ^{m6}A methylation playing crucial roles in dynamically adapting the gene expression landscape in an infected state.

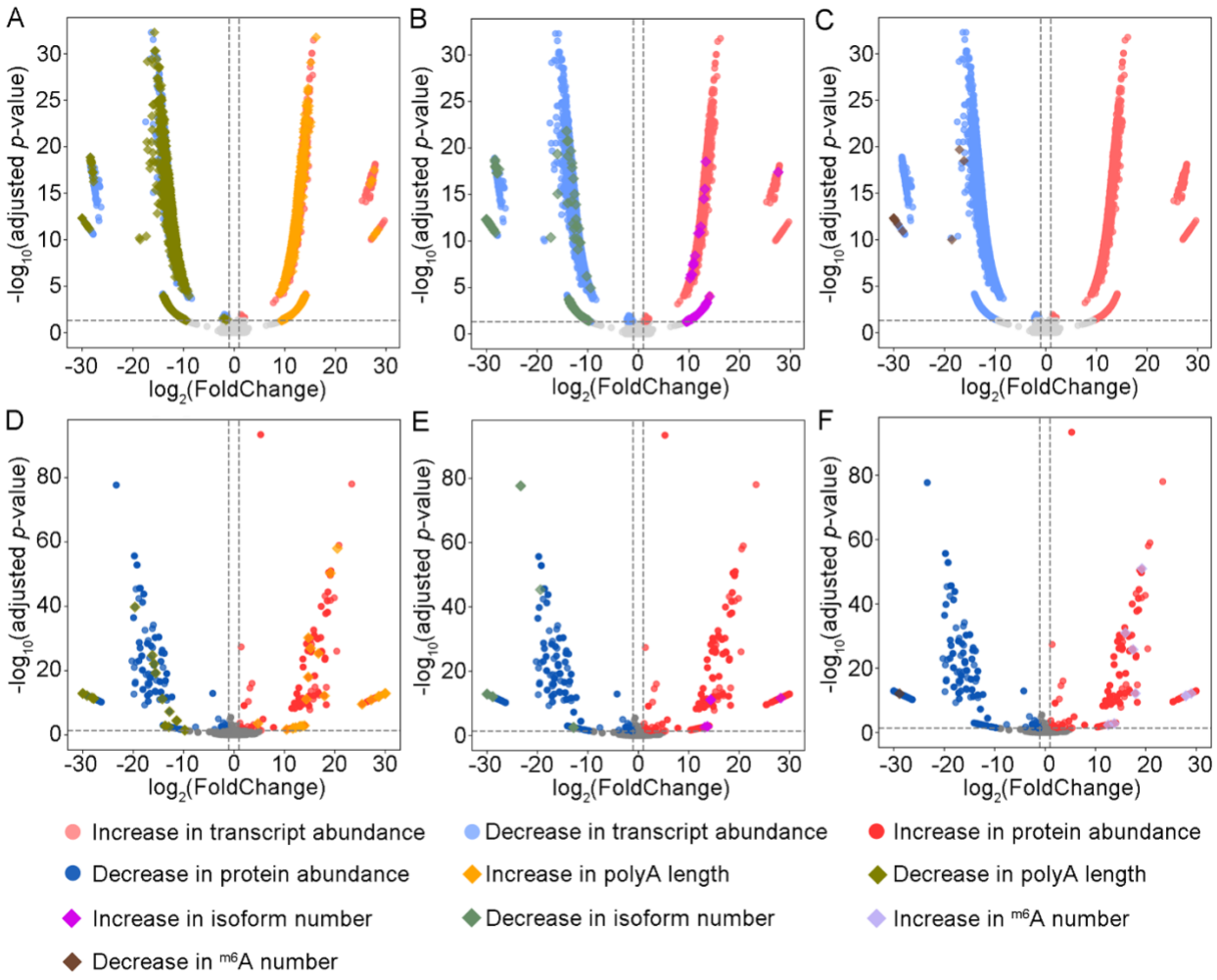

Figure 3.6 Differential Expression and Modifications in SINV-Infected A549 Cells.

Volcano plots show significant changes in transcript abundance, protein abundance, isoform number, poly(A) tail length, and ^{m6}A modification frequency in A549 cells upon Sindbis virus (SINV) infection. Each plot depicts the $\log_2$(fold change) on the x-axis and the -$\log_{10}$(adjusted p-value) on the y-axis, with data points colored according to the specific type of molecular change.

3.4.6 Functional Enrichment Analysis of Differentially Expressed Genes in SINV-Infected A549 Cells

Gene Ontology (GO) and KEGG pathway enrichment analyses highlight significant changes in biological processes, cellular components, molecular functions, and pathways in A549 cells following Sindbis virus (SINV) infection (Figure 3.7). These analyses reveal specific host pathways that are either activated or repressed in response to viral infection, shedding light on the host's regulatory mechanisms to counteract viral activity.

The GO enrichment analysis shows an upregulation in several immune-related biological processes and cellular components in response to SINV infection. Increased expression is observed in pathways associated with immune responses, such as "positive regulation of innate immune response," "natural killer cell activation," "type I interferon signaling pathway," and "interleukin-6 receptor binding." These pathways indicate an active antiviral immune response, as type I interferon signaling and innate immune pathways are central to the host's immediate defense mechanisms against viral infections. Additional enrichment in processes related to "nuclear-transcribed mRNA catabolic processes" and "nucleocytoplasmic transport" suggests that SINV infection might enhance mRNA degradation and transport activities, potentially as a way for the host to degrade viral RNA or manage the increased mRNA demand during infection.

