Early Events in RNA Virus Infection: Profiling the Pre-Replicated Interactome of Chikungunya Virus

By

Sarah Arcos

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

December 12th, 2020

Nashville, Tennessee

Approved:

Manuel Ascano, Ph.D

Scott Hiebert, Ph.D

Martin Egli, Ph.D

Yi Ren, Ph.D

Eric Skaar, Ph.D

John Karijolich, Ph.D

To my loving and brilliant husband, Nick

To my cat, Flumpert

and

To my family

ACKNOWLEDGEMENTS

First, I would like to acknowledge the Chemical Biology of Infectious Diseases Training Grant (T32AI11254) for funding this research.

I would like to thank my mentor, Dr. Manny Ascano, for his guidance throughout this degree. He has provided the support and feedback necessary to make me a better scientist. I appreciate the invaluable time he has spent in training me over the last five years. I would also like to thank the members of my committee for their insights and encouragement. I deeply appreciate their time and interest in my research.

I am grateful to the members of the Ascano lab for always being there to share a cup of coffee or a failed western blot. Much of the work presented here would not have been possible without the unwavering brilliance and mentorship of Dr. Byungil Kim. Byungil, I am thankful for your patient and gracious feedback, and I feel truly privileged to have worked with and learned from you during my time at Vanderbilt. Katie, thank you for your friendship and for being an amazing role model— if I become half as good a scientist as you, I will have succeeded. Caroline, thank you for the nice Havarti, and for your unfailing wit. Sam, thank you for the dance moves and vent sessions, and Robert, thank you for always thinking outside the box. I'm so excited to see what you both achieve with your incredible work ethic and creativity. And thank you to Jeff, for your inspiring perseverance and kindness.

I would also like to thank the many members of the greater Vanderbilt community for creating an environment of collaboration and scholarship, though I will not have the space to thank every individual who took time to help me with my research through providing reagents,

advice, knowledge, or equipment. However, I would like to particularly thank the members of the Hodges lab, in particular Kelly Barnett, Bob Chen, Tyler Hansen, and Lindsey Guerin, for lending me reagents and for help with everything related to R; I would like to thank the Vanderbilt Mass Spectrometry Research Center, especially Dr. W. Hayes McDonald and Dr. Kristie L. Rose of the Proteomics Core, for their assistance and advice in analyzing the VIR-CLASP proteomics data sets; I would like to thank all the members of R Fridays for their helpful code review and commitment to rigorous bioinformatics analysis; I would like to thank the Vanderbilt Institute for Infection, Immunology, and Infectious disease and the leadership of the CBID training grant (especially Dr. Eric Skaar) for creating an outstanding training program for graduate students.

Lastly, and most importantly, I would like to thank my family for their love, encouragement, and patience throughout these last five years. Thank you to my parents, Janet and John, for your support and for always being there. Thank you to my wonderful husband, Nick, for the laughs, hikes, statistics questions, and ice cream. Your creativity, joy, work ethic, and kindness inspire me every day. I am so lucky to have you in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

RNA viruses represent a broad class of human pathogens. They cause diseases from the common cold to hemorrhagic fevers with astonishing mortality. They can spread via air-borne droplets, a mosquito bite, or strictly through prolonged contact with infected blood. Yet the seemingly arbitrary grouping of these pathogens does encapsulate shared biology. The composition of RNA virus genomes makes them uniquely suited to the rapid co-option of cellular machinery. RNA processing components within the host cytoplasm can accommodate many of the RNA metabolism steps of RNA virus replication, with the notable exception of RNA transcription. Recently, host RNA modification metabolism has been revealed as an important determinant of RNA virus replication, and in particular the regulation of the modification N6-methyladenosine (m^6A) (Gokhale and Horner, 2017; Horn and Sarnow, 2017).

The dynamic m^6A RNA modification is found on the genomes of RNA viruses (Gokhale and Horner, 2017) and is the most widespread RNA modification in mammalian mRNA and lncRNA transcripts (Yue et al., 2015). m^6A is involved in regulating many aspects of mRNA metabolism, including stability (X. Wang et al., 2014), translation (X. Wang et al., 2015a), localization (Ries et al., 2019), and splicing (Xiao et al., 2016). Dynamic regulation of the m^6A modification contributes to the coordination of cell-state transitions, such as stem cell reprogramming (Aguilo et al., 2015), heat shock response (Zhou et al., 2015), differentiation (Geula

et al., 2015), and X-chromosome inactivation (Patil et al., 2016). However, comparatively little research has been done on the contribution of m^6A to innate immune activation or viral infection.

The battle between viral RNA and host factors begins the instant that the viral genome enters a cell. Prior to the transcriptional upregulation of interferons, cytokines, and antiviral genes, a cell must rely on mRNAs and proteins that already exist in its cytoplasmic arsenal for protection against infection. The virus also benefits from any mechanism it can utilize to co-opt host processes for replication before the onslaught of new defenses from transcription alters the internal cellular environment. However, the molecular events that comprise these primary interactions between host protein and viral RNA are not well understood. In this dissertation, I will present the development of a new method to identify proteins that directly bind to pre-replicated, primary viral RNA genomes that we term VIR-CLASP for **Vir**al **C**ross-**L**inking **A**nd **S**olid-phase **P**urification (B. Kim et al., 2020). This approach is amenable to essentially any RNA virus and captures interactions that occur within minutes of viral entry. In Chapter III, I will show how this approach identified hundreds of host RNA-binding proteins (RBPs) that interact with the primary Chikungunya virus (CHIKV) RNA genome, including the YTHDF family of N^6-methyladenosine (m^6A) binding proteins. I will then examine how the YTHDF proteins can in turn regulate CHIKV replication. In Chapter IV, I will investigate discrepancies between the activity of the YTHDF proteins *in vitro* and *in vivo*, ultimately demonstrating that these proteins can bind both m^6A-modified and unmodified RNA sites in certain cellular contexts. In Chapter V, I will discuss the impact of these results on our understanding of RNA viruses and RNA modifications, and I will provide future directions for study of these complex systems. In this chapter, I will introduce the concepts of the innate immune response, RNA virus replication (specifically CHIKV), and the m^6A RNA modification as it relates to RNA virus infection.

Overview of the innate immune response to RNA virus infection

The innate immune system comprises the cells and signaling pathways responsible for the rapid detection and response to infection or cellular stress. The sensors and signaling pathways of this system provide a first line of defense upon infection, prior to activation of an adaptive immune response (Murphy et al., 2017). Abnormal function of the innate immune system can lead to increased susceptibility to infection or to autoimmune disease. The traditional approach to studying innate immunity involves characterizing the cytokine and chemokine responses downstream of pathogen recognition or cellular damage. However, less is known about the role of post-transcriptional gene regulation (PTGR) in coordinating this acute transcriptional response. Because many viral pathogens rely on host machinery in order to replicate, regulation of host or viral nucleic acids through the activity of RNA-binding proteins (RBPs) plays a key role in the host response to infection. Cellular RBPs involved in the initial recognition events after viral entry are critical to the speed and effectiveness of a host response, yet this subset of RBP interactions is not well understood.

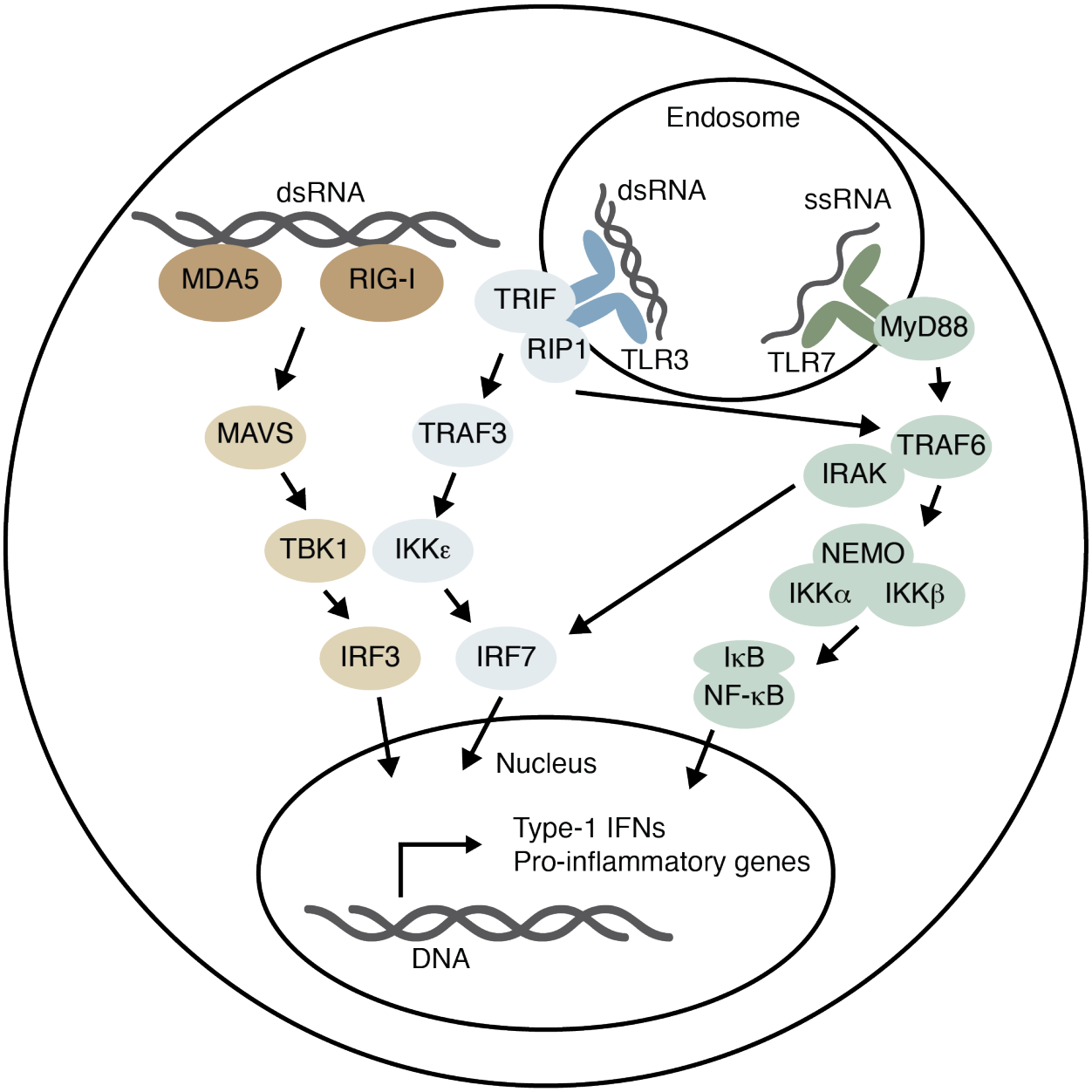


Figure 1.1 RNA sensors in the innate immune system. The main RNA sensors of the innate immune system include MDA5, RIG-I, TLR3, and TLR7. Upon recognition of pathogen RNA, these sensors stimulate intracellular signaling cascades leading to IRF3- or IRF7-dependent upregulation of Type-1 interferons and interferon-stimulated genes. Activation of NF-κB upregulates transcription of pro-inflammatory genes.

Pattern-recognition receptors (PRRs) are an important class of innate immune signaling proteins, and those that recognize RNA-based patterns are also classified as RNA-binding proteins (Nan et al., 2014). The main PRRs responsible for recognizing and responding to RNA virus infection include RIG-I, MDA5, TLR3, and TLR7 (Nan et al., 2014) (Figure 1.1). These receptors act by first recognizing specific molecular patterns on RNA (5' triphosphorylation, double-strandedness, 5' cap structures (or lack of), etc.), then stimulating IRF3- or IRF7-depependent intracellular signaling pathways leading to the transcription and translation of type-1 IFNs. This is followed by upregulation of interferon-stimulated genes (ISGs) including pro-inflammatory cytokines and chemokines.

Yet, the structural similarity of RNA virus genomes to cytoplasmic RNAs challenges the innate immune system, which cannot rely on cytoplasmic localization of alpha-helical structure or a missing hydroxyl group for nucleic acid recognition as with DNA viruses or prokaryotic genomes. Notably, RNA modifications are known mechanisms by which RNA viruses can "mimic" host RNAs to avoid innate immune detection. The 5' 7-methylguanosine cap on mRNA protects host RNAs from recognition by RIG-I; many RNA viruses either encode methyltransferases or use more subtle methods (such as cap-snatching by Influenza A virus) to protect their own viral RNAs with such a cap (Beachboard and Horner, 2016; De Vlugt et al., 2018). Similarly, the poly-A tail, which blocks mRNAs from degradation by exonucleases, has also been adopted by a number of RNA viruses (Jan et al., 2016). Recently, m^6A and has been shown to inhibit the binding and conformational changes necessary for RIG-I activation (Durbin et al., 2016). While this study was *in vitro*, other recent work with HCV and HPMV have shown that m^6A can block RIG-I recognition of RNA viruses in cells (G.-W. Kim et al., 2020; M. Lu et al., 2020). Therefore, it is possible that the m^6A modification could function as a distinguishing mark between host and virus

RNA. Yet given that not all cellular RNAs contain m^6A, this seems unlikely. Existence of multiple m^6A readers with diverse activities (Edupuganti et al., 2017; X. Wang et al., 2015a; 2014; Yuan Zhang et al., 2019) also indicates that there is no single function for m^6A -- thus whether m^6A contributes to, inhibits, or has no effect on immune detection is likely context-dependent. Indeed, recent papers have found that m^6A can both promote and inhibit innate immune signaling during RNA virus infection (Y. Liu et al., 2019; Winkler et al., 2019; Yuan Zhang et al., 2019). Specific details of how m^6A regulates innate immune responses and RNA virus replication can be found in following introductory sections.

RNA viruses also directly inhibit innate immune signaling through a variety of mechanisms including regulating host translation (Walsh et al., 2013), proteolysis of host innate immune proteins (Jurczyszak et al., 2020; Morazzani et al., 2019), and inhibition of mRNA nuclear export (Sandri-Goldin, 2004). In all, these reports draw attention to the ongoing evolutionary battle between host and virus to gain the upper hand during infection. In particular, these studies highlight the importance of RNA-binding proteins as front-line effectors of many pro- and anti-viral strategies.

Chikungunya virus: a re-emerging pathogen

Chikungunya virus (CHIKV) is a single-stranded, positive-sense RNA virus of the family *Togaviridae*; it is transmitted to humans via infected mosquitos. Common symptoms are high fever and debilitating joint pain which can last long after the acute infection resolves (Weaver and Lecuit, 2015). Chikungunya virus (CHIKV) infected roughly 2 million people in the Americas

between 2014 and 2016, after being first reported in the Americas only one year earlier (Silva and Dermody, 2017).

Chikungunya was first identified and isolated from human patients in present-day Tanzania in 1952 (Lumsden, 1955; Robinson, 1955). The name Chikungunya is derived from the Kimakonde word chikungunya, translated as "that which stoops up", which describes the arthritic posture and gait of afflicted individuals (Robinson, 1955). Many of the symptoms of CHIKV infection resemble those of the related arbovirus Dengue, such as high fever, headache, and rash (Carey, 1971; Halstead, 2015). Thus, it is likely that CHIKV outbreaks have occurred as early as 1779 but were mistakenly attributed to Dengue fever; phylogenetic studies indicate that CHIKV first emerged in Africa as long as 500 years ago (Volk et al., 2010). While CHIKV has a low fatality rate ($< 0.01\%$), persistent arthritis and polymyalgia can lead to chronic disability, leading to significant quality of life and socioeconomic considerations (Couturier et al., 2012; Renault et al., 2008). CHIKV infection mostly occurs in Africa, Asia, and India, and despite the increasing prevalence of CHIKV worldwide, there are currently no vaccines or drugs available to treat CHIKV infection (Weaver and Lecuit, 2015).

CHIKV is an enveloped virus with a positive-sense genome of ~11,800 nucleotides (nts) with a 3' poly-A tail and a 5' 7-methylguanosine cap (Khan et al., 2002) (Figure 1.2). The genome encodes for four non-structural protein subunits of the RNA polymerase (nsP1, nsP2, nsP3, nsP4), and three main structural proteins (Capsid; the glycoproteins E1 and E2). The genome also contains untranslated regions at the 3' and 5' ends.

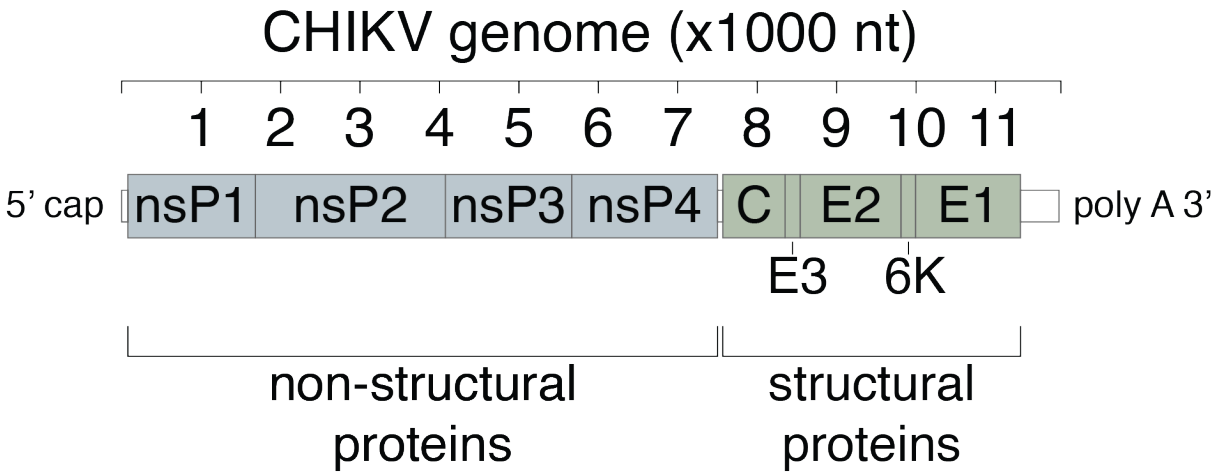


Figure 1.2 The genome structure of Chikungunya virus. The ~11,000 nucleotide CHIKV genome encodes four non-structural proteins and three main structural proteins. The 5' end contains a 7-methylguanosine cap, and the 3' end has a poly-A tail.

CHIKV replication occurs in much the same way as other alphaviruses (Figure 1.3). CHIKV entry requires the viral E2 protein, but the host entry receptor is currently unknown (van Duijl-Richter et al., 2015). After clathrin-mediated endocytosis, the vRNA is released into the cytosol via endosome acidification and insertion of the fusion peptide of E1 into the membrane of the endosome (van Duijl-Richter et al., 2015). Translation of the non-structural proteins results in the P1234 polyprotein, which is then cleaved by a virally-encoded protease in nsP2 to form the individual subunits of the viral polymerase (J. H. Strauss and E. G. Strauss, 1994). The viral polymerase then transcribes the negative-sense RNA, which can then be used as a template for the genomic RNA and the subgenomic RNA encoding the structural proteins (J. H. Strauss and E. G. Strauss, 1994). During translation of the structural proteins, autoproteolysis releases Capsid protein; cleavage of E3-E2-6K-E1 occurs after translation and relies on host proteases (Solignat et al., 2009). After post-translational processing of E3-E2 and release of E3, these proteins assemble into spikes at the plasma membrane, and form particles ready for budding upon association with intact nucleocapsid (Capsid associated with vRNA) (Ryan et al., 1998). CHIKV replication occurs mainly in "replication factories" composed of vesicles that first localize at the plasma membrane early in infection, then move deeper into the cytoplasm as replication progresses (Frolova et al., 2010).