Conversely, downregulated processes include cellular processes linked to basic cell maintenance and structure, such as "cell-cell adhesion," "negative regulation of stress fiber assembly," and "cell-matrix adhesion." This downregulation likely reflects the host's strategy to conserve energy and resources by suppressing non-essential

structural and maintenance functions during viral infection. Reducing energy expenditures on cellular adhesion and cytoskeletal organization may allow the cell to redirect resources towards immune defense and other critical responses to the viral challenge.

KEGG pathway analysis reveals that SINV infection leads to upregulation in multiple immune signaling pathways, including "JAK-STAT signaling pathway," "Cytokine-cytokine receptor interaction," "RIG-I-like receptor signaling pathway," and "Toll-like receptor signaling pathway." These pathways are crucial for detecting viral components and activating downstream immune responses. The enrichment of "Natural killer cell mediated cytotoxicity" and "B cell receptor signaling pathway" further supports the activation of antiviral immune responses, suggesting that the host is mobilizing both innate and adaptive immune components to control the infection. Metabolic pathways like "Alanine, aspartate, and glutamate metabolism" and "Tryptophan metabolism" are also upregulated, which may indicate that SINV infection alters host metabolic processes to support viral replication or meet the heightened energy demands associated with an activated immune state.

In the lower-left panel, GO terms associated with downregulated genes indicate a repression in structural components and cell division processes, including "mitotic spindle assembly," "sister chromatid cohesion," and "cytoplasmic microtubule organization." Downregulation of these processes suggests that SINV infection disrupts cell division and structural integrity, potentially as a result of viral manipulation of the host cell cycle or as a host-driven response to prevent viral replication. Additionally, downregulated metabolic processes, such as "sulfurtransferase activity" and

"sulfotransferase activity," indicate a broad suppression of specific metabolic pathways, likely reflecting the host's effort to limit resources available to the virus.

The KEGG pathway analysis reveals downregulation in various metabolic and structural pathways, including "Citrate cycle (TCA cycle)," "Oxidative phosphorylation," and "Proteasome." These suppressed pathways indicate a shift in energy production and protein degradation systems, as the host likely reallocates resources away from energy-intensive processes that are not essential during an infection. Suppression of energy production and protein processing pathways may also serve as a defensive measure, depriving the virus of metabolic resources and essential cellular machinery needed for its replication.

In summary, these analyses reveal a robust and selective reprogramming of cellular processes in response to SINV infection in A549 cells. Upregulated immune pathways underscore the host’s activation of innate and adaptive defenses, while downregulation of cellular structural and metabolic pathways indicates a shift in resource allocation aimed at limiting viral replication. This functional reorganization reflects the dynamic interplay between the virus and host, highlighting potential therapeutic targets that could support or enhance the host's antiviral response.

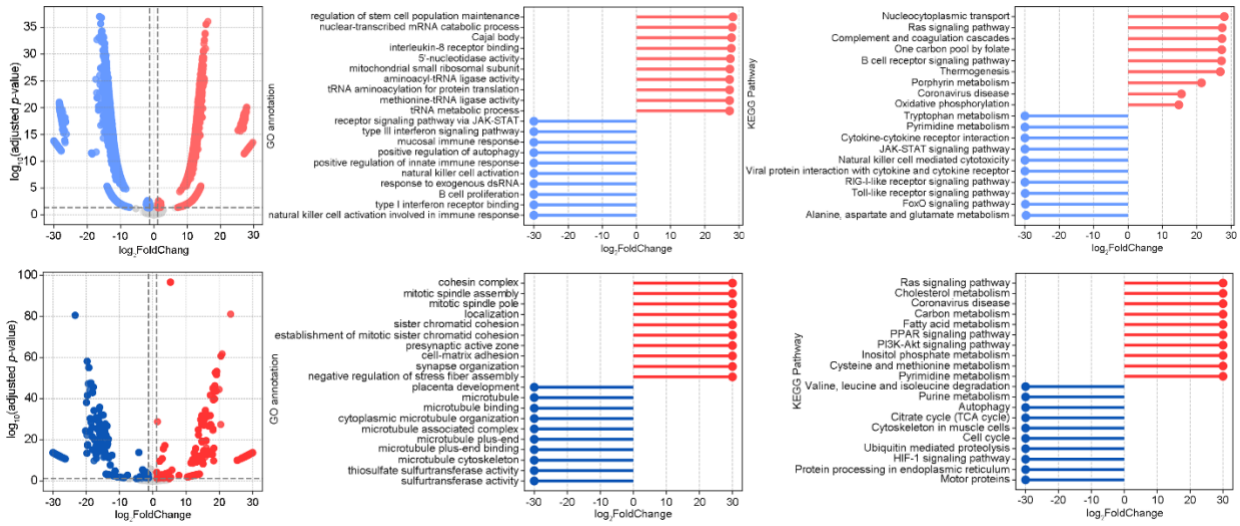

Figure 3.7 Functional Enrichment Analysis of Differentially Expressed Genes in SINV-Infected A549 Cells.

3.5 Discussion

The results of this study highlight the intricate and multifaceted response of A549 human lung epithelial cells to Sindbis virus (SINV) infection, revealing significant changes across transcriptomic, epitranscriptomic, and proteomic levels. The combined approach of nanopore direct RNA sequencing and proteomic analysis provided a comprehensive view of how SINV infection influences host and viral RNA modifications, alternative splicing, and protein expression. These findings reveal a complex interplay between host defense mechanisms and viral strategies, with SINV leveraging specific molecular changes to facilitate its replication while the host enacts counteractive responses at various regulatory levels.

One of the key findings is the increased abundance of immune-related transcripts in response to SINV infection. Upregulated transcripts, particularly those linked to innate immune pathways, illustrate the activation of host antiviral responses. At the same time, the repression of transcripts associated with cellular structure and maintenance reflects the host's strategy to allocate resources selectively, prioritizing immune functions over non-essential processes. This selective transcriptional modulation suggests a coordinated cellular reprogramming, in which immune activation and energy conservation work together to enhance the host's resilience against viral exploitation [104, 105].

Post-transcriptional modifications, particularly ^{m6}A methylation and changes in poly(A) tail length, further underscore the host's adaptive response to infection [106]. An increase in ^{m6}A-modified transcripts in infected cells suggests that ^{m6}A methylation plays a role in stabilizing immune-related RNAs, facilitating their translation and persistence in

the cell. By stabilizing these transcripts, ^{m6}A modifications likely support the rapid production of proteins essential for antiviral defenses. However, the weak correlations between ^{m6}A levels, transcript abundance, and protein abundance indicate that ^{m6}A acts in concert with other factors, enhancing RNA stability and translational efficiency without fully controlling protein output. Additionally, the reduction in poly(A) tail lengths in lncRNAs and protein-coding RNAs implies that SINV manipulates RNA stability by selectively targeting host RNAs, potentially limiting the host's capacity to sustain a prolonged immune response [107]. This targeted modulation reflects a sophisticated viral strategy, where SINV-induced changes in RNA stability enable it to create a cellular environment favorable to its replication.

The proteomic data reveals a parallel trend, with infected cells exhibiting elevated immune-related proteins. This suggests that the upregulation of immune transcripts translates into protein-level changes that enhance the host's ability to counter SINV infection. Despite the observed weak positive correlation between transcript and protein abundance, these findings indicate that other regulatory mechanisms contribute to the translation process. The slight association between poly(A) tail length and protein abundance suggests that while poly(A) tail length aids in RNA stability and translational efficiency, it is one of several factors influencing protein levels. This further reinforces the idea that SINV infection drives a complex translational control environment in which multiple layers of regulation collectively shape the host's proteomic response [108].

A notable aspect of the study is the conservation of the GGACU motif at ^{m6}A-modified sites across both control and infected conditions. This stability implies that SINV infection does not disrupt the host's ^{m6}A methylation machinery, allowing the host

to maintain core regulatory functions while selectively modulating immune-related transcripts [109]. By targeting this conserved sequence, the host can regulate critical ^{m6}A-dependent functions, such as RNA stability and splicing, even during viral infection. This suggests an evolutionary resilience in ^{m6}A regulation, where the host maintains its essential post-transcriptional controls even as it modulates immune-related transcripts to counteract SINV's effects [38, 110].

Gene Ontology (GO) and KEGG pathway analyses provide further insights into how SINV infection reorganizes cellular functions. The upregulation of immune-related pathways, such as interferon signaling and cytokine production, highlights the host's efforts to mobilize a robust antiviral response. Conversely, the downregulation of cellular structure and metabolic pathways suggests a strategic redirection of resources to sustain immune responses. By suppressing cellular maintenance functions, such as cell adhesion and cytoskeletal organization, the host appears to limit energy expenditures, reallocating these resources to defend against the viral invasion. Furthermore, the downregulation of oxidative phosphorylation and TCA cycle pathways indicates that SINV infection leads to a shift in energy production systems, perhaps as a defensive response to restrict viral access to metabolic resources.

This study offers valuable insights into the host-pathogen dynamics of SINV infection, showing how the virus can exploit the host's molecular machinery while the host mounts a multi-tiered defense. While the findings advance our understanding of SINV-host interactions, they also highlight the need for further research, particularly given the weak correlations observed between post-transcriptional modifications and protein abundance. Future studies could explore these interactions in other cell types or

under varying SINV infection conditions to provide a more comprehensive view of SINV's impact on cellular processes. This work underscores the potential of targeting RNA modifications, such as ^{m6}A, as part of antiviral strategies, as these modifications play an essential role in fine-tuning the host's response to infection.

In summary, SINV infection induces a coordinated molecular response in A549 cells characterized by heightened transcriptional activity, selective ^{m6}A methylation, and adaptive protein expression changes. The nuanced regulation of ^{m6}A methylation and poly(A) tail length, along with the host's strategic allocation of resources, illustrates a sophisticated defense strategy that balances immune activation with energy conservation. This study's findings emphasize the importance of post-transcriptional modifications in shaping the host's antiviral response and suggest that targeting these modifications could hold therapeutic potential in managing SINV and related viral infections.