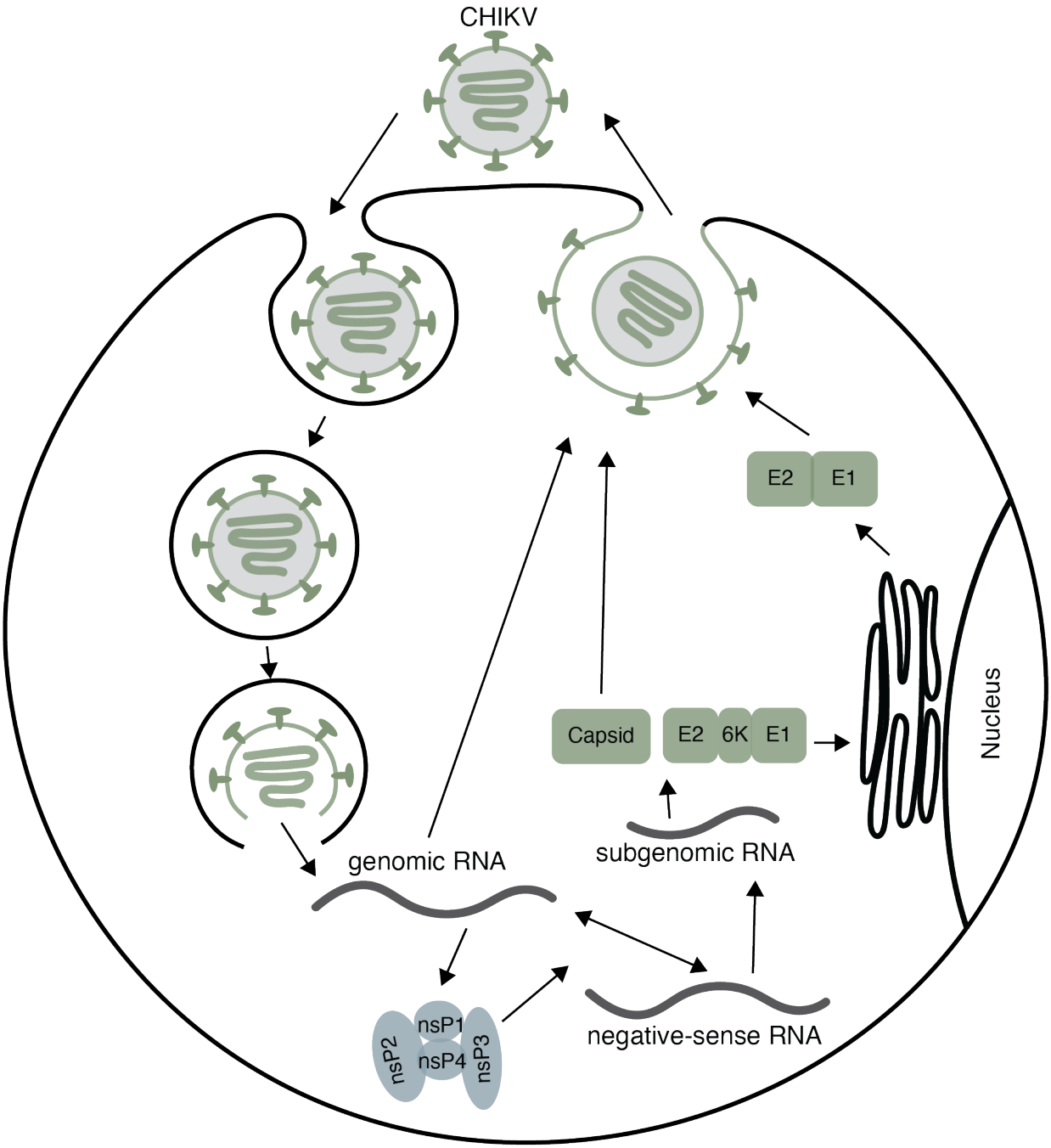


Figure 1.3 The replication cycle of Chikungunya virus in human cells. Chikungunya virus enters human cells via clathrin-mediated endocytosis following interaction of coat proteins with an unknown host receptor. After release of genomic RNA, translation produces the viral RNA polymerase comprised of nsP1-4, followed by transcription of the negative-sense RNA intermediate. The positive-sense subgenomic RNA encodes the structural proteins, which are processed by the ER prior to association with encapsidated genomic RNA and formation of complete viral particles.

RNA viruses use a variety of strategies to evade host innate immunity, but many of these strategies remain a mystery for CHIKV. Upon infection, CHIKV stimulates a rapid and strong innate immune response, characterized by high expression of type-1 IFNs and pro-inflammatory cytokines (Fros and Pijlman, 2016). However, CHIKV is also capable of blocking translation of these IRF3-dependent mRNAs through an unknown mechanism (White et al., 2011), and a decreased IFN response is a known risk factor for severe disease (Couderc et al., 2008). While it is clear that CHIKV is capable of modulating host innate immunity, many of the details regarding this molecular interplay are yet undiscovered.

Current methods to profile RNA-protein interactions

The current techniques that are available to study interactions between host RBPs and viral RNAs are limited in their resolution of the different stages of viral replication. Cross-linking and IP approaches have succeeded in mapping the binding sites of a number of different RBPs along viral RNA genomes. Recently, YTHDF binding sites were mapped on Hepatitis C virus (HCV), a member of the Flaviviridae family, 48 hours post-infection (Gokhale et al., 2016). The identified binding sites therefore represent interactions between YTHDF and the HCV genome both pre- and post-replication. The combination of RNA antisense purification and mass-spectrometry (RAP-MS) was recently used to discover the repertoire of RBPs that interact with the Dengue viral genome (Phillips et al., 2016). However, this study infected cells for 30 hours prior to cross-linking and lysis, leading to identification of RBP-RNA interactions at all stages of viral replication. The insights into the host RBP and viral RNA interactions are invaluable to a comprehensive understanding of how host and virus communicate throughout the viral life cycle.

Yet, the unique events that occur prior to viral replication and before the host cell can mount a full innate immune response remain unknown.

This dissertation will present a technique developed in the Ascano lab (VIR-CLASP) (B. Kim et al., 2020), which addresses this gap by utilizing 4-thiouridine (4SU) labeled virus to infect unlabeled host cells (Figure 1.4). After crosslinking the infected cells with 365nm light just after infection, the RBP-viral RNA interactions are purified using solid-phase reverse immobility (SPRI) beads. The resulting proteins are analyzed by mass-spectrometry, leading to identification of proteins that specifically interact with the infecting, pre-replicated viral genome. VIR-CLASP improves upon the related techniques RAP-MS (Phillips et al., 2016) and thiouracil crosslinking mass-spectrometry (TUX-MS) (Lenarcic et al., 2013) in two important ways. First, VIR-CLASP specifically purifies interactions with the incoming, pre-replicated virion RNAs. Second, VIR-CLASP uses sequence-independent purification based on SPRI bead technology with protein-denaturing buffer conditions. This enables VIR-CLASP to be used with essentially any RNA virus. Our initial experiments with CHIKV yielded identification of both known and novel RNA binding proteins, including the YTHDF family of m^6A-binding proteins.

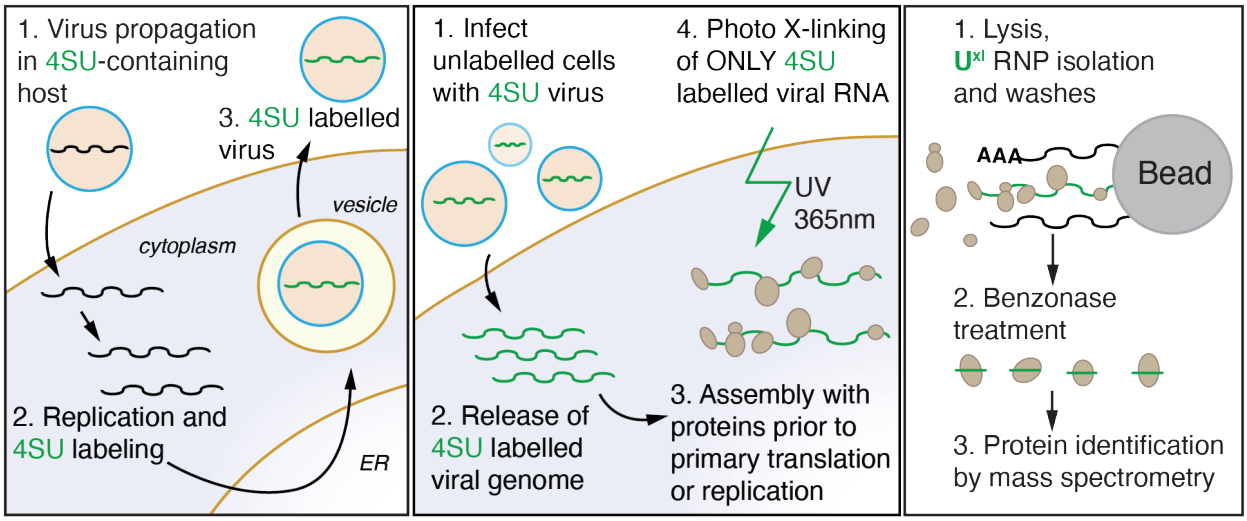


Figure 1.4 VIR-CLASP: Viral Cross-Linking And Solid-phase Purification. VIR-CLASP utilizes 4-thiouridine (4SU) to label the genomes of RNA viruses. Labeled viral particles are then infected into unlabeled host cells. RNA-protein interactions are crosslinked using 365nm light. The crosslinked complexes are purified using SPRI beads, followed by RNA digestion and identification of proteins using LC-MS/MS.

The RNA modification N^6-methyladenosine

N^6-methyladenosine (m^6A) is the most widespread internal RNA modification in eukaryotic mRNA and long non-coding RNA (lncRNA) transcripts (Yue et al., 2015). When m^6A was identified in the 1970s, it was shown to be present not only in cellular RNAs, but also in the genomes of RNA viruses that replicate in the nucleus (Wei et al., 1975). Interest in m^6A waned after this initial discovery due to technological limitations and the belief that the modification was stable throughout the life of a transcript after it was added co-transcriptionally. However, recent identification of proteins that "read", "write", and "erase" m^6A on mRNA and lncRNA have confirmed that this modification can be dynamically regulated (J. Liu et al., 2014) (Figure 1.5). The methyltransferase complex responsible for catalyzing m^6A formation contains two distinct methyltransferases, METTL3 and METTL14, as well as Wilms-tumour associated protein 1 (WTAP), and a number of unidentified subunits (Ping et al., 2014). Two enzymes, fat-mass and obesity associated protein (FTO) and AlkB homolog 5 (ALKBH5) have been identified as demethylases that specifically act on m^6A to break it down into its oxidized forms, N^6-hydroxymethyladenosine and N^6-formyladenosine (Jia et al., 2011; G. Zheng et al., 2013). The YTH-domain family of RNA-binding proteins specifically bind m^6A-modified RNA through a direct interaction between the conserved YTH domain and the modified adenosine (Luo and Tong, 2014) (Figure 1.5).

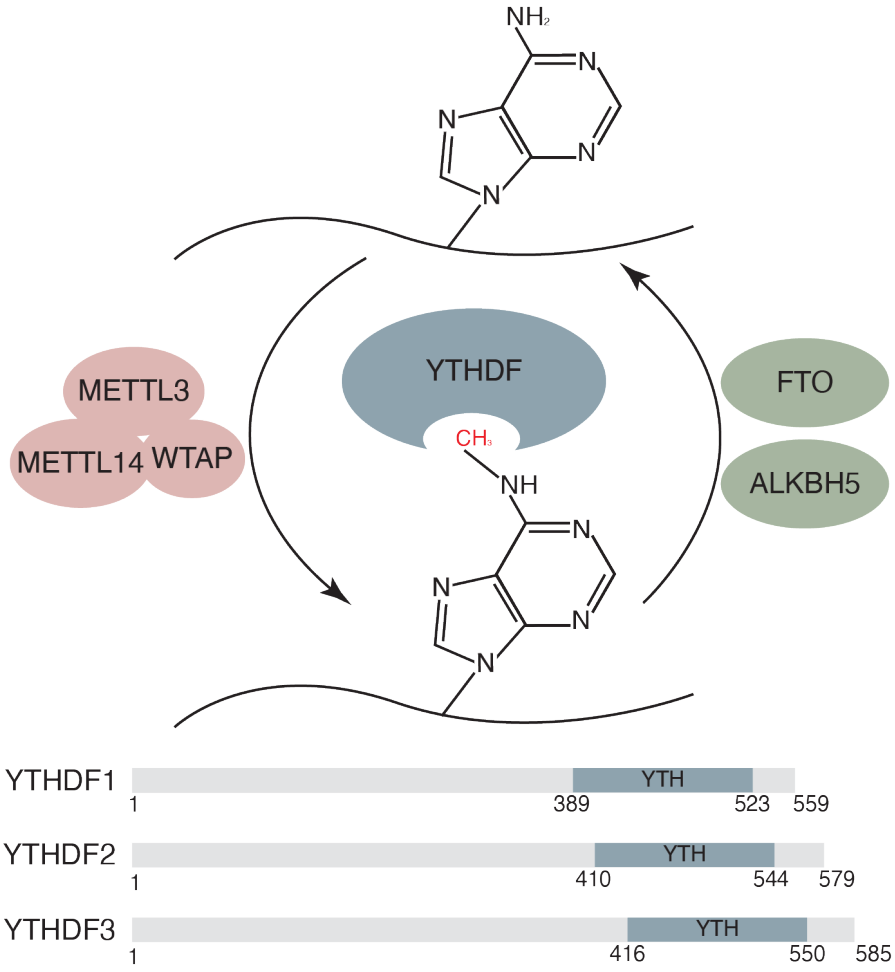


Figure 1.5 The N^6-methyladenosine RNA modification machinery . N^6-methyladenosine (m^{6}A) is formed on RNA by the "writer" complex comprising METTL3, METTL14, WTAP, and other adaptor proteins. m^{6}A is removed from RNA by the demethylase "erasers" FTO and ALKBH5. The main m^{6}A "readers" are YTHDF1-3, which contain a conserved C-terminal YTH domain and an uncharacterized N-terminal region.

Recent studies have investigated how recognition of m^{6}A by YTHDF proteins and other RNA binding proteins may impact the fate of mRNA transcripts during normal cellular function. In all, the YTHDF proteins and other m^{6}A readers have been found to regulate diverse aspects of mRNA metabolism including translation (X. Wang et al., 2015a), stability (X. Wang et al., 2014), or localization (Ries et al., 2019) (Figure 1.6). Characterization of one family member, YTHDF2, revealed that YTHDF2-bound mRNAs have a decreased half-life (X. Wang et al., 2014). However, YTHDF2 binding had no effect on translation initiation or elongation. Though there is debate as to the mechanism of YTHDF2's impact on RNA stability, at least one group has shown that YTHDF2 binding recruits the CCR4-NOT deadenylase complex, leading to a shortening of the mRNA poly-A tail and subsequent degradation (Du et al., 2016). In contrast, YTHDF1 has been shown to alter the efficiency of translation for specific target mRNAs (X. Wang et al., 2015b). This occurs through direct interaction with protein translation machinery, yet YTHDF1 had no measurable impact on mRNA half-life. Another recent paper identified a single, distinct m^{6}A site located in the 5' UTR of some mRNAs that can be directly bound by eukaryotic initiation factor 3 (eIF3) (Meyer et al., 2015). This leads to recruitment of the 43S complex and cap-independent translation. While these studies have illuminated some of the mechanisms by which the YTHDF proteins and other potential m^{6}A-binding proteins can alter the fate of target mRNA transcripts during normal cellular function, the role that m^{6}A and the YTHDF proteins play during dynamic cellular conditions, and in particular viral infection and innate immune activation, is not well understood.

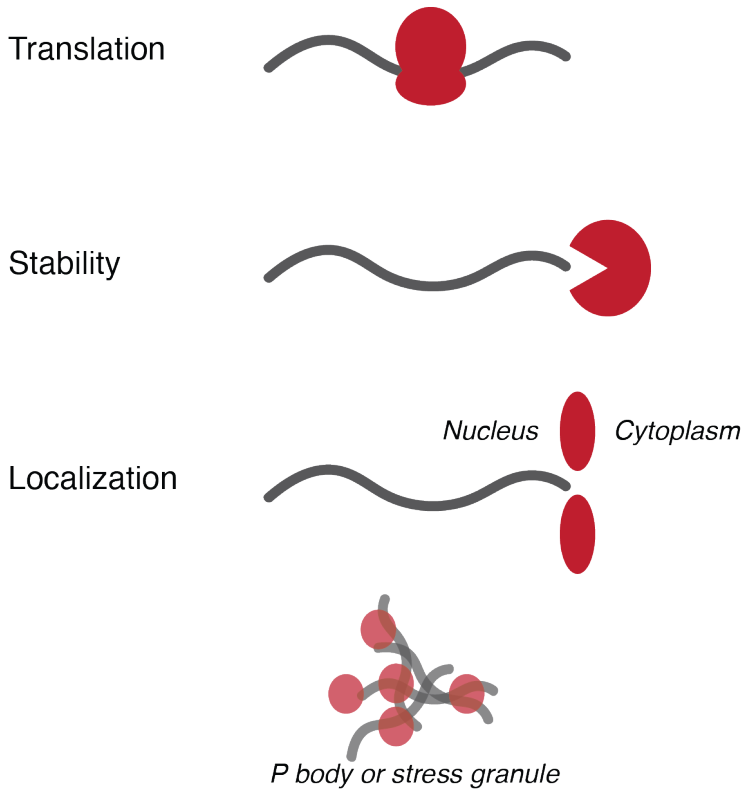


Figure 1.6 Main functions of m^{6}A in RNA metabolism. Through recognition by various m^{6}A readers, m^{6}A can regulate mRNA translation, stability, or localization (nuclear export, association with P-bodies or stress granules).

The roles of YTHDF proteins in RNA virus replication appear to be specific to the virus, viral replication stage, and cell type. Further, these effects may or may not depend on m^{6}A modification. Reports from multiple groups (Gokhale et al., 2016; X. Wang et al., 2015a; 2014; Yuan Zhang et al., 2019) demonstrate that YTHDF proteins do not exclusively bind to modified RNA sequences *in vivo*, in both mRNA and viral RNAs. Some findings show that up to ~80% of YTHDF sites identified by CLIP assays do not overlap with an identified m^{6}A site, and $>$ 55% do not overlap with the canonical m^{6}A motif, RRACH (Gokhale et al., 2016). This calls into question the determinants for YTHDF binding on various RNA species. The functions of large regions of the YTHDF proteins remain poorly characterized (Figure 1.5), partially due to their inability to form crystals (Zhu et al., 2014). Yet disordered regions are known to be important for RNA-binding in a number of other mammalian proteins (Calabretta and Richard, 2015). Therefore, it is possible that these disordered regions within YTHDF proteins may contribute to the observed non-m^{6}A binding in cells. YTHDF proteins form protein-protein interactions with a range of binding partners (Du et al., 2016; X. Wang et al., 2015a), which could enable YTHDF proteins to interact with unmethylated RNAs. In all, while there is substantial evidence for YTHDF binding of unmethylated RNA *in vivo*, the mechanisms and outcomes of these binding events are unexplored.

m^{6}A is present on the genomes of many RNA viruses, including Human Immunodeficiency virus (HIV) (Kennedy et al., 2017; Lichinchi et al., 2016a; Tirumuru et al., 2016), Influenza A virus (IAV) (Courtney et al., 2017; Krug et al., 1976), Flaviviruses (Gokhale et al., 2016; Lichinchi et al., 2016b), and others. Yet the ultimate effect of m^{6}A on the replication of these viruses appears to be specific to both virus and cell-type. In the remainder of this chapter I will profile the diverse roles of m^{6}A and its writers, readers, and erasers in RNA virus infection, paying specific attention

to how these viruses attain m^{6}A modifications and the diverse effects of m^{6}A on viral replication, translation, and infectious particle production. Through careful consideration of historical and recent studies, I will show that 1) it is unlikely that the canonical m^{6}A methyltransferase complex is solely responsible for RNA virus methylation, 2) studies of "global" m^{6}A effects may be useful for development of potential anti-viral drugs, but have limited biological insight due to the combinatorial or competing activities of m6A readers, and 3) the regulation of RNA virus infection by m^{6}A is highly dependent upon the virus, cell type, strain, and stage of replication.