Chapter 4 Summary and Outlook

4.1 Summary

This research provides an in-depth investigation into the role of ^{m6}A RNA modifications in modulating host cellular responses to viral infections, focusing specifically on influenza virus and Sindbis virus (SINV) interactions within human epithelial cells. Utilizing advanced nanopore direct RNA sequencing and comprehensive proteomic analyses, the study elucidated how ^{m6}A modifications, poly(A) tail length dynamics, and subsequent protein expression alterations contribute to both viral replication strategies and host defense mechanisms. Findings demonstrated that influenza and SINV infections trigger distinct epitranscriptomic and proteomic shifts in host cells, reflecting a finely balanced interaction where both viruses utilize host cellular machinery to promote replication while the host implements adaptive mechanisms to limit viral spread.

A critical finding was the dual role of ^{m6}A modifications in stabilizing immune-related RNAs in influenza-infected cells and supporting viral RNA stability in SINV-infected cells. In the context of influenza infection, increased ^{m6}A modifications on immune-related transcripts led to enhanced transcript stability and translation efficiency, thereby amplifying antiviral immune responses. Conversely, in SINV-infected cells, ^{m6}A modifications were observed not only on viral RNAs, aiding their stability and translation, but also on specific host immune transcripts. This dual modulation suggests that SINV may strategically leverage ^{m6}A methylation to stabilize its own transcripts while selectively regulating host immune transcripts to create a more favorable environment for its replication. These findings highlight the sophisticated interplay of ^{m6}A

modifications as both viral mechanisms and host defenses, underscoring the complexity of epitranscriptomic regulation in viral pathogenesis.

A significant insight was the conservation of the GGACU motif at ^{m6}A-modified sites across both control and infected samples. This motif conservation indicates that the underlying sequence specificity for ^{m6}A methylation remains robust despite viral infection, allowing essential ^{m6}A-regulated processes, such as RNA stability, splicing, and translation, to persist. Such conservation suggests that while viral infection can influence the quantity and distribution of ^{m6}A modifications, it does not disrupt the fundamental targeting machinery of ^{m6}A methylation, thereby preserving the host's ability to utilize ^{m6}A as a defense mechanism. Additionally, weak correlations observed between ^{m6}A levels, poly(A) tail length, and protein abundance imply that ^{m6}A and poly(A) modifications, while contributing to RNA stability and translational control, likely operate within a broader regulatory network. This network may involve other RNA modifications and RNA-binding proteins that collectively refine the host’s response to infection by modulating transcript processing and protein production in a context-specific manner.

4.2 Outlook

The insights gained from this study highlight the potential of ^{m6}A modifications as therapeutic targets in the development of antiviral strategies. Given their role in regulating the stability and translation of both viral and immune-related transcripts, targeting ^{m6}A pathways presents a promising avenue for therapeutic intervention. Future research could focus on identifying and developing small-molecule inhibitors or modulators of ^{m6}A "writers" (such as METTL3 and METTL14), "readers" (e.g., YTHDF

proteins), and "erasers" (such as FTO and ALKBH5) to selectively inhibit viral replication or enhance host antiviral responses [111, 112]. Modulating these proteins with precision could enable selective regulation of ^{m6}A marks on viral versus host transcripts, offering the ability to disrupt viral RNA stability while simultaneously promoting the stability and translation of immune transcripts.

Further research is also needed to unravel the complex regulatory interactions between ^{m6}A modifications, poly(A) tail dynamics, and other epitranscriptomic markers within the context of viral infection. The weak correlations found between ^{m6}A levels, transcript abundance, and protein expression suggest that additional post-transcriptional controls, potentially involving alternative RNA modifications such as pseudouridylation, are influential in defining gene expression outcomes during infection. Future studies could leverage long-read sequencing and single-cell approaches to capture these interactions in high resolution, enhancing our understanding of how different RNA modifications coordinate to optimize host responses against diverse viral pathogens [113-115].

Another promising direction lies in the exploration of ^{m6}A modifications in non-coding RNAs, which have increasingly been recognized for their regulatory roles in immune response pathways. Long non-coding RNAs (lncRNAs), in particular, exhibit significant ^{m6}A methylation shifts in response to viral infections, potentially modulating gene networks that are crucial for cellular defenses. Future research could explore how ^{m6}A modifications on lncRNAs influence their structure, stability, and interactions with immune signaling molecules, uncovering new regulatory layers within antiviral responses that might serve as therapeutic targets.

Lastly, while this study focused on A549 human epithelial cells, further investigation in other relevant cell types, such as primary immune cells or respiratory epithelial cells, will be essential to generalize these findings across different physiological environments. Such studies could investigate the variability of ^{m6}A-related responses among cell types, as well as the differential effects of ^{m6}A modifications on host-pathogen dynamics in respiratory versus non-respiratory infections. By broadening the cellular and tissue contexts, future research could better inform the design of targeted therapies aimed at enhancing ^{m6}A-mediated antiviral responses or inhibiting viral exploitation of these modifications in various infectious diseases.

In conclusion, this research underscores the significance of ^{m6}A modifications and poly(A) tail dynamics as integral components of the host's regulatory network during viral infection. As the understanding of RNA modifications in virus-host interactions deepens, there is potential to develop innovative, targeted therapies that exploit these regulatory pathways. By precisely modulating ^{m6}A-related proteins or fine-tuning ^{m6}A deposition on specific transcripts, it may be possible to disrupt viral lifecycles while reinforcing the host's immune defenses, offering new hope for combating viral infections through an epitranscriptomic lens.

Reference

1. Dunning, J., R.S. Thwaites, and P.J. Openshaw, *Seasonal and pandemic influenza: 100 years of progress, still much to learn.* Mucosal immunology, 2020. **13**(4): p. 566-573.
2. Williams, B.A., et al., *Outlook of pandemic preparedness in a post-COVID-19 world.* npj Vaccines, 2023. **8**(1): p. 178.
3. Lofgren, E., et al., *Influenza seasonality: underlying causes and modeling theories.* Journal of virology, 2007. **81**(11): p. 5429-5436.
4. Petrova, V.N. and C.A. Russell, *The evolution of seasonal influenza viruses.* Nature Reviews Microbiology, 2018. **16**(1): p. 47-60.
5. Omenn, G.S., *Evolution and public health.* Proceedings of the National Academy of Sciences, 2010. **107**(suppl_1): p. 1702-1709.
6. Morens, D.M., G.K. Folkers, and A.S. Fauci, *The challenge of emerging and re-emerging infectious diseases.* Nature, 2004. **430**(6996): p. 242-249.
7. Chumakov, K., et al., *Old vaccines for new infections: Exploiting innate immunity to control COVID-19 and prevent future pandemics.* Proceedings of the National Academy of Sciences, 2021. **118**(21): p. e2101718118.
8. Khan, T. and S. Raza, *Exploration of computational aids for effective drug designing and management of viral diseases: A comprehensive review.* Current Topics in Medicinal Chemistry, 2023. **23**(17): p. 1640-1663.
9. Pichlmair, A., et al., *Viral immune modulators perturb the human molecular network by common and unique strategies.* Nature, 2012. **487**(7408): p. 486-490.
10. Mohite, P., et al., *Revolutionizing Cancer Treatment: Unleashing the Power of Viral Vaccines, Monoclonal Antibodies, and Proteolysis-Targeting Chimeras in the New Era of Immunotherapy.* ACS omega, 2024. **9**(7): p. 7277-7295.
11. Speranza, E., *Understanding virus–host interactions in tissues.* Nature Microbiology, 2023. **8**(8): p. 1397-1407.
12. Fan, X., et al., *Decoding epitranscriptomic regulation of viral infection: mapping of RNA N 6-methyladenosine by advanced sequencing technologies.* Cellular & Molecular Biology Letters, 2024. **29**(1): p. 42.

13. Bradley, R.K. and O. Anczuków, *RNA splicing dysregulation and the hallmarks of cancer.* Nature Reviews Cancer, 2023. **23**(3): p. 135-155.
14. Cao, X., et al., *Identification of RNA structures and their roles in RNA functions.* Nature Reviews Molecular Cell Biology, 2024: p. 1-18.
15. Sun, H., et al., *Regulation and functions of non-m6A mRNA modifications.* Nature Reviews Molecular Cell Biology, 2023. **24**(10): p. 714-731.
16. Rogalska, M.E., C. Vivori, and J. Valcárcel, *Regulation of pre-mRNA splicing: roles in physiology and disease, and therapeutic prospects.* Nature Reviews Genetics, 2023. **24**(4): p. 251-269.
17. Price, A.M., et al., *Direct RNA sequencing reveals m6A modifications on adenovirus RNA are necessary for efficient splicing.* Nature communications, 2020. **11**(1): p. 6016.
18. Batista, P.J., et al., *m6A RNA modification controls cell fate transition in mammalian embryonic stem cells.* Cell stem cell, 2014. **15**(6): p. 707-719.
19. Williams, G.D., N.S. Gokhale, and S.M. Horner, *Regulation of viral infection by the RNA modification N6-methyladenosine.* Annual review of virology, 2019. **6**(1): p. 235-253.
20. Cross, S.T., et al., *RNA regulatory processes in RNA virus biology.* Wiley Interdisciplinary Reviews: RNA, 2019. **10**(5): p. e1536.
21. McFadden, M.J. and S.M. Horner, *N6-methyladenosine regulates host responses to viral infection.* Trends in biochemical sciences, 2021. **46**(5): p. 366-377.
22. Wang, A., et al., *m6A modifications regulate intestinal immunity and rotavirus infection.* Elife, 2022. **11**: p. e73628.
23. Lu, M., et al., *N 6-methyladenosine modification enables viral RNA to escape recognition by RNA sensor RIG-I.* Nature microbiology, 2020. **5**(4): p. 584-598.
24. Han, D. and M.M. Xu, *RNA modification in the immune system.* Annual review of immunology, 2023. **41**(1): p. 73-98.
25. Berdasco, M. and M. Esteller, *Clinical epigenetics: seizing opportunities for translation.* Nature Reviews Genetics, 2019. **20**(2): p. 109-127.
26. Mitrea, D.M., et al., *Modulating biomolecular condensates: a novel approach to drug discovery.* Nature Reviews Drug Discovery, 2022. **21**(11): p. 841-862.