The role of m^{6}A in RNA virus replication

Mechanisms of m^{6}A deposition on RNA virus genomes

m^{6}A RNA modifications are not strictly deposited in the nucleus. This model is supported by the existence of RNA viruses that replicate exclusively in the cytoplasm, which do not encode their own methyltransferases, yet contain genomic m^{6}A (Gokhale et al., 2016; Lichinchi et al., 2016b). The mechanism of m^{6}A deposition on cytoplasmic RNAs is an open question. One model proposes that the methyltransferases METTL3 and METTL14 can re-locate to the cytoplasm under cellular conditions such as viral infection (Gokhale et al., 2016; Hao et al., 2019), or that in some cell lines they are present in the cytoplasm under steady-state conditions (Lichinchi et al., 2016b). Although the mechanism remains unclear, the knockdown of METTL3 or METTL14 during viral infection decreases m^{6}A in released virions of HCV (Gokhale et al., 2016). This model is also supported by the observation that knockdown of METTL3, METTL14, or ALKBH5 were each able to increase or decrease methylation of specific adenosines in the RRE element of HIV-

1, respectively (Lichinchi et al., 2016a). Another possibility is that other uncharacterized cellular methyltransferases are responsible for cytoplasmic methylation. Recent research has shown that the enzyme METTL16 may participate in methylation of specific host RNAs (U6 snRNA) (Pendleton et al., 2017), though whether it or other methyltransferases contribute to RNA virus methylation is unknown.

Further evidence for separate methylation pathways used by different RNA viruses comes from early studies of 3-deazaadenosine (3DAA) and its carbocyclic analog (CC-3DAA). Though these drugs are chemically similar, they have vastly different antiviral activities. 3DAA potently inhibits the nuclear-replicating viruses IAV and IBV (Influenza B Virus), while having no effect on the cytoplasmic-replicating viruses WNV (West Nile Virus), SFV (Semlicki Forest Virus), PV (Poliovirus), or NDV (Newcastle Disease Virus) (Fischer et al., 1990). The potency of 3DAA on IAV is also strain-specific. CC-3DAA has no or little effect on IAV or IBV, yet inhibits the replication of MMV (Measles Morbillivirus), RSV (Respiratory Syncitial Virus), and PIV3 (Parainfluenza Virus Type 3) (de Clercq and Montgomery, 1983; Wyde et al., 1990). Both drugs inhibit cellular methylation reactions but through different mechanisms: 3DAA is metabolized to 3-deazaadenosyl homocysteine (3DAA-Hcy), which then can bind to and inhibit methyltransferases as an analog of s-adenosyl homocysteine (SA-Hcy); SA-Hcy participates in product inhibition of methyltransferases under normal cellular conditions (Backlund et al., 1986; Montgomery et al., 1982) (Figure 1.7). CC-3DAA acts by inhibiting Hcy hydrolase, resulting in an accumulation of SA-Hcy levels, and increased product inhibition of methyltransferases (Backlund et al., 1986; Montgomery et al., 1982) (Figure 1.7). One proposed model to explain these differing effects is that the analog 3DAA-Hcy can only bind to *specific* methyltransferases, while SA-Hcy can bind and potentially inhibit all methyltransferases (Fischer et al., 1990). If this

is correct, then it suggests that not all viruses acquire m^6A via the same process: either they use different methyltransferases, or they rely on some other process such as nucleoside recycling of methylated adenosines.

Recycling of methylated adenosines from mRNA degradation may contribute to m^6A on RNA virus genomes. If m^6AMP can be recycled and incorporated into nascent RNAs, then it must be an efficient substrate of nucleotide kinases, and m^6ATP must be an efficient substrate of RNA-dependent RNA polymerases. Work in Arabidopsis thaliana, while of course having limited relevance to mammalian systems, does indicate that m^6AMP can be recycled into nascent mRNA in plants (Chen et al., 2018). m^6ATP as a substrate of RNA virus polymerases has not been tested, but due to the ready production of in vitro transcribed RNAs containing m^6A using RNA polymerases from various species (Meyer, 2019; Molinie et al., 2016), there is precedent for incorporation of m^6ATP using similar enzymes. If m^6ATP is indeed incorporated into viral RNA in human cells, future studies should examine whether or how selectivity of m^6A incorporation is achieved on viral transcripts, and how m^6ATP impacts human RNA polymerase processivity, error rate, or termination. Whether recycled m^6ATP can be trafficked to the nucleus to contribute to m^6A in mammalian mRNA is also unknown.

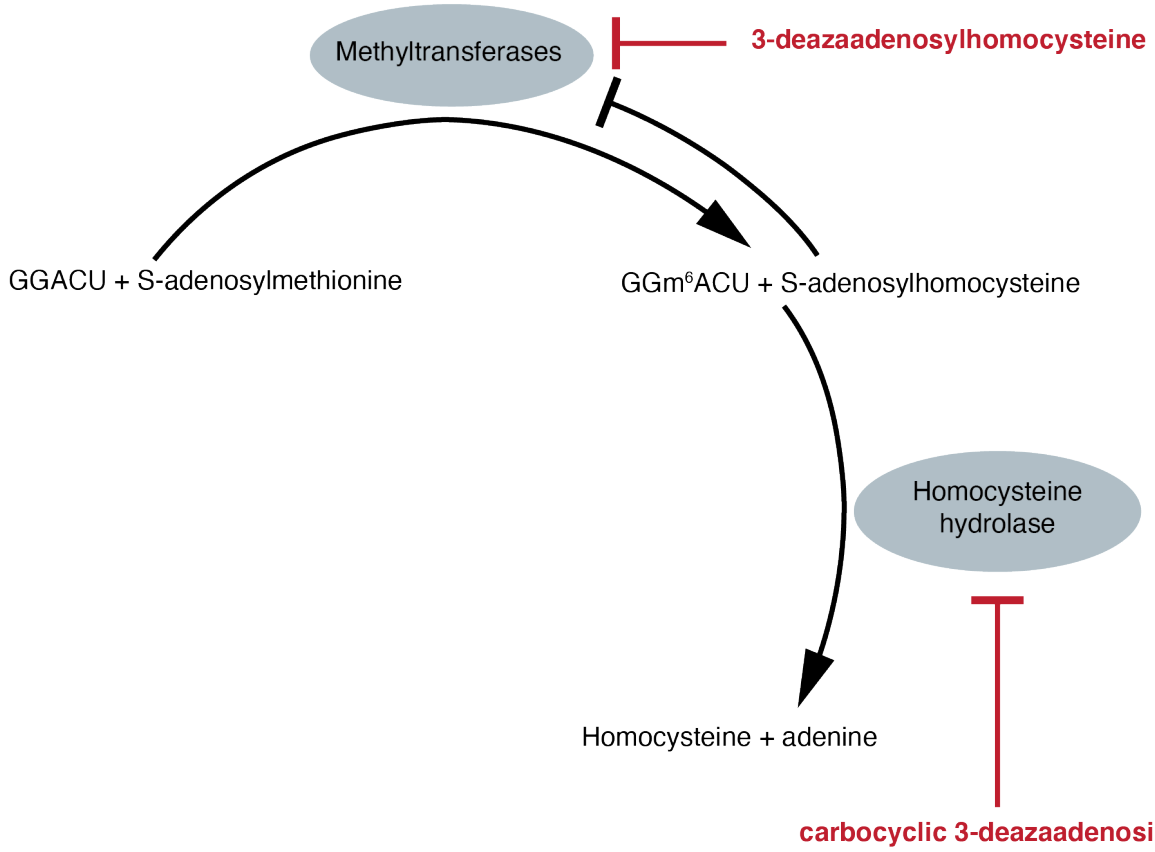


Figure 1.7 Mechanism of 3-deazaadenosine and its carbocyclic analog in inhibiting cellular methylation. 3-deazaadenosine is metabolized to 3-deazaadenosylhomocysteine, which inhibits cellular methylation as an analog of S-adenosylhomocysteine. Carbocyclic 3-deazaadenosine cannot be metabolized, and thus inhibits cellular methylation by inhibiting homocysteine hydrolase and increasing the concentration of S-adenosylhomocysteine.

The role of m^{6}A in Retrovirus replication

Studies of HIV-1 have revealed diverse m^{6}A locations and functions based on the cell line or the stage of viral replication. Three groups profiled the location of m^{6}A modifications along the HIV-1 genome in different human cell lines or primary cells (Kennedy et al., 2017; Lichinchi et al., 2016a; Tirumuru et al., 2016) (Figure 1.8A). All groups used the HIV-1 strain pNL3-4, thus the discrepancy in m^{6}A sites is likely due to the different cell lines used, the time post infection, or the technique used. While no peak is perfectly shared among all four m^{6}A maps, they all contain peaks in close proximity within the Nef coding region.

Each group used genetic manipulation or reporter assays to discover that global m^{6}A promotes most stages of HIV-1 replication. Lichinchi et al showed that the m^{6}A sites within the RRE stem loop II in the Env gene contribute to nuclear export and promote HIV-1 replication (Lichinchi et al., 2016a). Interestingly, they also observed that knockdown of METTL3 in HEK293T cells led to increased Rev expression by western blot, while the same knockdown in MT4 cells led to decreased gp120 expression (qRT-PCR) and p24 expression (western blot), indicating either cell-type or gene-specific responses to m^{6}A (Lichinchi et al., 2016a). Tirumuru et al found that silencing of the m^{6}A writers in HEK293T cells decreased Gag protein levels, extracellular p24 levels, and release of infectious viral particles (Tirumuru et al., 2016). Kennedy et al also found that m^{6}A in the HIV-1 3' UTR enhanced HIV replication (Kennedy et al., 2017). Together, these studies reveal that globally, m^{6}A promotes HIV-1 replication in most cell lines and at most, but not all, stages of viral replication. However, these studies also demonstrate that m^{6}A does not have a single function on HIV RNA: specific m^{6}A sites in the HIV genome can have unique functions in HIV replication.

A recent study of potential g-quadruplex forming sequences (PQSs) found that a number of PQSs overlap with m^6A sites in the HIV-1 3' UTR (Fleming et al., 2019). Both the m^6A sites and the PQSs are highly conserved. PQSs are capable of blocking the viral polymerases of ZIKV and HCV when exposed to a stabilizing drug (Fleming et al., 2016; Jaubert et al., 2018; S.-R. Wang et al., 2016). Whether m^6A drives or inhibits G-quadruplex formation is unknown, but the interplay of m^6A sites and PQSs in the HIV 3' UTR may contribute to the enhanced translation and replication observed by Kennedy et al.

Experiments to understand the role of the YTHDF m^6A readers in HIV-1 replication have yielded contrasting results. While Kennedy et al found that YTHDF2 and YTHDF3 enhance HIV-1 protein and RNA levels in HEK293T cells, YTHDF1 only did so to a minor extent (Kennedy et al., 2017). In contrast, Tirumuru et al observed that the YTHDFs inhibit HIV-1 infectivity in HeLa cells (Tirumuru et al., 2016). Silencing of the YTHDFs in Jurkat or primary CD4+ T cells promoted HIV-1 infectivity, with YTHDF1 and YTHDF3 having a much stronger effect than YTHDF2 (Tirumuru et al., 2016). However, Gag mRNA levels in Jurkat cells were only affected by YTHDF3, not YTHDF1 or YTHDF2 (Tirumuru et al., 2016). These conflicting results illustrate the importance of context (such as cell type, time post infection, or MOI) in understanding the role of m^6A and its reader proteins in HIV-1 infection.

It is clear that observations based on perturbation of m^6A at a global level will be influenced by the combinatorial activities of m^6A readers both known and undiscovered. One model that may explain some of these discrepancies is that YTHDF proteins can have different roles based on the stage of viral replication. Lu et al propose that YTHDF proteins enhance release of infectious HIV-1 viral particles, while inhibiting HIV-1 replication immediately after infection into target cells (W. Lu et al., 2018). To test this model, they silenced the YTHDF proteins in

HEK293T cells and found that while YTHDF1 and YTHDF3 decreased Gag and processed p24 protein levels, the produced viral particles were more infectious in target cells compared to control (W. Lu et al., 2018). While the mechanism is unclear, this indicates that YTHDFs may have roles in first inhibiting HIV-1 replication, then in facilitating HIV-1 viral packaging. Future studies should investigate whether this switch in YTHDF function is due to host factors, viral factors, or changes in m^6A location on HIV-1 RNA.

It will be important for all future m^6A-seq studies to be explicit in the viral replication stage captured and the cell lines (both propagating and target) used. In fact, if the differing observations of YTHDF function are ultimately due to changes in host factors, then use of pseudotyped rather than wild-type HIV-1 may explain the discrepancies. This is due to the interesting observation by Lichinchi et al that HIV-1 infection leads to an increase in host m^6A levels (Lichinchi et al., 2016a). In 2019, Tirumuru et al followed up on this observation and discovered that the mechanism for this increase in host m^6A is binding of gp120 to the CD4+ receptor: application of exogenous gp120 was sufficient to increase host m^6A, but infection with VSV-G pseudotyped HIV-1 or with heat-inactivated HIV-1 was not (Tirumuru and Wu, 2019). Thus, comparisons between YTHDF function on HIV must take into account not only cellular context, but also the pseudotype of the virus.

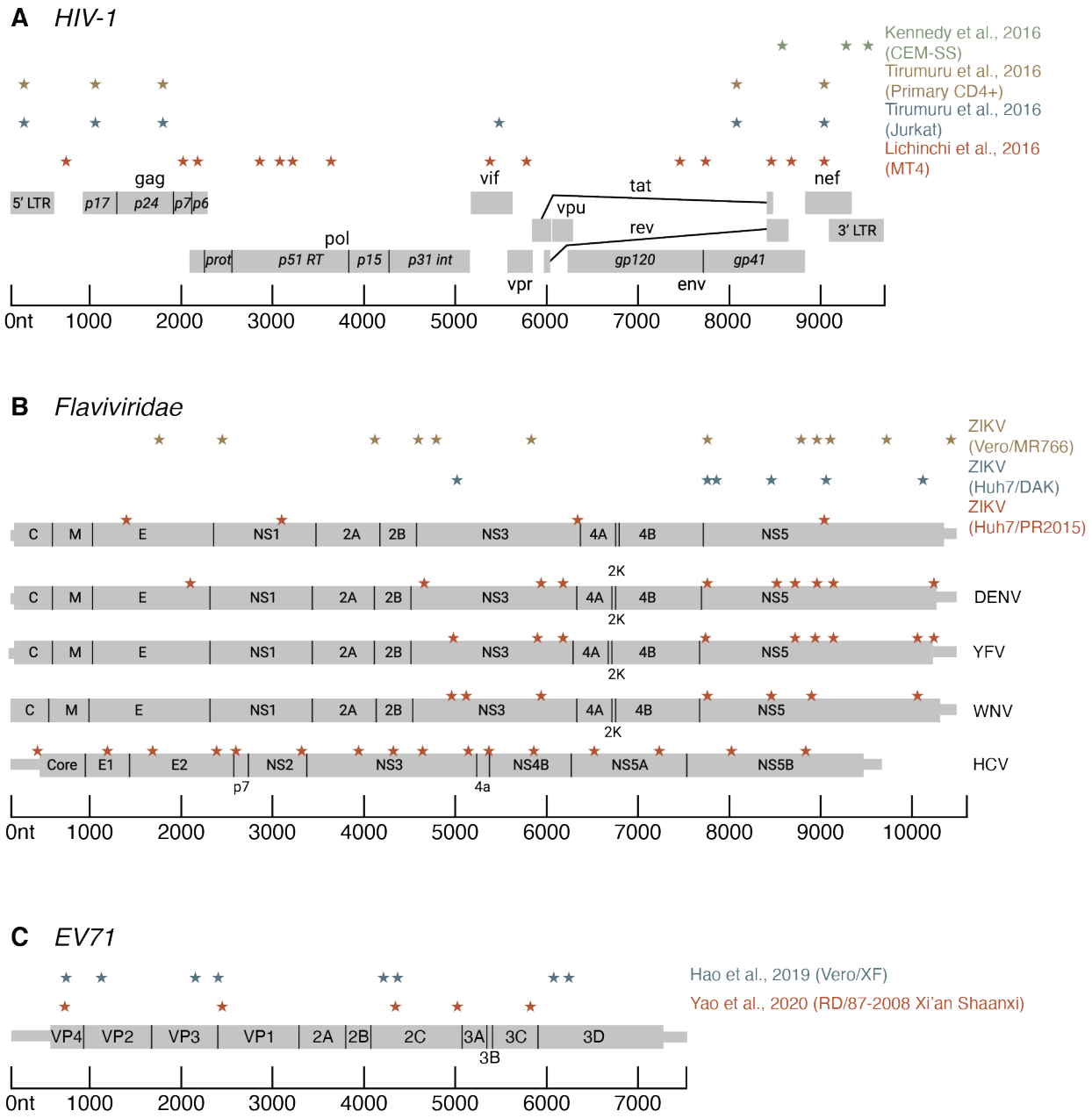


Figure 1.8 Cell-type and strain specific location of m^{6}A on RNA viruses. A) m^{6}A modifications on the HIV-1 genome post-infection into human cells (listed on right hand side). Stars indicate m6A peak, genome structure is shown below. B). m^{6}A modifications on Flavivirus genomes indicated as in (A). For DENV, YFV, WNV, and HCV, viruses were infected into Huh7 cells. For ZIKV, cell line and strain are listed on right hand side C) m^{6}A modifications on Enterovirus 71 (EV71) genomes as in (A). Cell lines and strains are listed on right hand side. Strain 87-2008 Xi'an Shaanxi is a clinical isolate. In (A,C), m^{6}A peak locations were derived from published datasets, and references are indicated in the figure. In B, all peaks were derived from (Gokhale et al., 2016) except for ZIKV (Vero/MR766), which was derived from (Lichinchi et al., 2016).

The role of m^{6}A in negative-strand RNA virus replication

IAV was one of the first RNA viruses found to contain genomic m^{6}A modifications (Krug et al., 1976). Early work using 3DAA established that IAV replication is promoted or inhibited at various lifecycle stages by methylation (Fischer et al., 1990). Treatment with 3DAA inhibited translation of late proteins like HA, while stimulating translation of mRNA that encodes early proteins (Fischer et al., 1990). In particular, translation of NS2 mRNA was potently stimulated by 3DAA treatment (Fischer et al., 1990). 3DAA also enhanced splicing of the M mRNA (Fischer et al., 1990). Yet overall, 3DAA decreased production of infectious IAV, though the extent of this inhibition was strain specific (Fischer et al., 1990). As noted above, the analog CC-3DAA had no activity against IAV or IBV (Wyde et al., 1990), indicating that the effects of 3DAA may be due to a specific methyltransferase, rather than inhibition of all cellular methylation.

Thirty years later, Courtney et al made use of technological advances to map the m^{6}A sites along the IAV genome and mRNA (Courtney et al., 2017). They found that m^{6}A-abrogating mutations to the HA gene resulted in decreased expression of HA mRNA and protein (consistent with 3DAA), and a 20%-40% drop in HA incorporation into released virions (Courtney et al., 2017). However, virus spread was not inhibited and IFNB1 levels in the target cell were unchanged (Courtney et al., 2017). Courtney et al also examined whether YTHDF proteins promote or inhibit IAV replication. They observed that overexpression of YTHDF2, but not YTHDF1, led to increased IAV protein and mRNA levels, and production of infectious viral particles (Courtney et al., 2017). Together, these experiments illustrate that m^{6}A has varying effects on different stages of IAV replication, due to the competing or combinatorial behavior of the m^{6}A reader proteins.