27. Zheng, J., et al., *Epitranscriptomic modifications in mesenchymal stem cell differentiation: advances, mechanistic insights, and beyond.* Cell Death & Differentiation, 2024. **31**(1): p. 9-27.
28. Wu, Y., et al., *Transfer learning enables identification of multiple types of RNA modifications using nanopore direct RNA sequencing.* Nature Communications, 2024. **15**(1): p. 4049.
29. Yang, S.L., et al., *Comprehensive mapping of SARS-CoV-2 interactions in vivo reveals functional virus-host interactions.* Nature communications, 2021. **12**(1): p. 5113.
30. Niu, X., et al., *Infectious bursal disease virus VP5 triggers host shutoff in a transcription-dependent manner.* Mbio, 2024. **15**(3): p. e03433-23.
31. Ashraf, U., et al., *Advances in analyzing virus-induced alterations of host cell splicing.* Trends in microbiology, 2019. **27**(3): p. 268-281.
32. Depledge, D.P., et al., *Direct RNA sequencing on nanopore arrays redefines the transcriptional complexity of a viral pathogen.* Nature communications, 2019. **10**(1): p. 754.
33. Dawson, A.R., et al., *Post-translation regulation of influenza virus replication.* Annual review of virology, 2020. **7**(1): p. 167-187.
34. Ge, Y., et al., *Degradation of WTAP blocks antiviral responses by reducing the m6A levels of IRF3 and IFNAR1 Mrna.* EMBO reports, 2021. **22**(11): p. e52101.
35. Nie, L., et al., *Integrative analysis of transcriptomic and proteomic data: challenges, solutions and applications.* Critical reviews in biotechnology, 2007. **27**(2): p. 63-75.
36. Hegde, P.S., I.R. White, and C. Debouck, *Interplay of transcriptomics and proteomics.* Current opinion in biotechnology, 2003. **14**(6): p. 647-651.
37. Li, N. and T.M. Rana, *Regulation of antiviral innate immunity by chemical modification of viral RNA.* Wiley Interdisciplinary Reviews: RNA, 2022. **13**(6): p. e1720.
38. Zhang, C., J. Fu, and Y. Zhou, *A review in research progress concerning m6A methylation and immunoregulation.* Frontiers in immunology, 2019. **10**: p. 922.

39. Zhang, K., et al., *Caspases switch off the m6A RNA modification pathway to foster the replication of a ubiquitous human tumor virus.* Mbio, 2021. **12**(4): p. 10.1128/mbio. 01706-21.
40. Shen, S. and L.-S. Zhang, *The regulation of antiviral innate immunity through non-m6A RNA modifications.* Frontiers in Immunology, 2023. **14**: p. 1286820.
41. Gokhale, N.S., et al., *Altered m6A modification of specific cellular transcripts affects flaviviridae infection.* Molecular cell, 2020. **77**(3): p. 542-555. e8.
42. Maida, C.M., et al., *Detection of influenza virus in urban wastewater during the season 2022/2023 in Sicily, Italy.* Frontiers in Public Health, 2024. **12**: p. 1383536.
43. Valiveti, C.K., *Polymeric Nanovaccine Delivery System for Influenza Vaccine.* 2022.
44. Gaucherand, L. and M.M. Gaglia, *The role of viral RNA degrading factors in shutoff of host gene expression.* Annual review of virology, 2022. **9**(1): p. 213-238.
45. Statello, L., et al., *Gene regulation by long non-coding RNAs and its biological functions.* Nature reviews Molecular cell biology, 2021. **22**(2): p. 96-118.
46. Kalamara, V. and G.A. Garinis, *The epitranscriptome: reshaping the DNA damage response.* Trends in Cell Biology, 2024.
47. Boulias, K. and E.L. Greer, *Biological roles of adenine methylation in RNA.* Nature Reviews Genetics, 2023. **24**(3): p. 143-160.
48. Rand, U., et al., *Uncoupling of the dynamics of host–pathogen interaction uncovers new mechanisms of viral interferon antagonism at the single-cell level.* Nucleic acids research, 2014. **42**(13): p. e109-e109.
49. Yang, E. and M.M. Li, *All about the RNA: interferon-stimulated genes that interfere with viral RNA processes.* Frontiers in immunology, 2020. **11**: p. 605024.
50. Tan, B., et al., *Viral and cellular N6-methyladenosine and N6, 2'-O-dimethyladenosine epitranscriptomes in the KSHV life cycle.* Nature microbiology, 2018. **3**(1): p. 108-120.
51. Schulz, T.F. and E. Cesarman, *Kaposi Sarcoma-associated Herpesvirus: mechanisms of oncogenesis.* Current opinion in virology, 2015. **14**: p. 116-128.

52. Liu, J.e., et al., *The m6A methylome of SARS-CoV-2 in host cells.* Cell research, 2021. **31**(4): p. 404-414.
53. Xue, M., et al., *Viral RNA N 6-methyladenosine modification modulates both innate and adaptive immune responses of human respiratory syncytial virus.* PLoS pathogens, 2021. **17**(12): p. e1010142.
54. Zhang, S., et al., *Targeting epigenetic regulators for inflammation: Mechanisms and intervention therapy.* MedComm, 2022. **3**(4): p. e173.
55. Feng, J., et al., *Global profiling reveals common and distinct N6-methyladenosine (m6A) regulation of innate immune responses during bacterial and viral infections.* Cell Death & Disease, 2022. **13**(3): p. 234.
56. Potužník, J.F. and H. Cahová, *It's the little things (in viral RNA).* MBio, 2020. **11**(5): p. 10.1128/mbio. 02131-20.
57. Tan, L., et al., *Analysis of bacterial transcriptome and epitranscriptome using nanopore direct RNA sequencing.* Nucleic acids research, 2024. **52**(15): p. 8746-8762.
58. Wang, Y., et al., *Nanopore sequencing technology, bioinformatics and applications.* Nature biotechnology, 2021. **39**(11): p. 1348-1365.
59. Ying, Y.-L., et al., *Nanopore-based technologies beyond DNA sequencing.* Nature nanotechnology, 2022. **17**(11): p. 1136-1146.
60. Diensthuber, G., et al., *Enhanced detection of RNA modifications and read mapping with high-accuracy nanopore RNA basecalling models.* Genome Research, 2024: p. gr. 278849.123.
61. Hendra, C., et al., *Detection of m6A from direct RNA sequencing using a multiple instance learning framework.* Nature methods, 2022. **19**(12): p. 1590-1598.
62. Pater, A.A., et al., *High throughput nanopore sequencing of SARS-CoV-2 viral genomes from patient samples.* Journal of biological methods, 2021. **8**(COVID 19 Spec Iss).
63. Decombe, A., P. El Kazzi, and E. Decroly, *Interplay of RNA 2'-O-methylations with viral replication.* Current Opinion in Virology, 2023. **59**: p. 101302.
64. Abebe, J.S., R. Verstraten, and D.P. Depledge, *Nanopore-based detection of viral RNA modifications.* Mbio, 2022. **13**(3): p. e03702-21.

65. Depledge, D.P. and A.C. Wilson, *Using direct RNA nanopore sequencing to deconvolute viral transcriptomes.* Current protocols in microbiology, 2020. **57**(1): p. e99.
66. Ahsan, M.U., et al., *A signal processing and deep learning framework for methylation detection using Oxford Nanopore sequencing.* Nature Communications, 2024. **15**(1): p. 1448.
67. Wick, R.R., L.M. Judd, and K.E. Holt, *Performance of neural network basecalling tools for Oxford Nanopore sequencing.* Genome biology, 2019. **20**: p. 1-10.
68. Kalikar, S., et al., *Accelerating minimap2 for long-read sequencing applications on modern CPUs.* Nature Computational Science, 2022. **2**(2): p. 78-83.
69. Su, Y., et al., *Comprehensive assessment of mRNA isoform detection methods for long-read sequencing data.* Nature Communications, 2024. **15**(1): p. 3972.
70. Yu, B., et al., *m6ATM: a deep learning framework for demystifying the m6A epitranscriptome with Nanopore long-read RNA-seq data.* Briefings in Bioinformatics, 2024. **25**(6): p. bbae529.
71. Love, M., S. Anders, and W. Huber, *Differential analysis of count data–the DESeq2 package.* Genome Biol, 2014. **15**(550): p. 10-1186.
72. Shulman, Z. and N. Stern-Ginossar, *The RNA modification N 6-methyladenosine as a novel regulator of the immune system.* Nature immunology, 2020. **21**(5): p. 501-512.
73. Roignant, J.-Y. and M. Soller, *m6A in mRNA: an ancient mechanism for fine-tuning gene expression.* Trends in Genetics, 2017. **33**(6): p. 380-390.
74. Heward, J.A. and M.A. Lindsay, *Long non-coding RNAs in the regulation of the immune response.* Trends in immunology, 2014. **35**(9): p. 408-419.
75. Zhang, L., X. Xu, and X. Su, *Noncoding RNAs in cancer immunity: functions, regulatory mechanisms, and clinical application.* Molecular cancer, 2020. **19**: p. 1-12.
76. Spinetti, G., et al., *Cardiovascular complications of diabetes: role of non-coding RNAs in the crosstalk between immune and cardiovascular systems.* Cardiovascular Diabetology, 2023. **22**(1): p. 122.

77. Perales, S.G., et al., *Therapy of infectious diseases using epigenetic approaches*, in *Epigenetics in Human Disease*. 2024, Elsevier. p. 853-882.
78. Barbieri, I. and T. Kouzarides, *Role of RNA modifications in cancer.* Nature Reviews Cancer, 2020. **20**(6): p. 303-322.
79. Baquero-Perez, B., et al., *The Tudor SND1 protein is an m6A RNA reader essential for replication of Kaposi's sarcoma-associated herpesvirus.* Elife, 2019. **8**: p. e47261.
80. Kong, J., et al., *Alphavirus infection triggers antiviral RNAi immunity in mammals.* Cell Reports, 2023. **42**(5).
81. Shakya, R., et al., *Inkoo and Sindbis viruses in blood sucking insects, and a serological study for Inkoo virus in semi-domesticated Eurasian tundra reindeer in Norway.* Virology Journal, 2022. **19**(1): p. 99.
82. Heidary, F. and R. Gharebaghi, *Ivermectin: a systematic review from antiviral effects to COVID-19 complementary regimen.* The Journal of antibiotics, 2020. **73**(9): p. 593-602.
83. Suchowiecki, K., et al., *Persistent joint pain following arthropod virus infections.* Current rheumatology reports, 2021. **23**: p. 1-12.
84. Mou, C., et al., *GETV nsP2 plays a critical role in the interferon antagonism and viral pathogenesis.* Cell Communication and Signaling, 2023. **21**(1): p. 361.
85. Dominguez, F., et al., *Alphavirus-based replicons demonstrate different interactions with host cells and can be optimized to increase protein expression.* Journal of Virology, 2023. **97**(11): p. e01225-23.
86. Fernandez, G.J., J.M. Ramírez-Mejía, and S. Urcuqui-Inchima, *Transcriptional and post-transcriptional mechanisms that regulate the genetic program in Zika virus-infected macrophages.* The International Journal of Biochemistry & Cell Biology, 2022. **153**: p. 106312.
87. McFadden, M.J., et al., *Post-transcriptional regulation of antiviral gene expression by N6-methyladenosine.* Cell reports, 2021. **34**(9).
88. Palmer, C.S., *Innate metabolic responses against viral infections.* Nature Metabolism, 2022. **4**(10): p. 1245-1259.