A recent paper confirmed that METTL3 promotes IAV mRNA expression through knockdown experiments, and further showed that that this effect can be rescued by treatment with ruxolitinib, an inhibitor of IFNB signaling (Winkler et al., 2019). This finding expands upon the previous report to suggest that the effects of knockdown or overexpression of m^6A machinery on IAV replication may be due to activities in host mRNA in addition to activity directly with IAV RNA. This observation may explain why 3DAA but not knockdown of METTL3 was able to decrease infectious IAV particle production. An important next experiment would be to examine whether the effects of 3DAA at different stages of IAV replication can also be rescued by ruxolitinib treatment.

m^6A and the YTHDF proteins promote Human metapneumovirus (HMPV) replication and infectious particle production (M. Lu et al., 2020). Overexpression of YTHDF3 and YTHDC1 increased G protein by 12 hpi, with YTHDF1 and YTHDF2 also having an effect by 18 hpi (M. Lu et al., 2020). All YTHDFs and YTHDC1 increased N protein and viral titers (M. Lu et al., 2020). One mechanism for how m^6A promotes HPMV replication was by inhibiting recognition by RIG-I: abrogating mutations in both the genome and antigenome led to increased binding by RIG-I, induction of the RIG-I conformational change, and increased IFNB mRNA levels (M. Lu et al., 2020).

Respiratory syncytial virus (RSV) contains m^6A modifications in virion RNA and in its post-infection gRNA and mRNA (Xue et al., 2019). Overall, m^6A in RSV was found to increase viral replication and infectivity, as demonstrated by genetic manipulation of the m^6A writers and erasers, and construction of m^6A-abrogating mutants (Xue et al., 2019). The abrogation mutants also revealed striking differences based on whether infection was in cell culture, human airway epithelial cultures (HAEs) or cotton rats. In cell culture, a single abrogating mutation had a

stronger negative effect on VSV replication than a combined double mutant (Xue et al., 2019). In HAEs and cotton rats however, the double mutant was substantially more attenuated than either single mutant (Xue et al., 2019). This observation directly supports the model that individual m^{6}A sites can have unique function, since their effects are not additive in all contexts.

No studies have directly examined the function of m^{6}A in vesicular stomatitis virus (VSV), though multiple recent groups have found roles for m^{6}A in regulating innate immunity (Yuan Zhang et al., 2019; Q. Zheng et al., 2017) or host cellular metabolism (Y. Liu et al., 2019) during VSV infection. Importantly, at least one group has confirmed by m^{6}A-seq that VSV RNA does contain m^{6}A (Y. Liu et al., 2019).

The role of m^{6}A in positive-strand RNA virus replication

The role of m^{6}A in flavivirus replication has been extensively studied, revealing a diversity of biological effects even within the same virus family (Gokhale et al., 2016; Lichinchi et al., 2016b) (Figure 1.8B). Knockdown of the m^{6}A methyltransferase complex increased HCV infectious particle production and NS5A protein levels (Gokhale et al., 2016). In contrast, knockdown of the YTHDF reader proteins did not change NS5A levels or HCV replication, but did increase infectious particle production (Gokhale et al., 2016). Finally, m^{6}A abrogating mutations in the E1 gene led to a 3-fold increase in infectious particle production (Gokhale et al., 2016). This illustrates that m^{6}A and the YTHDF proteins specifically inhibit the latest stages of the HCV lifecycle.

Maps of m^{6}A modifications (Figure 1.8B) and YTHDF binding sites on the HCV genome revealed that few YTHDF binding sites overlap with m^{6}A modifications (Gokhale et al., 2016). Interestingly, Rios-Marco et al also observed that the YTHDF proteins can bind to *in vitro*

transcribed (and thus m^{6}A -free) HCV (Ríos-Marco et al., 2016). These observations suggest that YTHDF proteins may be capable of binding unmodified sites on HCV RNA during infection. It is possible that other protein factors or other domains may alter the binding preferences of the YTHDF proteins *in vivo*.

m^{6}A on the HCV genome was also shown to inhibit recognition by RIG-I leading to more efficient HCV replication, in contrast to the inhibition of infectious HCV production by YTHDF proteins (G.-W. Kim et al., 2020). Specifically, an m^{6}A site located 100nt upstream of a RIG-I PAMP in the 3' end of the HCV genome (Figure 1.8B) recruits YTHDF2 to block RIG-I recognition of HCV RNA (G.-W. Kim et al., 2020). An abrogating mutation in this site led to increased IFNB mRNA and phosphorylation of IRF3, and this effect was not enhanced by knockdown of METTL3 or METTL14; this indicates that the increase in IFNB signaling was due specifically to this m^{6}A site, and not to other m^{6}A sites in the HCV genome (G.-W. Kim et al., 2020). This is an important observation, as it demonstrates that studies of "global" m^{6}A in RNA virus infection do not capture the diversity of m^{6}A functions in viral or host RNAs. This study also found that overexpression of YTHDF2 and YTHDF3 could increase IFN signaling upon HCV infection, but YTHDF1 had no effect (G.-W. Kim et al., 2020). From these studies, it is clear that the YTHDF proteins have diverse roles in the HCV lifecycle and in host innate immunity. Identification of the proteins that interact with the YTHDFs at different stages of viral replication may yield insight into how the YTHDF proteins achieve such different roles

m^{6}A modification maps have also been constructed for other members of the *Flaviviridae* (Gokhale et al., 2016) (Figure 1.8B). These viruses (HCV, DENV, WNV, YFV, ZIKV) have a conserved m^{6}A peak in the NS3/NS5A gene, and distinct m^{6}A peaks along the remainder of their genomes (Gokhale et al., 2016). It is important to note that ZIKV showed strain- or cell-type-

specific differences in m^{6}A-modification sites (Gokhale et al., 2016; Lichinchi et al., 2016b). As with HCV, m^{6}A and the YTHDF proteins appear to increase ZIKV replication and infectious particle production (Lichinchi et al., 2016b). Knockdown of METTL3 led to increased ZIKV infectious particle production, and decreased m^{6}A density on the ZIKV genome (Lichinchi et al., 2016b). While this data supports the model that m^{6}A on ZIKV is added by the canonical cellular m^{6}A machinery, this effect could also be indirectly due to a depletion of m^{6}ATP recycled from host mRNAs. Knockdown of the YTHDF proteins also increased ZIKV infectivity, with YTHDF2 having a much stronger effect than YTHDF1 or YTHDF3 (Lichinchi et al., 2016b).

The recent study of potential G-quadruplex forming sequences also identified many PQSs on the ZIKV genome (Fleming et al., 2019). While the positive-sense ZIKV genome contained 70 PQSs, 7 of which were shared with all flaviviruses, the negative-sense strand contained zero PQSs (Fleming et al., 2019). The significance of G-quadruplex restriction to the positive-sense strand is unknown. Of the 12 m^{6}A peaks identified by Lichinchi et al, 8 overlapped with a PQS (Lichinchi et al., 2016b). Future studies should examine whether PQS context favors methylation or demethylation, and conversely whether m^{6}A stabilizes or destabilizes G-quadruplex structures. As noted above, stabilized G-quadruplexes are capable of inhibiting ZIKV RNA polymerase (Fleming et al., 2016). Thus, if m^{6}A drives or inhibits G-quadruplex formation, this may have implications in ZIKV replication.

One of the most recent RNA viruses to join the m^{6}A club is Enterovirus 71, the main causative agent of hand foot and mouth disease in eastern Asia. Two recent papers mapped the m^{6}A sites on both a laboratory strain (Hao et al., 2019) and a clinical isolate (Yao et al., 2020) of EV71 (Figure 1.8C). These maps show strain- or cell-type-specific and shared m^{6}A sites along the EV71 genome. Genetic manipulation of the m^{6}A machinery revealed striking cell-type differences

in YTHDF function. Knockdown of the YTHDFs in monkey-derived Vero cells increased viral copy numbers, while knockdown of YTHDF2 or YTHDF3 in human-derived RD cells led to decreased viral replication (Hao et al., 2019). Both papers found that abrogating mutations or knockdown of the methyltransferases led to decreased EV71 replication (Hao et al., 2019; Yao et al., 2020). In all, studies of m^6A in EV71 demonstrate the importance of cell-line considerations when interpreting the results of genetic manipulations of m^6A machinery components.

Viruses with indirect evidence of m^6A modification

Indirect evidence of m^6A on RNA viruses refers to evidence not based on HPLC, LC-MS, or m^6A antibodies. Such evidence includes sensitivity to 3DAA or CC-3DAA (Ebola (Bray et al., 2000), HPIV3 (Wyde et al., 1990), Measles (de Clercq and Montgomery, 1983), Poliovirus (Montgomery et al., 1982)) or computational prediction (SARS-CoV-2 (Jin et al., 2020)). Importantly, and with apologies for the necessary triple negative, insensitivity to 3DAA or its analogs is not sufficient to demonstrate a lack of m^6A, as WNV contains m^6A (Gokhale et al., 2016) yet was not inhibited by 3DAA (Fischer et al., 1990).

CHAPTER II

MATERIALS AND METHODS

Primers

Table 2.1 Primer and siRNA sequences used in this study. The tagged sequences used in ssqRT-PCR oligomers are underlined. The AT-rich flap sequences are shown in lowercase. The Att sequences used in the Gateway cloning system are shown in boldface. F, Forward primer; R, Reverse primer.

RT-qPCR with SYBR green		
Target	**Name**	**Sequence**
TUBA1A	TUBA1A qt F	GAGCGTCCAACCTATACTAACC
	TUBA1A qt F	GCAGCAAGCCATGTATTTACC
CHIKV	CHIKV qt F	AAAGGGCAAGCTTAGCTTCAC
	CHIKV qt R	GCCTGGGCTCATCGTTATTC
Hamster_*Gapdh*	Ham_Gapdh qt F	GACTTCAACAGTGACTCCCAC
	Ham_Gapdh qt R	TCTGTTGCTGTAGCCAAATTC
18S rRNA	18S rRNA qt F	GTAACCCGTTGAACCCCATT
	18S rRNA qt R	CCATCCAATCGGTAGTAGCG
28S rRNA	28S rRNA qt F	GGGTGGTAAACTCCATCTAAGG
	28S rRNA qt R	GCCCTCTTGAACTCTCTCTTC
YTHDF1	YTHDF1 qt F	GGACACCCAGAGAACAAAAGG
	YTHDF1 qt R	AGGAAATCCAATGGACGGCG
YTHDF2	YTHDF2 qt F	TAGCCAACTGCGACACATTC
	YTHDF2 qt R	CACGACCTTGACGTTCCTTT
YTHDF3	YTHDF3 qt F	TGACAACAAACCGGTTACCA
	YTHDF3 qt R	TGTTTCTATTTCTCTCCCTACGC

CHIKV (for 1000 nt)	CHIKV qt 981 F	GGTATGCGGTAACCCACCAC
	CHIKV qt 1042 R	GTCAACCGTGTCGGTAGTCTT
CHIKV (for 2000 nt)	CHIKV qt 1954 F	GGTGTACAACGAAAGAGAGTTCG
	CHIKV qt 2034 R	TCTTCGTCAGTGTTCAGGGC
CHIKV (for 3000 nt)	CHIKV qt 2958 F	TCTCTGGTGACCCGTGGATA
	CHIKV qt 3040 R	CTCCACCTCCCACTCCTTAAT
CHIKV (for 4000 nt)	CHIKV qt 3975 F	TATTTAGCAATTTTGACAATGGCAG
	CHIKV qt 4046 R	AGGCTGCATTCAGTTGATTGTT
CHIKV (for 5000 nt)	CHIKV qt 4932 F	TGTGTTCTTCATTTCCCCTTCCA
	CHIKV qt 5073 R	TCCTGTACAGACTCCTGGGAA
CHIKV (for 6000 nt)	CHIKV qt 5947 F	AGTCCCGACTTATCGGACCA
	CHIKV qt 6044 R	CATTACATGCTGCCACTGCG
CHIKV (for 7000 nt)	CHIKV qt 6927 F	GCTTCAAGTTCGGCGCTATG
	CHIKV qt 7029 R	AGACGATCTTCCAACACCCG
CHIKV (for 8000 nt)	CHIKV qt 7972 F	AAGCCAGCACACGTAAAGGG
	CHIKV qt 8061 R	CGCGCATTCAAGGTCGTACT
CHIKV (for 9000 nt)	CHIKV qt 8930 F	ACCACGACCCTCCTGTGATA
	CHIKV qt 9047 R	ACCTCTATCTCCTCGGCAGT
CHIKV (for 10000 nt)	CHIKV qt 9955 F	TTAGCCGTAATGAGCGTCGG
	CHIKV qt 10062 R	GCCCGGTCTGTTGACTAGAG
CHIKV (for 11000 nt)	CHIKV qt 10977 F	TGCGGTGCATTCGATGACTA
	CHIKV qt 11072 R	AGGGCCGTCGAAAAAGAGAT
Strand specific qRT-PCR with Taqman		
Target	**Name**	**Sequence**
	CHIVK (-) RT with tag	GGCAGTATCGTGAATTCGATGCCGTGTGC TGTTCTCAGTA
	CHIVK (+) RT with tag	GGCAGTATCGTGAATTCGATGCACGAAAC CACTGTGTCA
	CHIKV Taqman probe	CATCGGTGTTCCATCTAAAGG
CHIKV (-) strand	RT tag qt	aataaatcataaGGCAGTATCGTGAATTCGATGC
	CHIKV (-) qt R	aataaatcataaACGAAACCACTGTGTCA

CHIKV (+) strand	CHIKV (+) qt F	aataaatcataaCGTGTGCTGTTCTCAGTA
	RT tag qt	aataaatcataaGGCAGTATCGTGAATTCGATGC
siRNA		
Target	**Name**	**Sequence**
YTHDF1	YTHDF1 si F	CCGCGUCUAGUUGUUCAUGAAUU
	YTHDF1 si R	UUCAUGAACAACUAGACGCGGUU
YTHDF2	YTHDF2 si F	AAGGACGUUCCCAAUAGCCAAUU
	YTHDF2 si R	UUGGCUAUUGGGAACGUCCUUUU
YTHDF3	YTHDF3 si F	AUGGAUUAAAUCAGUAUCUAAUU
	YTHDF3 si R	UUAGAUACUGAUUUAAUCCAUUU
Plasmid construction		
Target	**Name**	**Sequence**
YTHDF1	YTHDF1 F	**GAAGACAAGTTTGTACAAAAAAGCAGGCT** TAATGTCGGCCACCAGC
	YTHDF1 R	**GAAGACCACTTTGTACAAGAAAGCTGGGT** TTAGCATTGTTTGTTTCGACTCTG
YTHDF2	YTHDF2 F	**GAAGACAAGTTTGTACAAAAAAGCAGGCT** TAATGTCGGCCAGCAGC
	YTHDF2 R	**GAAGACCACTTTGTACAAGAAAGCTGGGT** TTATTTCCCACGACCTTGACG
YTHDF3	YTHDF3 F	**GAAGACAAGTTTGTACAAAAAAGCAGGCT** TAATGTCAGCCACTAGCGTGG
	YTHDF3 R	**GAAGACCACTTTGTACAAGAAAGCTGGGT** TTATTGTTTGTTTCTATTTCTCTCCCTAC

Software

Table 2.2 Software packages used in this study. Main software packages used for analysis of mass-spectrometry and RNA-sequencing datasets are listed in this table with version numbers (where appropriate) and links to documentation or download.

Package Name	Version	Documentation/Download
MaxQuant	1.5.035	https://www.maxquant.org

R	3.6.0 – 4.6.1	https://www.r-project.org
limma	3.44.3	10.18129/B9.bioc.limma
clusterProfiler	3.16.1	10.18129/B9.bioc.clusterProfiler
DOSE	3.14.0	10.18129/B9.bioc.DOSE
STAR	2.7.2	https://github.com/alexdobin/STAR
UpSetR	1.4.0	https://cran.r-project.org/web/packages/UpSetR/index.html
Trim-galore	0.6.0	https://github.com/FelixKrueger/TrimGalore
PARpipe	NA	https://github.com/ohlerlab/PARpipe
MACS2	2.2.4	https://github.com/macs3-project/MACS
ggplot2	3.2.0	https://ggplot2.tidyverse.org
DEseq	1.39.0	10.18129/B9.bioc.DESeq
Cuffdiff	2.0	http://cole-trapnell-lab.github.io/cufflinks/cuffdiff/
RCAS	(development)	https://github.com/BIMSBbioinfo/RCAS

Cell lines and culture

All cell lines except Huh-7 (male) were obtained from ATCC. BHK-21 cells (male), Vero cells (female), A549 cells (male) and HEK-293T cells (female) were maintained in DMEM (Gibco) supplemented to contain 10% fetal bovine serum (FBS from Peak Serum) and 2 mM L-Glutamine (Gibco). U2OS cells (female) were cultured in McCoy's 5A medium (Gibco) with 10% FBS and 2 mM L-Glutamine. CRISPR knock-in cell line Huh-7 HA-EMC4 (clon: DU12C3) and knockout cell line Huh-7 EMC4 KO (clon: DU13A1) were generated as previously described(Barrows et al., 2019).

Cells were treated with the following reagents: Dinaciclib (HY-10492) was from MedChemExpress; Actinomycin D (A9415) and Cerulenin (219557) were from Sigma; Ruxolitinib Phosphate (sc-396768) was from Santa Cruz Biotech.

Viruses

CHIKV strain 181/25 was provided by Terence S. Dermody (University of Pittsburgh School of Medicine). CHIKV was propagated in BHK-21 cells with or without 1mM 4-thiouridine (4SU). Virus stocks were purified by ultracentrifugation of clarified supernatants through a 20% sucrose cushion in TNE buffer (50 mM Tris-HCl [pH 7.2], 0.1 M NaCl, and 1 mM EDTA) at ~125,000 × *g* for 4 hr in a Beckman SW32Ti rotor. To remove remaining free 4SU, virus pellets were washed three times with TNE buffer, and then re-suspended in virus dilution buffer (DMEM medium containing 10 mM HEPES [Gibco] supplemented to contain 1% FBS), aliquoted, and stored at −70°C. Virus titers were determined by plaque assay using Vero cells.

Encephalomyocarditis virus (EMCV; VR-129B) and zika virus (ZIKV; strain: PRVABC59; VR-1843) were obtained from ATCC, murine hepatitis virus (MHV; strain: MHV-A59) was provided by Mark R. Denison (Vanderbilt University Medical Center), influenza A virus (IAV; strain: California/7/2004 (H3N2)), vesicular stomatitis virus (VSV; strain: Indiana) and rift valley fever virus (RVFV; strain: MP-12) were provided by Thomas G. Voss and James E. Crowe (Vanderbilt Vaccine Center). EMCV, ZIKV, VSV and RVFV were propagated in Vero cells; MHV was propagated in DBT-9 cells; IAV was propagated with TPCK trypsin (2 µg/ml, Pierce) in MDCK cells. 4SU was added during propagation, and the viruses were purified as with CHIKV.

VIR-CLASP

For VIR-CLASP, cells were infected with 1000 MOI of non-labeled or 4SU-labeled virus for 1 hr at 4°C and uninfected virus was washed away with cold PBS. The infected cells were

incubated for the indicated times at 37°C, prior to 365 nm ultraviolet irradiation. For pretreatment with interferon, IFN was added to the media 16 hr before viral infection.

For CLASP, 100 μM 4SU with or without IFN (500 U/mL) was added to the cells and incubated for 16 hr before irradiation. One 15-cm plate was used for each CLASP experiment.

$UV_{365\ nm}$ crosslinking

To irradiate with UV_{365nm}, the growth medium was removed and cells were washed with PBS. Cells were irradiated on ice with 365 nm UV light (0.6 J/cm^2 x 2 times) in a Stratalinker 2400 (Stratagene). Cells were scraped off in 2.5 ml PBS per plate.

CLASP

Cells were lysed in denaturation buffer (50 mM Tris–HCl, pH 6.8, 10% glycerol, 2.5% SDS, 0.66% NP-40), incubated for 10 min at 95 °C and subsequently slowly cooled to 25°C. Crosslinked RNA-protein complexes were purified by Solid-Phase Reversible Immobilization (SPRI) (Hawkins et al., 1994) beads (GE Healthcare, cat# 65152105050250) under denaturing SPRI buffer (30 mM Tris–HCl, pH 6.8, 6% glycerol, 1.5% SDS, 0.4% NP-40, 1 M NaCl, 8% PEG-8000). To each sample, 0.66× (e.g. 660 μl of beads for 1ml of sample) of SPRI beads (1mg/ml SPRI beads in 10 mM Tris-HCl, pH 8.0, 1 M NaCl, 18% PEG-8000, 1 mM EDTA and 0.055% Tween 20) were added, and samples were incubated at room temperature for 10 minutes. The SPRI beads and complexes were washed 5 times with denaturing SPRI buffer. The crosslinked RNA-protein complexes were eluted

for 5min at 37 °C in denaturation buffer (lysis buffer). To reduce non-specific binding on the beads, SPRI purification was repeated.

Benzonase digestion of RNA from crosslinked RNA-protein complexes

An equal volume of 4x Benzonase buffer (80 mM Tris-HCl, pH 7.5, 600 mM NaCl, 20 mM $MgCl_2$, 4 mM DTT, 40% Glycerol) and 2x volume of water were added to eluted samples, followed by the addition of Benzonase (EMD Millipore, cat# 70746-4) to a final concentration of 50 U/ml, and incubation for 2 hr at 37°C. Proteins were precipitated by methanol and chloroform and then re-suspended in 2x NuPAGE LDS Sample Buffer (Thermo Fisher Scientific, cat# NP0007) with 50 mM DTT.

Proteinase K digestion of protein from crosslinked RNA-protein complexes

To measure the efficiency of total RNA purification, with 10% of eluted samples from CLASP, 0.1mg/ml proteinase K (3115879001, Sigma) was treated in proteinase K buffer (50 mM Tris-HCl, pH 7.5, 6.25 mM EDTA-NaOH, pH 8.0, 75 mM NaCl and 1% SDS) for 2 hr at 55°C. RNA was purified using TRIzol™ LS Reagent (Ambion). The RNA was resolved on agarose gel, visualized (ChemiDoc™ MP), and analyzed using ImageJ.

Viral titer with UV_{365nm} crosslinking

Plaque assay with UV_{365nm} crosslinking for CHIKV

Plaque assay with UV_{365nm} crosslinking was performed on U2OS cells. U2OS cells were plated 1 day before. IFN was added to the media 16 hr before viral infection. CHIKV was serially diluted in DMEM medium with 1% FBS and 10mM HEPES and absorbed to U2OS cells for 1 hr at 4 °C. After plates were washed with PBS, infected cells were incubated in cell growth medium at 37 °C until UV_{365} nm crosslinking. At the indicated time point, the growth medium was removed and cells were washed with PBS. Cells were irradiated on ice with 365 nm UV light (0.6 J/cm^2 x 2 times) in a Stratalinker 2400 (Stratagene). Then, cells were overlaid with DMEM with 0.6% SeaPlaque™ Agarose (Lonza, cat# 50100) containing 5% FBS and incubated for 72 hr. Cells were then fixed with 4% paraformaldehyde and stained with 0.1% crystal violet. The resulting plaques were counted.

Mass spectrometric analysis

Gel lanes were cut and diced into 1mm^3 cubes. Proteins were treated for 30 minutes with 45 mM DTT, and available Cys residues were carbamidomethylated with 100mM iodoacetamide for 45 minutes. Gel pieces were further destained with 50% MeCN in 25mM ammonium bicarbonate, and proteins were digested with trypsin (10ng/uL) in 25mM ammonium bicarbonate overnight at 37°C. Peptides were extracted by gel dehydration with 60% MeCN, 0.1% TFA, the extracts were dried by speed vac centrifugation, and reconstituted in 0.1% formic acid. Peptides were analyzed by LC-coupled tandem mass spectrometry (LC-MS/MS). An analytical column was packed with 22cm of C18 reverse phase material (Jupiter, 3 μm beads, 300Å, Phenomenox) directly into a laser-pulled emitter tip. Peptides were loaded on the capillary reverse phase analytical

column (360 µm O.D. x 100 µm I.D.) using a Dionex Ultimate 3000 nanoLC and autosampler. The mobile phase solvents consisted of 0.1% formic acid, 99.9% water (solvent A) and 0.1% formic acid, 99.9% acetonitrile (solvent B). Peptides were gradient-eluted at a flow rate of 350 nL/min, using a 120-minute gradient. The gradient consisted of the following: 1-98 min, 2-40% B; 98-108 min, 40-95% B; 108-110 min, 95% B; 110-111 min, 95-2% B; 111-120 min (column re-equilibration), 2% B. A Q Exactive Plus mass spectrometer (Thermo Scientific), equipped with a nanoelectrospray ionization source, was used to mass analyze the eluting peptides using a data-dependent method. The instrument method consisted of MS1 using an MS AGC target value of 3e6, followed by up to 15 MS/MS scans of the most abundant ions detected in the preceding MS scan. The MS2 AGC target was set to 1e5, dynamic exclusion was set to 20s, HCD collision energy was set to 28 nce, and peptide match and isotope exclusion were enabled.

Bioinformatics analysis of Mass Spectrometry data

Raw data files from the LC-MS/MS instrument were processed and searched using MaxQuant (v1.5.0.35) (Cox and Mann, 2008) to generate peak lists and identify peptide-spectrum matches. Searches were performed using a Uniprot/Swissprot database for *Homo sapiens* with only reviewed proteins included (downloaded on Feb. 28th, 2018), with added sequences for Benzonase nuclease (Uniprot #P13717), and for CHIKV proteins (Capsid, E1, E2, E3, 6k, nsP1, nsP2, nsP3, nsP4) from strain 181/25 (TSI-GSD-218), and IAV proteins (H9XN78, 79, 80, 81, 83, 84, 85). For the purified CHIKV viral particles, the database used was the Uniprot reference proteome for *Mesocricetus auratus* (UP000189706- downloaded July 26th, 2019, with one protein sequence per gene) with sequences added for CHIKV proteins (Capsid, E1, E2, E3, 6k, nsP1, nsP2,

nsP3, nsP4) from strain 181/25 (TSI-GSD-218). The search parameters for Andromeda were: full tryptic specificity, two missed cleavages allowed, carbamidomethyl (C) fixed modification, and acetylation (N terminal) variable modification. Match between runs was selected, and LFQ normalization was performed in separate parameter groups (+IFN and +4SU, -IFN and +4SU, +IFN and -4SU, and -IFN and -4SU). All other settings used were default, resulting in a protein FDR of $< 1\%$ for each dataset. To define the set of pre-replicated CHIKV interacting proteins, we computed peptide intensity ratios between +4SU and -4SU samples for proteins with at least two distinct peptides (Garcia-Moreno et al., 2019; Sysoev et al., 2016). The average log2-intensity ratio for each protein was then tested to be different from 0 using the moderated t-test implemented in the R/bioconductor package limma (Smyth, 2004). p-values were then corrected for multiple testing using the Benjamini-Hochberg method. Proteins with an adjusted p-value $<$ 0.01 and a fold change greater than five were classified as "VIR-CLASP RBPs". Due to the low complexity of the -4SU samples, peptide intensity ratios could not be calculated for all proteins. For proteins with intensity values of zero in the -4SU samples, we performed a semiquantitative method that makes the assumption that peptides with no intensity values are below the detection threshold (Garcia-Moreno et al., 2019; Sysoev et al., 2016). This approach determines the number of replicates for +4SU and -4SU samples in which a peptide has an intensity value. For CHIKV, this leads to a matrix of 12 different groups, for peptides that were detected in 0, 1, 2, or 3 +4SU samples and 0, 1, or 2 -4SU samples. For IAV, the matrix contains 9 different groups, representing the 2 replicates performed for +4SU and -4SU each. The FDRs were estimated as described (Sysoev et al., 2016), as the ratios resulting from the division of the transposed matrix. A protein is determined to be a "candidate VIR-CLASP RBP" if it comprises peptides found in cells with an FDR < 0.01.

Correlations between replicates were performed using Pearson correlation and the "pairwise.complete.obs" option in the base R cor function to allow for missing values. Mapping of GO biological process, molecular function, and cellular component, and KEGG pathways was performed using the R/bioconductor packages clusterProfiler (Yu et al., 2012) and DOSE (Yu et al., 2015). The background used was all human genes, and statistical significance was determined using the Fisher's exact test and Benjamini-Hochberg method to correct for multiple testing. Analysis of enriched PFAM domains was performed using the DAVID Bioinformatics Resource (D. W. Huang et al., 2009a; 2009b) using the Fisher's exact test and Bonferroni correction for multiple testing. Visualization of overlaps between datasets was performed in R using the UpSetR package (Conway et al., 2017) for UpSet diagrams, and the VennDiagram package for 2-way Venn diagrams. All other data visualizations were made using the R package ggplot2 (Wickham, 2016). All code for the R analysis can be found at https://github.com/saraharcos.

Virion RNA sequencing and alignment

CHIKV was purified by ultracentrifugation through a 20% sucrose cushion in TNE buffer as described in methods section of viruses. CHIKV was re-suspended in TNE buffer and further purified twice by discontinuous sucrose gradient (30%-60%) in TNE buffer at ~125,000 × g for 3 hr in a Beckman SW32Ti rotor, and concentrated by 30% sucrose cushion. Trizol reagent (Ambion) was added to the virus pellets and viral RNA was extracted. Total RNA was converted into cDNA and sequenced using NEBNext DNA Library Prep Kit for Illumina on the Illumina NovaSeq6000 platform using PE150 at the Vanderbilt Technologies for Advanced Genomics (VUMC VANTAGE). Fastq files were pre-processed with trim-galore with the default settings

(http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) to remove any adapter contamination and then aligned to the Chikungunya virus strain genome (TSI-GSD-218, GenBank: L27661.3) and *Mesocricetrus auratus* (GCA_000349665.1) using STAR(Dobin et al., 2013) with the default settings. Reference annotation file (GTF) does not exist file for Chikungunya virus strain 181/25 (TSI-GSD-218).

Antibodies and immunoblotting

The antibody to YTHDF1 (anti-YTHDF1; 17479-1-AP) was from Proteintech; anti-FASN(3180), anti-GM130 (12480) and KDM3B (3100) were from Cell Signaling; anti-YTHDH2 (ab170118), anti-YTHDH3 (ab103328), anti-ELAVL1 (ab200342) and anti-TUBA4A (ab7291) were from abcam; anti-IFI16 (sc-8023) and LARP1 (sc-515873 were from Santa Cruz Biotechnology; anti-Flag (F1804) was from Sigma-Aldrich; anti-HA (901502) and WDHD1 (630302) were from BioLegend; anti-FXR1 (MAB2160) was from Millipore.

Samples were separated by SDS-PAGE. After electrophoresis, proteins were semi-dry transferred (Bio-Rad) to nitrocellulose membranes (Hybond-ECL, GE Life Science). Protein membranes were taken through a standard immunoblot protocol followed by enhanced chemiluminescent detection (Luminata Forte ECL, Millipore) using a chemiluminescence imaging system (ChemiDoc MP, Bio-Rad).

Methylated RNA immunoprecipitation (MeRIP)

CHIKV was purified by ultracentrifugation through a 20% sucrose cushion in TNE buffer as described in method section of viruses. CHIKV was re-suspended in TNE buffer and further purified twice by discontinuous sucrose gradient (30%-60%) in TNE buffer at ~125,000 × g for 3 hr in a Beckman SW32Ti rotor and concentrated by 30% sucrose cushion. Trizol reagent (Ambion) was added to the virus pellets and viral RNA was extracted. The RNA was fragmented by RNaseT1 treatment (0.8 U/µl) for 15 min at 22°C. The fragmented RNA was dephosphorylated with CIP and radiolabeled with T4 PNK and [γ-^{32}P]-ATP. The RNA was incubated with m^{6}A antibody (202 003, Synaptic systems) for 2 hr at 4 °C and pulled down by magnetic protein A beads (Invitrogen). The RNA was eluted in Trizol and separated by 12% UREA gel and analyzed by phosphorimaging.

N^6-methyladenosine (m^{6}A) mapping

CHIKV was purified and immunoprecipitated with the m^{6}A antibody described above, with the following changes: the purified CHIKV RNA was treated with RNaseT1 (0.1 U/ µl) for 5 min at 4°C. The fragmented RNA was then extracted with phenol:chloroform and immunoprecipitated with the m^{6}A antibody. The RNA was eluted in Trizol, reverse transcribed using SuperScript III (ThermoFisher) with random hex primers, and quantified by RT-qPCR (see Methods section RT-qPCR analysis) using primers tiled across the CHIKV genome in 1000 nt segments (Table 2.1). Target Ct values were normalized to Input Ct values. Relative enrichment of CHIKV regions in the m^{6}A IP over the IgG IP was then calculated using the $\Delta\Delta$Ct method ($2\Delta\Delta$Ct).

Plasmids and stable cell lines

The Flag-HA-tag lentiviral inducible expression vector pLenti CMVtight Blast Flag-HA-DEST was constructed by insertion of Flag-HA-tag from pFRT_TO_DEST Flag-HA (#26361, Addgene) into the plasmid pLenti CMVtight Blast DEST (w762-1) (#26434, Addgene). The genes *YTHDF1* (NM_017798.3), *YTHDF2* (NM_001172828.1), and *YTHDF3* (NM_152758.4) were amplified by PCR from the plasmid HsCD00378781 for *YTHDF1*, HsCD00376823 for *YTHDF2* and HsCD00376805 for *YTHDF3* (the DNA Resource Core at Harvard Medical School), with gene-specific primers containing an attB sequence. Using the Gateway cloning system (Invitrogen), the amplified genes were cloned into pLenti CMVtight Blast Flag-HA-DEST vector.

For producing lentivirus, the plasmids were transfected with packaging vectors psPAX2 (12260, Addgene) and pMD2.G (12259, Addgene) by Lipofectamine 2000 (Invitrogen) in HEK-293T cells.

To generate inducible expression cell lines, first a stable tetracycline transactivator (rtTA) expressing U2OS cell line was created by transducing a lentivirus from the plasmid rtTA-N144 (66810, Addgene). Then the cDNA expressing lentiviruses were transduced into rtTA expressing U2OS cells and drug selected.

siRNA knockdown

siRNAs used in this study are listed in Table 2.1. siRNAs were transfected at a final concentration of 20 nM using Lipofectamine® RNAiMAX (Invitrogen) according to the manufacturer's instructions. Cells were incubated at 37 °C and 5% CO2 for 72 h before infection

with indicated MOI. For rescue experiments, a stable cell line of Flag-HA tagged YTHDF1 was used.

RT–qPCR analysis

RNA was collected from infected cells using Trizol (Ambion). The RNA concentration was determined using a NanoDrop 2000 (ThermoFisher). Equal amounts of total RNA were reverse-transcribed using SuperScript III (ThermoFisher) with random hex primers or CHIKV strand-specific primers (Plaskon et al., 2009). Real-time PCR reactions were done with FastSYBR Green Plus Master Mix (Applied Biosystems) or TaqMan® Fast Advanced Master Mix (Applied Biosystems) using a StepOnePlus qPCR machine (Applied Biosystems). Oligonucleotides used in this study are listed in Table 2.1. Target Ct values were normalized to *TUBA1A* Ct values and used to calculate ΔCt. Relative mRNA expression of target genes was then calculated using the ΔΔCt method (2ΔΔCt) except cellular RNA expression in CHIKV virion was calculated using the ΔCt method. To generate a standard curve, a plasmid containing a cDNA copy of the CHIKV strain 181/25 (pSinRep5-181/25ic, #60078, Addgene) was used.

Quantification and statistical analysis

Using the GraphPad PRISM 8 software, a two-tailed Student's t-test and one-way ANOVA (Tukey) were used for statistical analysis of all data presented except mass-spectrometry, RNA-seq, and bioinformatics analysis. Numbers of biological replicates of assays (n) are provided

and defined within the corresponding figures or figure legends. Error bars shown in the Figures represent means ± SD.

PAR-CLIP data analysis

Raw data were downloaded from the Gene Expression Omnibus using accession numbers GSM1197607 and GSM1197608. Raw sequencing data was processed, and peaks were called using PARpipe (https://github.com/ohlerlab/PARpipe) which trims adapters, aligns sequences using Bowtie V0.12.7 (Langmead et al., 2009) (to human genome version hg19), and finally calls peaks using the PARalyzer peak caller, V1.5 (Corcoran et al., 2011). A final set of YTHDF2 peaks was defined as the shared peaks between two biological replicates (requiring at least 1nt of overlap). For peaks with multiple overlaps, the T-C fraction used was the maximum value among the multiple overlaps. For all analyses after Figure 4.1, PAR-CLIP peaks were filtered for a T-C fraction > 0.20.

m^{6}A-seq data analysis

Raw data were downloaded from GEO using accession numbers GSM1135030 and GSM1135031 for input, and GSM1135032 and GSM1135033 for m^{6}A-IP. For identification of m^{6}A peaks, the m^{6}A-enriched regions of m^{6}A-IP over input were defined using the model-based analysis of ChIP-seq (MACS2) peak-calling algorithm (Yong Zhang et al., 2008), using parameters "--nomodel --extsize 100 --gsize 100e6". For m^{6}A peak calling with "relaxed" statistical cut-off for

Figure 4.1C, the additional parameter "--qvalue 0.20" was supplied. A final set of m^6A peaks was defined as the shared peaks between two biological replicates (requiring at least 1nt of overlap).

Ribosome profiling and RNA half-life data analysis

Semi-processed ribosome profiling and RNA half-life data were downloaded from the GEO using accession numbers GSM1197612 – GSM1197633. RPKM was calculated using DEseq (Anders and Huber, 2010) for ribosome profiling and Cuffdiff V2.0 (Trapnell et al., 2013) for mRNA half-life. Only genes with RPKM > 1 were considered for further analysis.

For ribosome profiling, the fold change in ribosome protected fragments was calculated by log2(siYTHDF2/siControl). Two biological replicates were used, and the reported fold change is the average of log2(siYTHDF2/siControl) for both replicates.

For calculation of mRNA half-life, RKPM values were converted to attomole by linear-fitting of an RNA spike-in. Then, the degradation rate k of the mRNA was estimated by:

$$log_2\left(\frac{A_t}{A_0}\right) = -kt \qquad (2.1)$$

where t is time after transcription inhibition(h), and A_t and A_0 are mRNA in attomoles at time t and time 0. The final mRNA half-life was calculated as the average of two k values (time 3 h versus time 0 h, time 6 h versus time 0 h).

$$t_{\frac{1}{2}} = \frac{2}{k_{3h} + k_{6h}} \qquad (2.2)$$

Statistical significance for RNA half-life and ribosome profiling was calculated using the Wilcoxon rank-sum test.

Integrative analysis of YTHDF2 sequencing data

The difference in distribution of T-C fraction for m^6A and non-m^6A YTHDF2 peaks was tested to be statistically significant using the Wilcoxon rank-sum test. *De novo* motif finding was performed using the function "runMotifDiscovery" from the R package RCAS (Uyar et al., 2017), using the parameters "sampleN = 10000, motifWidth = 6, motifN = 4". YTHDF2 coverage across genomic feature boundaries was plotted using the R package RCAS(Uyar et al., 2017) with the parameters "sampleN = 0, flankSize = 500". GO analysis was performed using the R package clusterProfiler (Yu et al., 2012) using default parameters and with the background of all human genes. The top five terms by statistical significance are plotted for each GO category. Overlaps with protein domains and secondary structure were calculated using databases downloaded from the UCSC genome browser Table Browser tool (Karolchik et al., 2004), with a minimum overlap of 1nt. The tracks used were "Uniprot Structure (unipStruct)" and "Uniprot Domains (unipDomain)" (UniProt Consortium, 2019).