89. Chen, Y., et al., *Interaction between N 6-methyladenosine (m 6 A) modification and noncoding RNAs in cancer.* Molecular cancer, 2020. **19**: p. 1-14.
90. Elsabbagh, R.A., et al., *Impact of N6-methyladenosine (m6A) modification on immunity.* Cell Communication and Signaling, 2022. **20**(1): p. 140.
91. Wang, J.-F., et al., *Pathogenic roles of m6A modification in viral infection and virus-driven carcinogenesis.* Endocrine, Metabolic & Immune Disorders-Drug Targets (Formerly Current Drug Targets-Immune, Endocrine & Metabolic Disorders), 2022. **22**(10): p. 1009-1017.
92. Mir, M.A., *Illuminating the pathogenic role of SARS-CoV-2: Insights into competing endogenous RNAs (ceRNAs) regulatory networks.* Infection, Genetics and Evolution, 2024: p. 105613.
93. Chen, J., et al., *Classical swine fever virus non-structural protein 5B hijacks host METTL14-mediated m6A modification to counteract host antiviral immune response.* Plos Pathogens, 2024. **20**(3): p. e1012130.
94. Rehman, U.U., et al., *Epigenetics regulation during virus-host interaction and their effects on the virus and host cell.* Microbial Pathogenesis, 2023: p. 106271.
95. Wright, D.J., et al., *Long read sequencing reveals novel isoforms and insights into splicing regulation during cell state changes.* BMC genomics, 2022. **23**: p. 1-12.
96. Shah, A., et al., *Benchmarking sequencing methods and tools that facilitate the study of alternative polyadenylation.* Genome biology, 2021. **22**: p. 1-21.
97. Gleeson, J., et al., *Accurate expression quantification from nanopore direct RNA sequencing with NanoCount.* Nucleic acids research, 2022. **50**(4): p. e19-e19.
98. Tan, L., et al., *Utilization of nanopore direct RNA sequencing to analyze viral RNA modifications.* Msystems, 2024. **9**(2): p. e01163-23.
99. Zhong, Z.-D., et al., *Systematic comparison of tools used for m6A mapping from nanopore direct RNA sequencing.* Nature Communications, 2023. **14**(1): p. 1906.
100. Leonardi, T. and A. Leger, *Nanopore RNA sequencing analysis.* RNA Bioinformatics, 2021: p. 569-578.

101. Tang, A.D., et al., *Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns.* Nature communications, 2020. **11**(1): p. 1438.
102. Huang, S., A.C. Wylder, and T. Pan, *Simultaneous nanopore profiling of mRNA m6A and pseudouridine reveals translation coordination.* Nature Biotechnology, 2024: p. 1-5.
103. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome biology, 2014. **15**: p. 1-21.
104. Rezinciuc, S., et al., *Dynamic metabolic reprogramming in dendritic cells: An early response to influenza infection that is essential for effector function.* PLoS pathogens, 2020. **16**(10): p. e1008957.
105. Hwang, Y., et al., *Viruses interact with hosts that span distantly related microbial domains in dense hydrothermal mats.* Nature Microbiology, 2023. **8**(5): p. 946-957.
106. Zhang, F., et al., *The novel roles of RNA m6A modification in regulating the development, infection, and oxidative DNA damage repair of Phytophthora sojae.* Plos Pathogens, 2024. **20**(9): p. e1012553.
107. Heaton, S.M., *Harnessing host–virus evolution in antiviral therapy and immunotherapy.* Clinical & translational immunology, 2019. **8**(7): p. e1067.
108. LaPointe, A.T., et al., *Identification and characterization of Sindbis virus RNA-host protein interactions.* Journal of virology, 2018. **92**(7): p. 10.1128/jvi. 02171-17.
109. Bhattacharya, T., et al., *Differential viral RNA methylation contributes to pathogen blocking in Wolbachia-colonized arthropods.* PLoS pathogens, 2022. **18**(3): p. e1010393.
110. Zhang, T., et al., *m6A-express: uncovering complex and condition-specific m6A regulation of gene expression.* Nucleic acids research, 2021. **49**(20): p. e116-e116.
111. Sikorski, V., et al., *The structure and function of YTHDF epitranscriptomic m6A readers.* Trends in Pharmacological Sciences, 2023. **44**(6): p. 335-353.

112. He, P.C. and C. He, *m6A RNA methylation: from mechanisms to therapeutic potential.* The EMBO journal, 2021. **40**(3): p. e105977.
113. Singh, M., et al., *High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes.* Nature communications, 2019. **10**(1): p. 3120.
114. Wu, S. and U. Schmitz, *Single-cell and long-read sequencing to enhance modelling of splicing and cell-fate determination.* Computational and Structural Biotechnology Journal, 2023. **21**: p. 2373-2380.
115. Hook, P.W. and W. Timp, *Beyond assembly: the increasing flexibility of single-molecule sequencing technology.* Nature Reviews Genetics, 2023. **24**(9): p. 627-641.