CHAPTER III

DISCOVERY OF WIDESPREAD HOST INTERACTIONS WITH THE PRE-REPLICATED GENOME OF CHIKV USING VIR-CLASP

Introduction

Emerging viruses threaten human and livestock populations across the globe. Thus it is imperative to understand differences in viral tropism across species and cell types. The earliest interactions between host protein and RNA viral genomes during an infection can determine viral tropism yet are unexplored. Before upregulation of interferons, cytokines, and antiviral genes, a cell must rely on mRNAs and proteins already in its cytoplasmic arsenal. These "intrinsic immunity" factors aid host cells in the race between viral replication and cellular production of new antiviral machines (Bieniasz, 2004; Sheehy et al., 2002). The virus also benefits from hijacking host mechanisms before transcription-dependent defenses alter the cellular environment.

Recent technological advances uncovered host RNA-binding proteins (RBPs) that interact with RNA viruses, yet no current method can identify interactions between the infecting, primary viral genome and host proteins (Lenarcic et al., 2013; Phillips et al., 2016). We developed a method to capture interactions between incoming viral RNA genomes and cellular proteins. VIR-CLASP (**VIR**al **C**ross-**L**inking **A**nd **S**olid-phase **P**urification) (B. Kim et al., 2020) differs from RNA-antisense purification mass-spectrometry (RAP-MS) (Phillips et al., 2016) and thiouracil crosslinking mass-spectrometry (TUX-MS) (Lenarcic et al., 2013) in two fundamental ways. First,

VIR-CLASP captures interactions with just the pre-replicated viral genome. Second, VIR-CLASP employs sequence-independent purification under protein-denaturing conditions. Here, we use VIR-CLASP to discover the initial host protein-viral RNA interactions between human cells and Chikungunya (CHIKV) or Influenza A (IAV). We selected these human pathogens to represent distinct types of RNA viruses with different strategies for viral replication. CHIKV is a member of the *Togaviridae* family of positive-sense RNA viruses (Weaver and Lecuit, 2015). CHIKV replication occurs in the cytoplasm, where the genome is a template for translation (Silva and Dermody, 2017). IAV is a member of the *Orthomyxoviridae* family of negative-sense RNA viruses. Upon infection, IAV translocates to the nucleus before initiating transcription of its segmented genome into the coding positive-sense RNA (Samji, 2009).

VIR-CLASP and mass spectrometry identified of hundreds of proteins that directly bind CHIKV and IAV. We found both shared and distinct host protein interactions with CHIKV and IAV. In-depth study of the CHIKV interactome uncovered condition-dependent interactions, including proteins that bind to CHIKV RNA following pretreatment of cells with interferon or under naïve conditions. We next explored the impact of three CHIKV interacting proteins on viral replication. VIR-CLASP revealed that the N^6-methyladenosine (m^6A)-binding proteins YTH-domain 1-3 (YTHDF1-3) interact with pre-replicated CHIKV RNA. We established that CHIKV genomes contain m^6A, an RNA modification found in eukaryotic and viral transcripts, which can regulate mRNA translation (X. Wang et al., 2015a), stability (X. Wang et al., 2014), and localization (Ries et al., 2019). Overexpression and knockdown studies showed that YTHDF1 suppresses CHIKV replication, while YTHDF2 has the opposite effect. VIR-CLASP uncovered extensive interactions between host proteins and the incoming genomes of CHIKV and IAV, and we found that these RBPs impact CHIKV replication and maturation. VIR-CLASP does not use

sequence-specific isolation of RNA; thus, we anticipate that this approach will facilitate study of many pathogenically relevant RNA viruses.

Results

Capturing interactions between host proteins and pre-replicated viral RNA

VIR-CLASP captures interactions specifically with pre-replicated viral RNA through infection of unlabeled host cells with 4-thiouridine (4SU)-labeled viral genomes (Figure 3.1A, Box 1). When infected host cells are irradiated with 365nm light, covalent crosslinks are formed between the 4SU-labeled viral RNAs and any interacting host or viral proteins (Figure 3.1A, Box 2). Solid-phase capture of these complexes with solid-phase reversible immbolization (SPRI) beads under protein denaturing conditions leads to purification of total RNA, and only proteins covalently crosslinked to the 4SU-labeled viral RNA (Figure 3.1A, Box 3). Lastly, RNAs are degraded using Benzonase nuclease, and the purified proteins are identified using LC-MS/MS (Figure 3.1A, Box 3). Only incoming viral genomes contain 4SU, so later events of viral replication or interactions between mRNA and protein are not captured. Since protein recovery is sequence-independent and achieved under high stringency, VIR-CLASP facilitates comparisons between time points, viral strains, and cell lines.

To show the versatility of VIR-CLASP we performed pilot experiments on viruses spanning seven families and representing plus- and minus-stranded RNA genomes: CHIKV (*Togaviridae*), EMCV (Encephalomyocarditis, *Picornaviridae*), MHV (Mouse hepatitis, *Coronaviridae*), ZIKV (Zika, *Flaviviridae*), RVFV (Rift Valley Fever, *Phenuiviridae*), IAV (subtype H3N2,

Orthomyxoviridae), and VSV (Vesicular stomatitis, *Rhabdoviridae*) (Figure 3.1B). By silver-stain, we observed diverse banding patterns across the viruses indicating interactions between cellular proteins and 4SU-labeled viral genomes. Negative controls represent cells infected with unlabeled virus, crosslinked, and processed using VIR-CLASP.

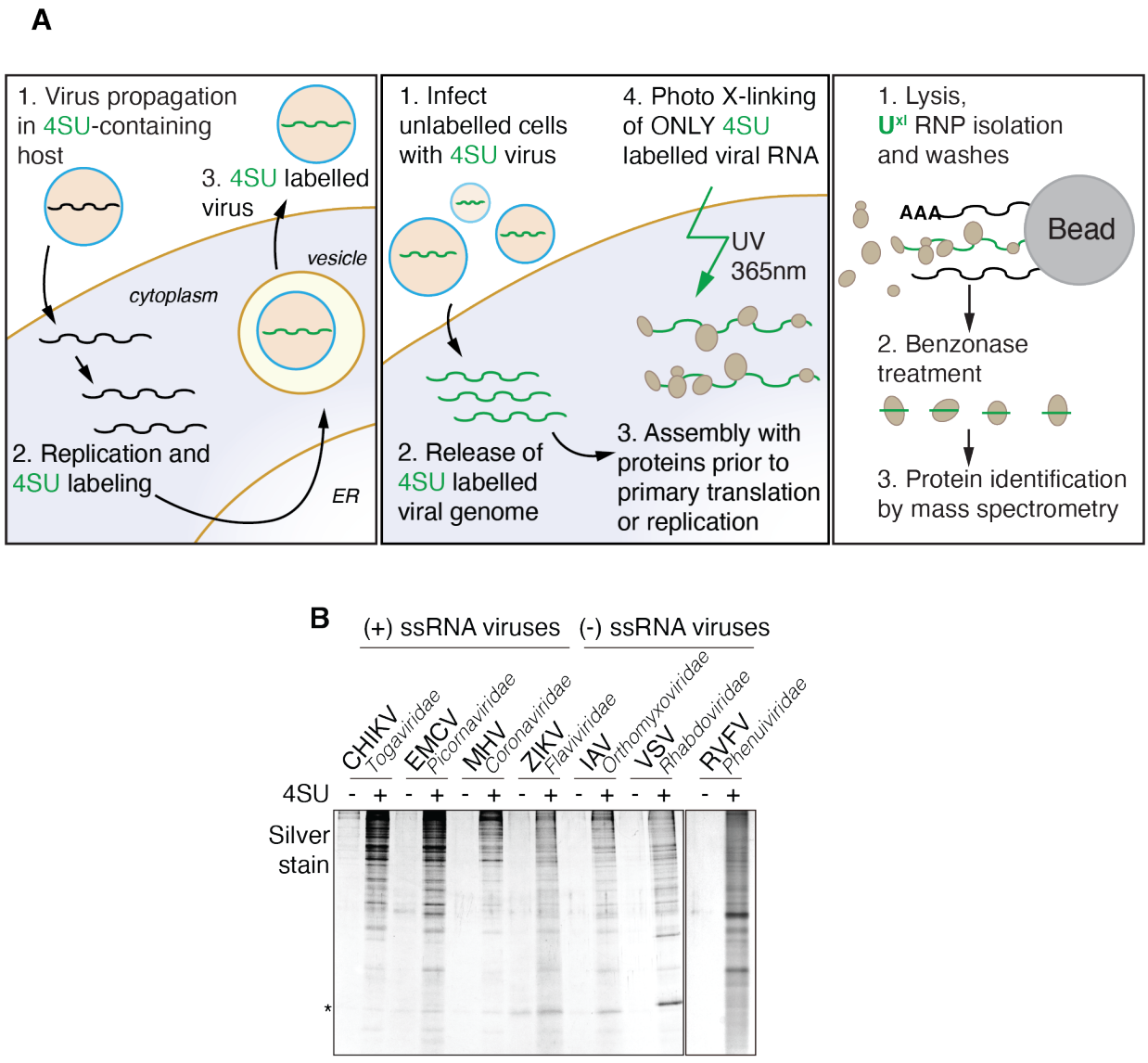


Figure 3.1 VIR-CLASP captures host proteins that directly interact with pre-replicated RNA virus genomes. A) Diagram showing VIR-CLASP technique. Viruses are labeled with 4-thiouridine (4SU), then infected into unlabeled host cells. RNA-protein complexes are crosslinked with 365nm light. Crosslinked complexes are purified, RNA is digested, and proteins are identified using LC-MS/MS. B) Silver stain of proteins purified from host cells using VIR-CLASP with various RNA viruses. Cells were infected with 4SU-labeled or unlabeled virus. Host cells for the RNA viruses were: U2OS for CHIKV; A549 for EMCV, IAV, VSV, RVFV; DBT-9 for MHV; Huh-7 for ZIKV. CHIKV, Chikungunya virus; EMCV, Encephalomyocarditis virus; IAV, Influenza A virus; VSV, Vesicular Stomatitis virus; RVFV, Rift Valley Fever virus; MHV, Mouse Hepatitis virus; ZIKV, Zika virus. *, Benzonase protein band. Panel (B) adapted from (B. Kim et al., 2020).

We chose to validate known interactions for CHIKV, IAV, and ZIKV using western blot. For CHIKV, we observed that ELAVL1 binds to the pre-replicated genome during the earliest timepoints of infection (Figure 3.2A). While ELAVL1 is known to interact with CHIKV RNA during infection (Dickson et al., 2012), this data further demonstrates that it can bind CHIKV RNA pre-replication. Importantly, the anti-viral RNA-binding protein IFIT1 does not interact with pre-replicated CHIKV, consistent with previous work showing that IFIT1 binds only to RNA viruses with a 5' cap-0 structure (Daffis et al., 2010). YTHDF1 is known to bind N6-methyladenosine (m^6A) modifications in IAV (Courtney et al., 2017) and ZIKV (Gokhale et al., 2016; Lichinchi et al., 2016b), and we expanded on these reports to show that it can bind to these viral genomes before replication (Figure 3.2B, C). ELAVL1 did not bind incoming IAV genomes (Figure 3.2B). For ZIKV, we observed distinct protein banding patterns after infection into either A549 or BEAS-2B cells; this demonstrates the utility of VIR-CLASP in identifying cell line-specific host factors that may influence viral tropism (Figure 3.2C). Lastly, TUBA4A did not bind pre-replicated CHIKV, IAV, or ZIKV (Figure 3.2A-C).

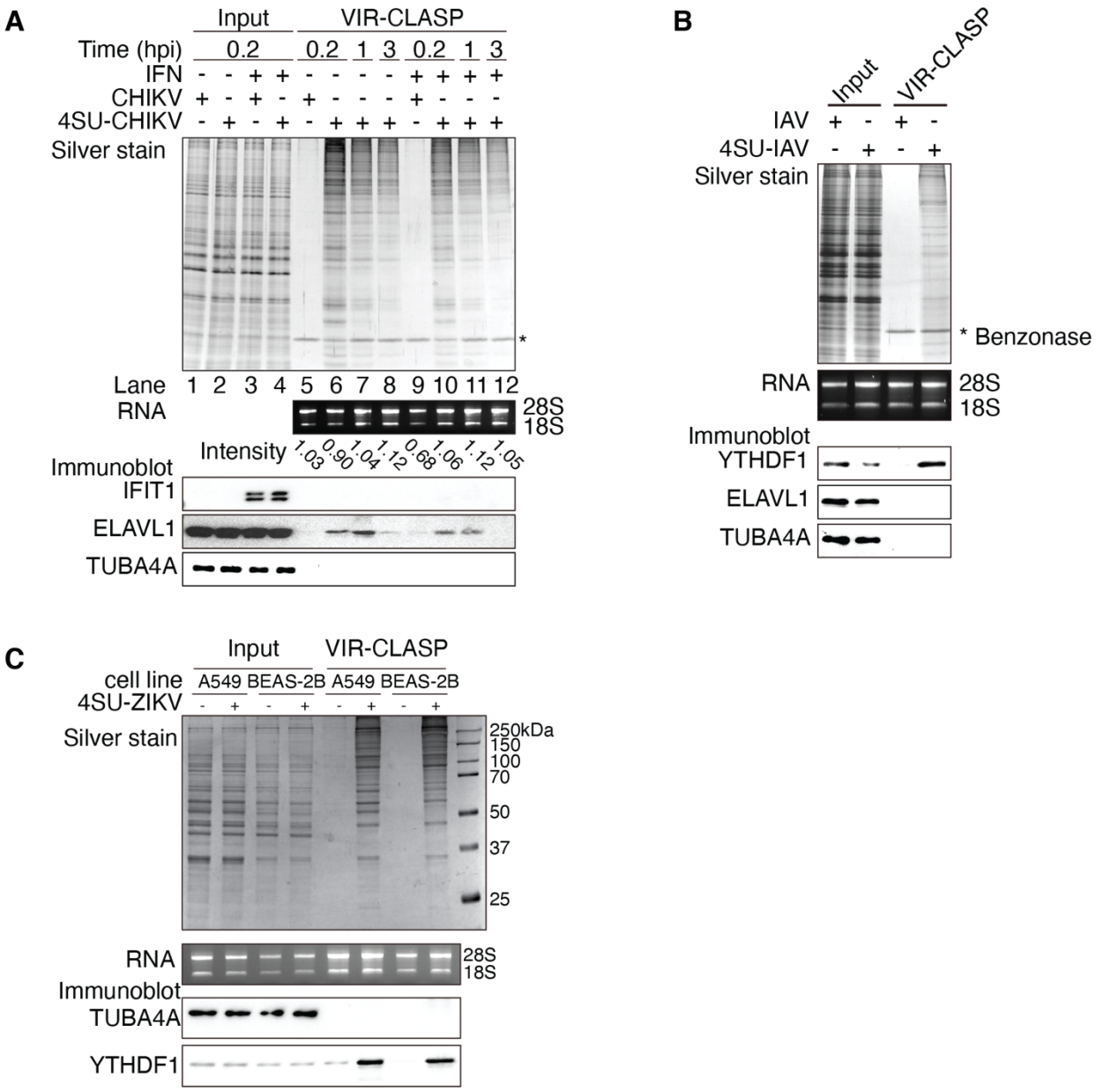
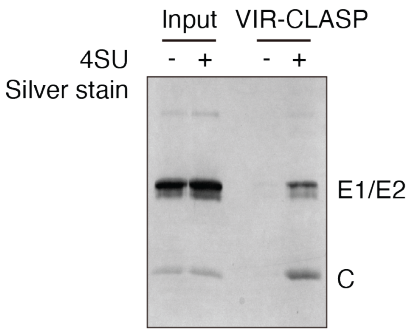


Figure 3.2 VIR-CLASP reveals host protein interactions with CHIKV, IAV, and ZIKV. A) Silver stain (top), agarose gel (middle), and immunoblot (bottom) showing the VIR-CLASP purification from U2OS cells infected with 4SU-labeled or unlabeled CHIKV. U2OS cells were stimulated with IFN (500U/mL recombinant interferon-β) or vehicle 16 h pre infection. rRNA band intensity quantified with ImageJ. Data represent three biologically independent replicates. B) Silver stain (top), agarose gel (middle), and immunoblot (bottom) showing the VIR-CLASP purification from A549 cells infected with 4SU-labeled or unlabeled IAV. C) Silver stain (top), agarose gel (middle), and immunoblot (bottom) showing the VIR-CLASP purification from A549 or BEAS-2B cells infected with 4SU-labeled or unlabeled ZIKV. Panels (A, B) adapted from (B. Kim et al., 2020).

To investigate whether host proteins or RNAs are packaged within the virion and are thus captured with VIR-CLASP, we performed VIR-CLASP using just purified 4SU-labeled CHIKV (Figure 3.3A), without infection into host cells. RNA packaged within the viral particle may lead to identification of interactions with non-viral RNAs. Using RNA-sequencing, we found that greater than 90% of reads over two replicates mapped to either CHIKV or the genome of the propagating BHK21 cell line, *Mesocricetus auratus* (Golden hamster) genomes (Figure 3.3B). Of these, 98.8% mapped to CHIKV, with the remaining ~1% mapping to *M. auratus*. Thus, BHK21 RNA represents a minority fraction of the RNA packaged within CHIKV virions. We conclude that the potential for identifying interactions with non-viral RNA packaged within virions using VIR-CLASP is likely minimal and stochastic.

A



B

	Total Mapped Reads 342,393,283	
	Reads	Percent
CHIKV aligned 1 time	337,800,476	98.7%
CHIKV aligned > 1 time	338,828	0.1%
Total alignment rate CHIKV	338,139,304	98.8%
M. auratus aligned 1 time	4,044,505	1.2%
M. auratus aligned > 1 time	209,474	0.1%
Total alignment rate *M. auratus*	4,253,979	1.2%

Figure 3.3 CHIKV virions contain few host proteins and RNAs. A). Silver stain showing VIR-CLASP purification from CHIKV virions. E1/E2, viral glycoproteins; C, viral capid protein. B) Table summarizing the RNA-sequencing reads mapped to CHIKV strain 181/25 and *Mesocricetus auratus* genomes. RNA was isolated from VIR-CLASP performed on purified CHIKV virions. Figure adapted from (B. Kim et al., 2020).

Proteomics analysis of the CHIKV and IAV pre-replicated interactomes

After validating known interactors and non-interactors, we performed LC-MS/MS analysis of VIR-CLASP with 4SU-labeled CHIKV at 0.2, 1, and 3 hpi, with naïve or IFN-treated cells (Figure 3.2A). A decrease in interacting protein recovery at 1 and 3 hpi suggests either that host proteins progressively bind less incoming viral RNA over time or that the incoming genomes are decaying. We also performed LC-MS/MS of VIR-CLASP with IAV to test the specificity of VIR-CLASP and to identify differences in host proteins that interact with pre-replicated plus- and minus-strand RNA viruses (Figure 3.2B).

To define candidate "VIR-CLASP RBPs" we calculated the peptide intensity ratios between +4SU and -4SU samples, leading to identification of ~400 significantly enriched proteins in each condition (0.01% FDR and fold change > 5) (Figure 3.4A). Due to the low complexity of the -4SU samples, we could not calculate peptide intensity ratios for all proteins. For the remaining proteins, we used a semiquantitative approach based on the assumption that peptides with no intensity value are below the detection threshold (Garcia-Moreno et al., 2019; Sysoev et al., 2016) (Figure 3.4B). We employed the same protocol to identify candidate "VIR-CLASP RBPs" for IAV, leading to identification of 316 proteins (Figure 3.5A, B).

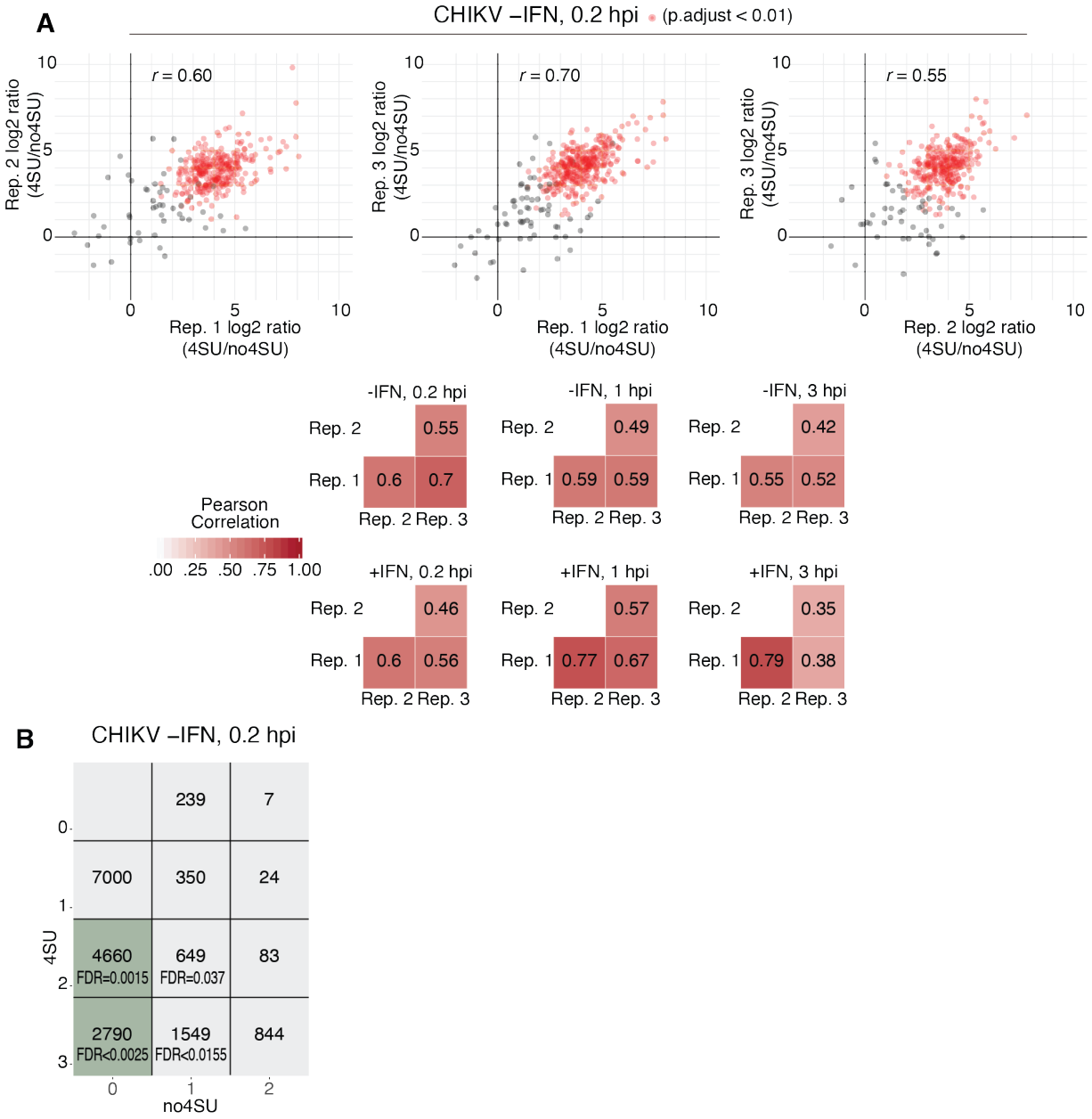


Figure 3.4 Proteomic analysis of the pre-replicated CHIKV interactome. A) Top: representative scatterplots of the log2-intensity ratios of quantified proteins in 4SU over no4SU samples for VIR-CLASP with CHIKV. Data represent three biological replicates (scatterplots show pairwise comparisons). Red points indicate significantly enriched proteins among all three replicates. Bottom: heatmaps summarizing the Pearson correlation coefficients for all 6 conditions of VIR-CLASP with CHIKV. P-values were calculated using a moderated t test, and the Benjamini-Hochberg correction for multiple testing was applied. B) Representative semi-quantitative analysis of proteins that did not have intensity values in all replicates of VIR-CLASP with CHIKV. Matrix shows the number of peptides found in 0, 1, or 2 replicates for no4SU samples, and 0, 1, 2, or 3 replicates for 4SU samples for the -IFN, 0.2 hpi condition. Green shading indicates cells with estimated FDRs of less than 0.01. Figure adapted from (B. Kim et al., 2020).

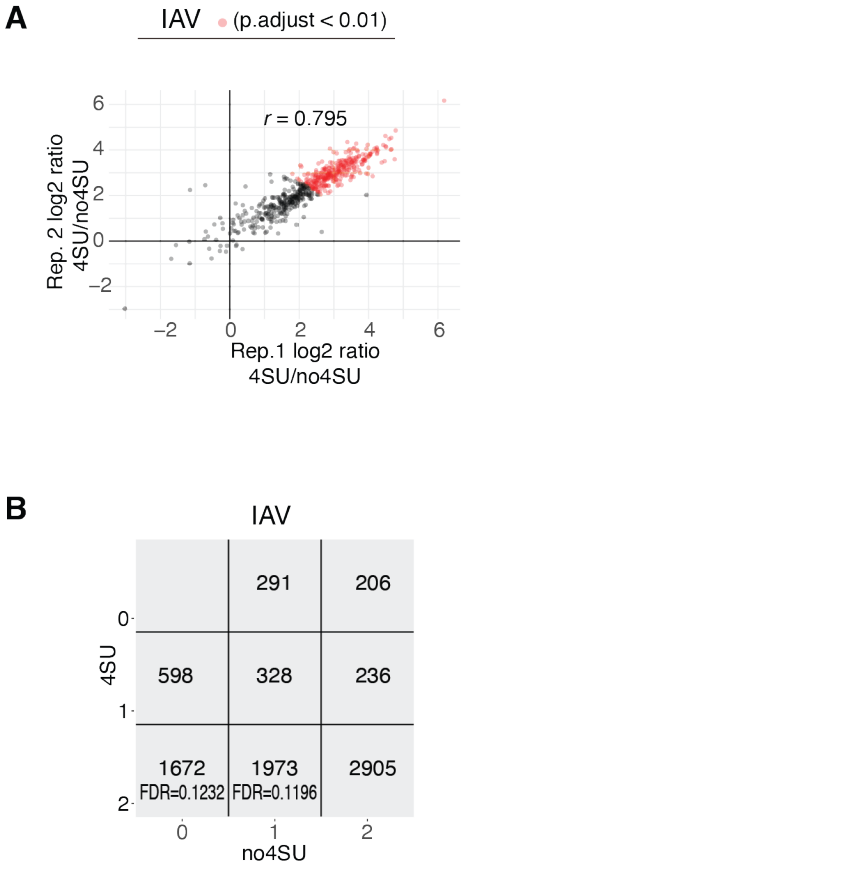


Figure 3.5 Proteomic analysis of the pre-replicated IAV interactome. A) Scatterplot as in Figure 1.4A depicting two biological replicates of VIR-CLASP with IAV. B) Semi-quantitative matrix as in Figure 1.4B for two biological replicates of VIR-CLASP with IAV. Figure adapted from (B. Kim et al., 2020).

To address the possibility that host proteins package within the virion we performed LC-MS/MS analysis of VIR-CLASP performed on the purified CHIKV virion (Figure 3.3A). We identified CHIKV Capsid, E1, and E2 proteins with very high intensities. We also compared all proteins with non-zero intensity values in one or both replicates of the virion LC-MS/MS to our CHIKV candidate RBPs. 40 proteins had identical peptides present in one or both replicates of the virion LC-MS/MS and the CHIKV candidate list in any condition. Many of the proteins found in both virion and VIR-CLASP are ribosomal proteins.

About 50% of proteins identified in VIR-CLASP for IAV were also identified for CHIKV in at least one condition (Figure 3.6A). The 193 shared proteins include known regulators of viral infection (STAU1 (de Lucas et al., 2010), ZC3HAV1 (Li et al., 2017), and LARP1 (Karlas et al., 2010; Suzuki et al., 2016), and RNA helicases (DHX9, DHX15, DHX29, DHX30, DHX36, and DHX38). We validated unique and overlapping IAV and CHIKV proteins using immunoblot, confirming that FXR1 and DDX21 are specific interactors of pre-replicated CHIKV, WDHD1 and KDMB3 are specific interactors of pre-replicated IAV, and LARP1 interacts with both (Figure 3.6B).

We identified viral proteins in the pre-replicated interactomes of both CHIKV and IAV. For CHIKV, we identified Capsid and E1 in all timepoints and conditions; we identified E2 only at 0.2 hpi and 1 hpi (in both +IFN and –IFN). For IAV we identified Nucleoprotein (NP), confirming its interaction with the IAV genome (Area et al., 2004; Das et al., 2010). For the remainder of this dissertation, we will focus on the CHIKV pre-replicated interactome due to the increased depth of the proteomics data.

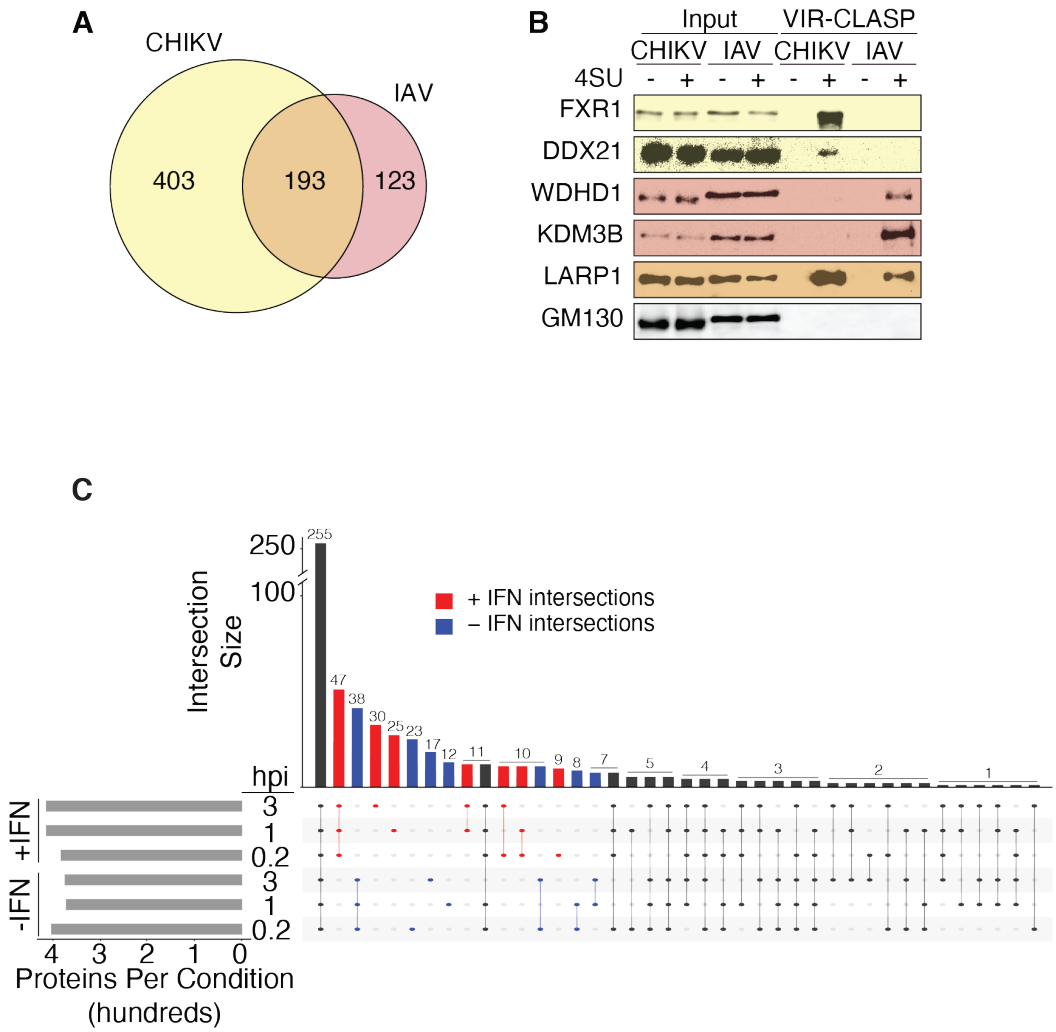


Figure 3.6 Comparisons between VIR-CLASP with different conditions and different viruses. A) Venn diagram showing the proteins identified by VIR-CLASP for CHIKV (composite of all 6 conditions tested) or IAV. B) Immunoblots validating shared and unique proteins in the CHIKV and IAV pre-replicated interactomes. C) UpSet diagram showing the shared and unique proteins among 6 conditions of VIR-CLASP with CHIKV. The top bar chart shows the number of protein identifications shared by the highlighted dots below. The side bar chart shows the total number of proteins identified for each condition. Red or blue coloring indicates proteins unique to either (-) or (+) IFN conditions. Figure adapted from (B. Kim et al., 2020).

The pre-replicated CHIKV interactome contains consistent and dynamic factors among the conditions tested. We found that ~255 of the candidate proteins interact with incoming CHIKV RNA throughout the first 3 hours of infection in both naïve and IFN-treated cells; ~340 proteins were present only in a subset of conditions (Figure 3.6C). We identified 142 proteins unique to the +IFN dataset, while 115 proteins were unique to –IFN (Figure 3.6C). The +IFN-unique proteins include known regulators of viral infection like SAMD9 (Liu et al., 2015) and PNPT1 (Dhir et al., 2018). 23 proteins were specific to the most "naïve" condition (0 hpi, -IFN), including EIF4G1 and RAC1.

Functional analysis of the CHIVK and IAV pre-replicated interactomes

To understand the functions of the proteins in the pre-replicated CHIKV interactome, we performed Gene Ontology (GO) analysis with candidate VIR-CLASP RBPs (Figures 3.7A-C). Many enriched GO molecular function, biological process, and cellular component terms were related to RNA binding and different aspects of RNA metabolism, including translation initiation (GO: 0006413) (Figures 3.7A-C). The CHIKV interactome was also enriched for terms relating to viral infection and immune response (Table 3.1). Enriched KEGG pathways in the CHIKV interactome include Spliceosome (hsa03040), Ribosome (hsa03010), and other RNA processing pathways (Figure 3.7D).

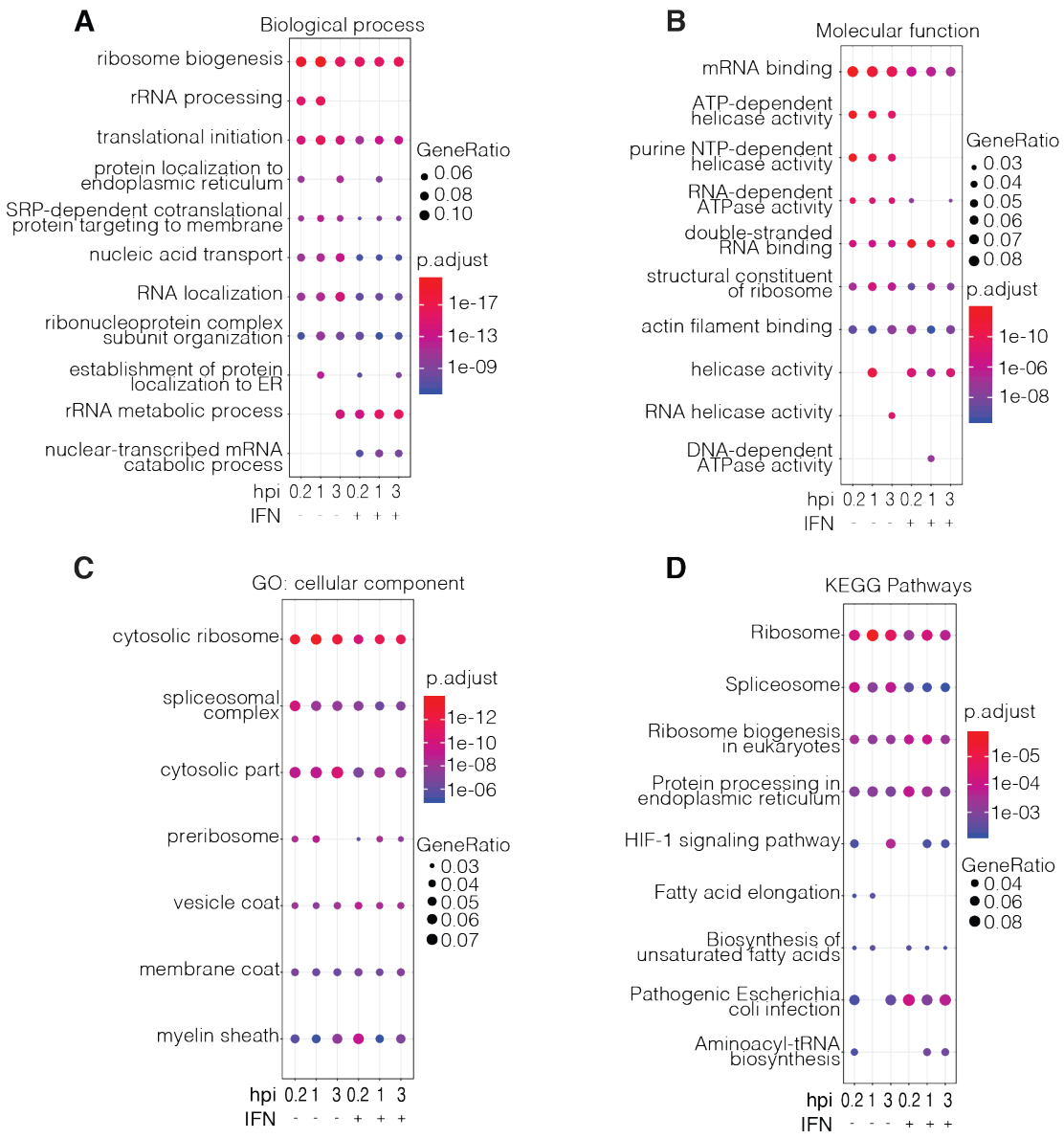


Figure 3.7 Functional analysis of the CHIKV pre-replicated interactome. A – D) GO analysis for enriched biological process (A), molecular function (B), or cellular component (C) terms, or KEGG pathways, in 6 conditions of VIR-CLASP for CHIKV. Dot plots indicate the Gene Ratio (size) and significance (color) the enriched terms. P-values were adjusted for multiple testing using the Benjamini-Hochberg correction. The background is all human genes. The most significant terms for each ontology are shown, and only significant results ($p < 0.01$) are plotted. In A – C, terms were collapsed by semantic similarity using the R/Bioconductor package clusterProfiler (Yu et al., 2012). Figure adapted from (B. Kim et al., 2020).

Over 50% of the CHIKV interactome proteins were interferon-stimulated genes (ISGs) (Rusinova et al., 2013) (Figure 3.8A). While the proportion of ISGs was consistent among the timepoints and conditions we tested, we were curious whether the ISGs represented potentially novel RBPs. About 60% of the CHIKV interactome is classified as an RBP by GO analysis or by previous interactome identification (Baltz et al., 2012; Castello et al., 2012; Gerstberger et al., 2014; R. Huang et al., 2018; Perez-Perri et al., 2018). The remaining 40% are potentially novel RBPs (Figure 3.8A). We observed that the proportion of novel RBPs that are also ISGs tended to increase over time or in response to IFN (Figure 3.8A). Comparison of the PFAM domains (El-Gebali et al., 2019) found in the proteins identified by VIR-CLASP for CHIKV revealed that the previously reported RBPs were enriched in helicase, RNA-recognition motif (RRM) and other nucleotide-binding domains, while the novel RBPs contained Gelsolin repeat, WD40 repeat, and FHA domains (Figure 3.8B).

GO analysis of the IAV pre-replicated interactome revealed enrichment for molecular function terms related to RNA binding, for biological process terms related to RNA localization and viral genome replication, and for KEGG pathways related to RNA processing (Figures 3.9A-D and Table 3.2). Over 50% of the proteins in the IAV interactome were previously classified as RBPs (Figure 3.9E).

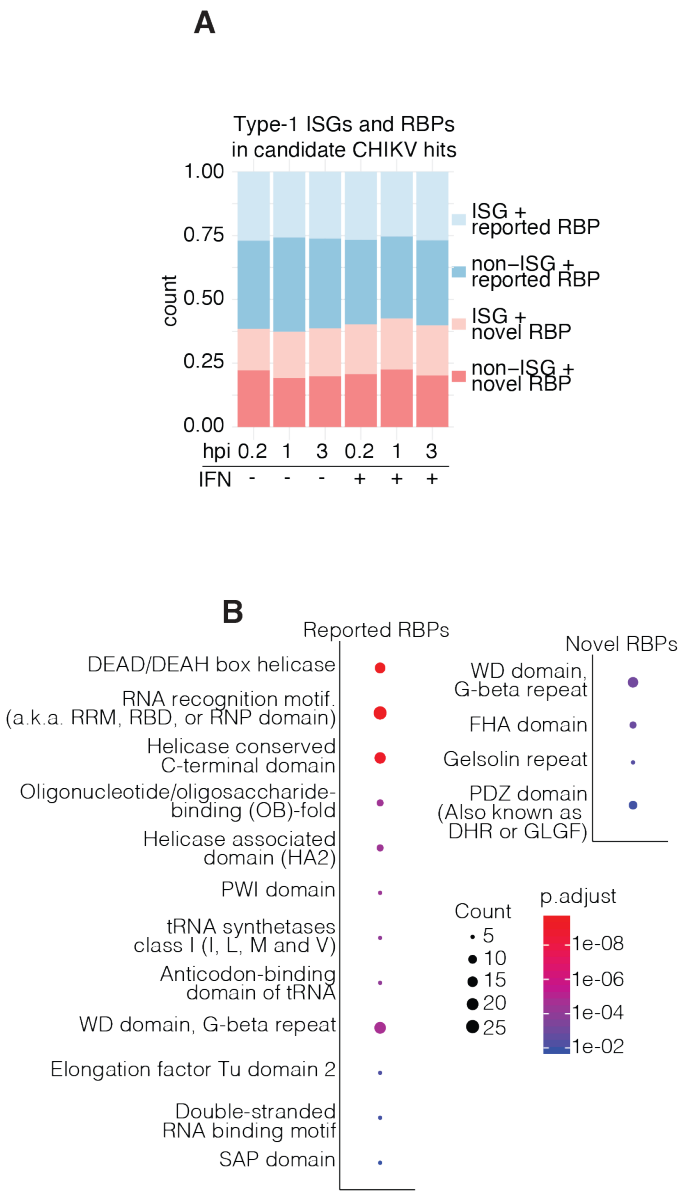


Figure 3.8 The CHIKV pre-replicated interactome contains annotated and novel RBPs. A) Bar chart showing previously annotated RNA-binding proteins (RBPs) and interferon-stimulated genes (ISGs) in the CHIKV pre-replicated interactome. RBP annotations are a compilation of GO and previous interactome-capture efforts. ISG annotation is based on the interferome (Rusinova 2013). B) Enriched PFAM domains by previously reported or novel RBPs in the CHIKV pre-replicated interactome. Dot plots show the number of proteins with the domain (size), and the statistical significance (color). P-values were calculated using Fisher's exact test with the Bonferroni correction. The background is all human genes. Figure adapted from (B. Kim et al., 2020).

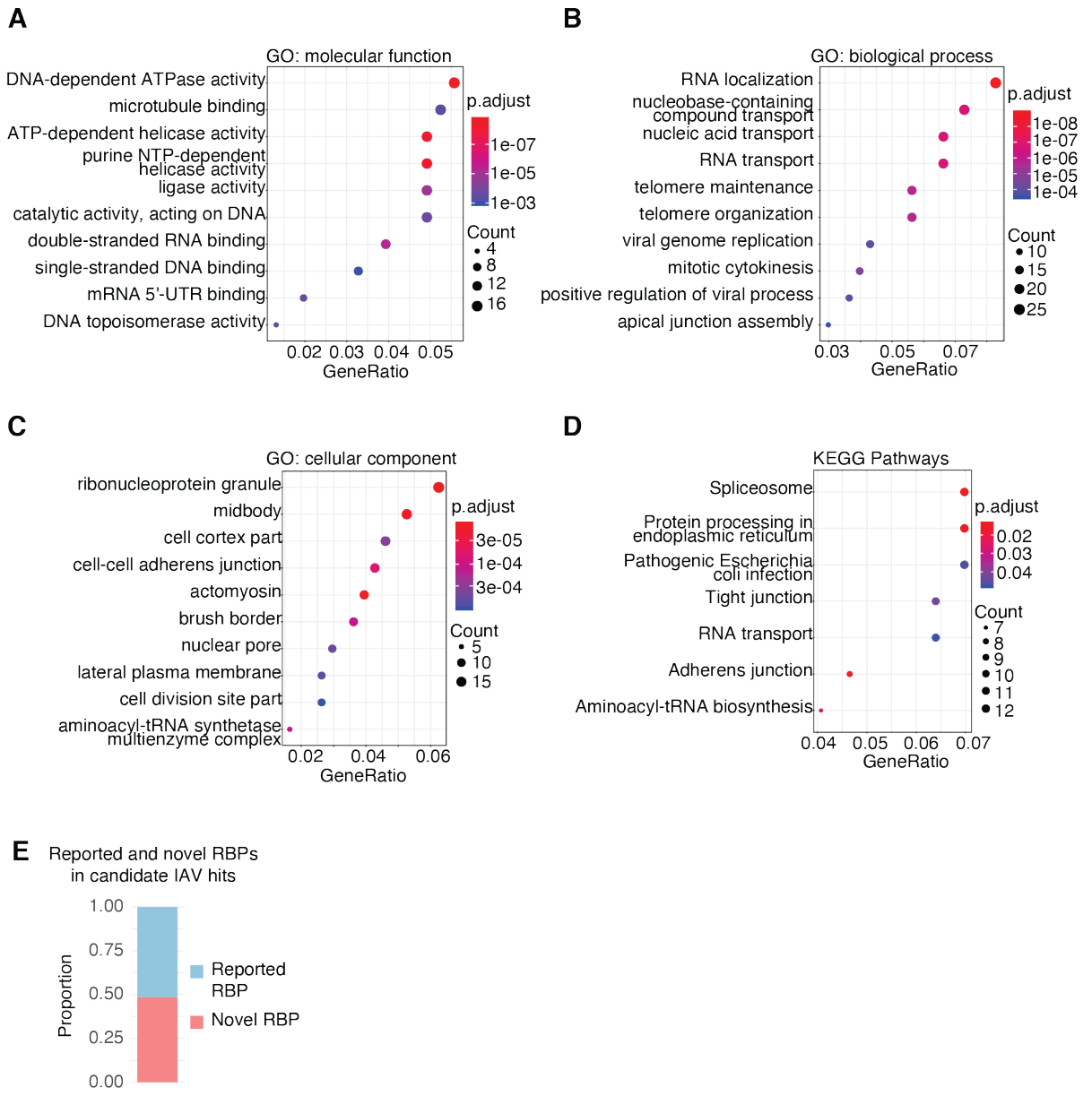


Figure 3.9 Functional analysis of the IAV pre-replicated interactome. A – D) Gene ontology analysis for enriched biological process (A), molecular function (B), or cellular component (C) terms, or KEGG pathways, in VIR-CLASP for IAV. Dot plots indicate the Gene Ratio (size) and significance (color) the enriched terms. P-values were adjusted for multiple testing using the Benjamini-Hochberg correction. The background is all human genes. The most significant terms for each ontology are shown, and only significant results ($p < 0.01$) are plotted. In A – C, terms were collapsed by semantic similarity using the R/Bioconductor package clusterProfiler (Yu et al., 2012). E) Bar chart showing the previously annotated RNA-binding proteins (RBPs) in the IAV pre-replicated interactome. RBP annotations are a compilation of GO and previous interactome-capture efforts. Figure adapted from (B. Kim et al., 2020).

We further explored the CHIKV interactome proteins annotated by the GO term "translation initiation" (GO: 0006413) (Figure 3.10A and Table 3.1). These included translation initiation factors, RNA helicases, and the m^6A -binding proteins YTHDF2 and YTHDF3. While m^6A exists in other plus-strand RNA viruses like flaviviruses(Gokhale et al., 2016), whether m^6A exists in alphavirus genomes is unknown. The identification by VIR-CLASP of a direct interaction of YTHDF2 and YTHDF3 with pre-replicated CHIKV (Figure 3.10B) suggests that CHIKV genomes contain m^6A.

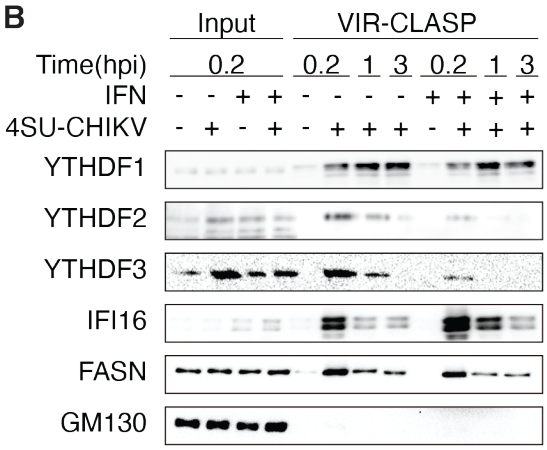
CHIKV VIR-CLASP hits include nucleic acid pattern-recognition receptors (PRRs) (Chow et al., 2018; Ma et al., 2018), such as DHX9, DDX21, OAS3, IFI16, XRCC6, and PRKDC (Figure 3.10C). Although CHIKV has an RNA genome, we found that pre-replicated CHIKV interacts with IFI16 and the DNA-PK complex (XRCC5, XRCC6, and PRKDC). RNA binding by the DNA-PK complex has been previously reported (Yoo and Dynan, 1998). IFI16 is an antiviral cytosolic-DNA sensor, yet was identified in all timepoints and conditions (Figure 3.10B). Recent research demonstrated that IFI16 restricts RNA virus infection by regulating transcription of type-I IFNs (Thompson et al., 2014). Identification of IFI16 as a VIR-CLASP hit and putative RBP suggests a novel mechanism for its restriction of RNA viruses.

A

GO biological process: translation initiation

ABCE1	EIF2A	TPR
DDX3X	EIF3A	**YTHDF2**
DHX29	EIF4G1	**YTHDF3**
DDX1	EIF5B	RPS6-8, 16, 23, 24
LARP1	NPM1	RPL3-7, 13A, 13-15, 18A, 26, 27A, 36A

B



C

Nucleic acid sensors found in CHIKV VIR-CLASP

	hpi	0.2	1	3	0.2	1	3	
	IFN	-	-	-	+	+	+	Ligand or pathogen interaction
RNA	DDX1	●	●	●	●	●	●	poly(I:C), IAV, and MRV
	DDX3	●	●	●	●		●	poly(I:C), VSV, SeV, HCV, WNV, JEV and HIV
	DDX21	●	●	●	●	●	●	poly(I:C), IAV, and MRV
	DHX9	●	●	●	●	●	●	poly(I:C), IAV, and MRV
	DHX15	●	●	●	●	●	●	poly(I:C), SeV, and MRV
	DHX36	●	●	●	●	●	●	poly(I:C), IAV, and MRV
	SNRNP200	●	●	●	●	●	●	poly(I:C) and SeV
	OAS3	●			●	●	●	poly(I:C), WNV, SINV, IAV, VACV and CHIKV
	OASL				●	●	●	poly(I:C), VSV, RSV, DENV and SeV
	PKR				●	●	●	poly(I:C), VACV, VSV and IAV
DNA	**IFI16**	●	●	●	●	●	●	Cytosolic DNA, HSV-1 and HCMV
	XRCC5	●	●	●	●	●	●	Cytosolic DNA and VACV
	XRCC6	●			●	●	●	Cytosolic DNA, HTLV-1 and VACV
	PRKDC	●	●	●	●	●	●	Cytosolic DNA and VACV

D

Previous screens	Virus	Protein
Karlas et al., 2016	CHIKV	ACLY, AP2M1, BOP1, CLTC, CNOT1 COPG1, DHX38, EEF2, **FASN**, GCN1 IQGAP1, KDM1A, PDLIM5, PPIL2, PTBP1 RACK1, RRBP1, SIN3A, STT3B, TECR
Ooi et al., 2013	SINV	ARCN1, LIMA1, OGDH, STK10, VARS, YARS
Radoshitzky et al., 2016	VEEV	CLTC, COPA, RAC1, STAU1, VCP
Schoggins et al., 2011	CHIKV	ADAR, GALNT2, OAS3

Figure 3.10 VIR-CLASP with CHIKV identifies proteins with known anti-viral roles. A) Table of CHIKV pre-replicated interactome proteins identified within the GO term “translation initiation” (GO: 0006413). B) Immunoblots validating the proteins found in VIR-CLASP with CHIKV. Data represent two independent biological replicates in (B). C) Table of nucleic acid sensors in the CHIKV pre-replicated interactome. PRR annotations were derived from recent review articles (Chow 2018, Ma 2018). D) Table of proteins in the CHIKV pre-replicated interactome that were identified in previous screens for antiviral factors. Figure adapted from (B. Kim et al., 2020).

We compared our CHIKV VIR-CLASP dataset to previous screens for regulators of CHIKV infection (Karlas et al., 2016; Ooi et al., 2013; Radoshitzky et al., 2016; Schoggins et al., 2011) (Figure 3.10D and Table 3.3). VIR-CLASP hits identified in at least one other screen include ADAR, ACLY, and FASN. FASN was also identified as a putative RBP in two other screens to identify novel cellular RBPs (Baltz et al., 2012; Castello et al., 2012). VIR-CLASP expands on these results to demonstrate that FASN binds viral RNA (Figure 3.10B). The connection between FASN's role in viral replication and its potential RNA-binding activity is unexplored.

For this report, we follow up on YTHDF1-3 to determine how their interactions with CHIKV affect viral replication.

YTHDF proteins have distinct regulatory roles in CHIKV replication and infection

We explored the function of YTHDF proteins on CHIKV through over- and underexpression studies. While YTHDF1 was not a CHIKV "candidate RBP", we decided to study its effect on CHIKV replication given that we identified the other YTHDF proteins by mass-spec, and all were identified by immunoblot (see Figure 3.10B). Strand-specific RT-qPCR distinguishes plus- and minus-strand CHIKV RNA (Figure 3.11A). Overexpression of YTHDF1 decreased both strands of CHIKV RNA (Figure 3.11B); knockdown of YTHDF1 increased both strands at 3 and 5 hpi (data normalized within each timepoint) (Figures 3.12A, B). The increase was highest in the minus-strand. By contrast, YTHDF2 overexpression increased both strands at 5 hpi, while overexpression of YTHDF3 resulted in little change (Figure 3.11B).

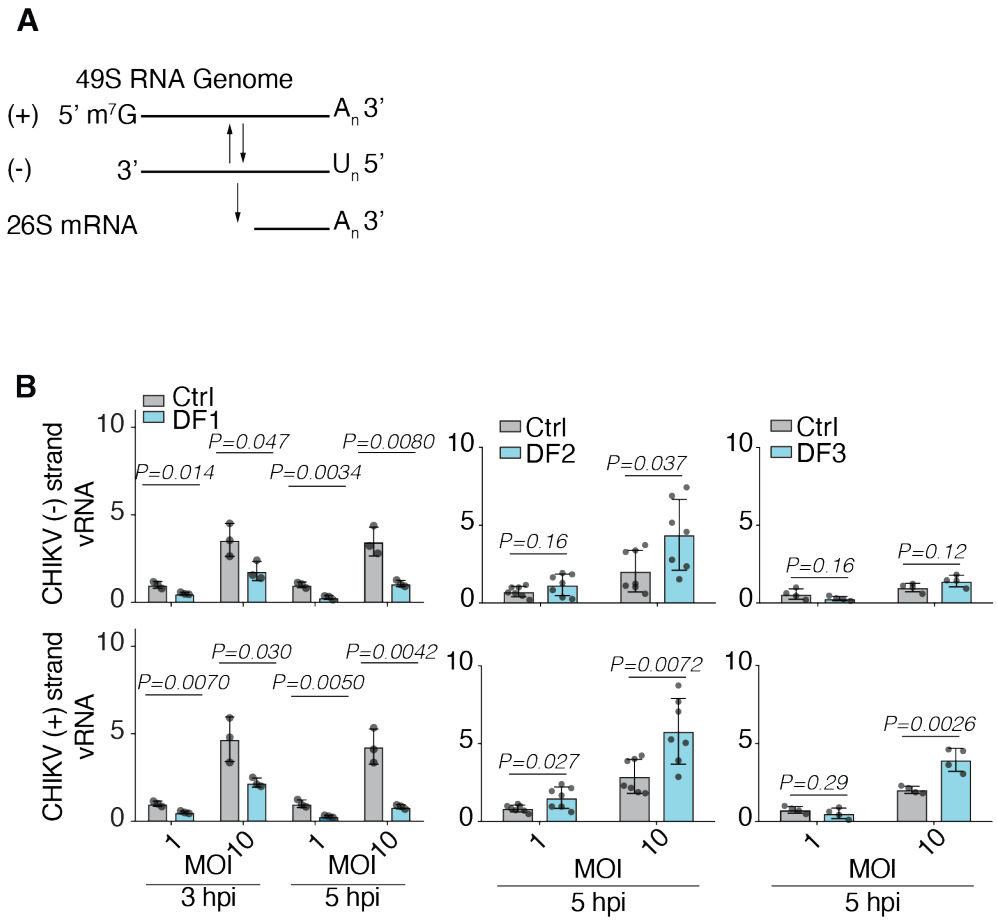


Figure 3.11 Overexpression of YTHDF proteins regulates CHIKV replication. A) Schematic showing the RNA species generated during CHIKV replication. B) Strand-specific qRT-PCR of CHIKV RNA levels in U2OS cells stably overexpressing Flag-HA tagged YTHDF1, YTHDF2, YTHDF3, or parental U2OS (Ctrl). Data were normalized to MOI 1 of parental U2OS, within each time point. n = 4. Figure adapted from (B. Kim et al., 2020).

Knockdown of YTHDF2 had little effect on CHIKV RNA levels at 5 hpi, and knockdown of YTHDF3 slightly increased CHIKV RNA (Figures 3.12A, B).

To examine possible indirect effects of YTHDF1 knockdown we rescued YTHDF1 levels by re-expression. Re-expression of YTHDF1 during siRNA knockdown of YTHDF1 rescued plus-strand CHIKV to the level of control, while the effect on minus-strand CHIKV was diminished, but not entirely to control levels (Figure 3.12C).

We next tested whether the early effects of YTHDF proteins on CHIKV persist to the release of new virions. Knockdown of YTHDF1 and YTHDF3 increased both extracellular viral RNA levels and mature virions; knockdown of YTHDF2 had the opposite effect (Figure 3.13A). These data show that YTHDF1 and YTHDF3 restrict CHIKV replication, while YTHDF2 promotes CHIKV replication.

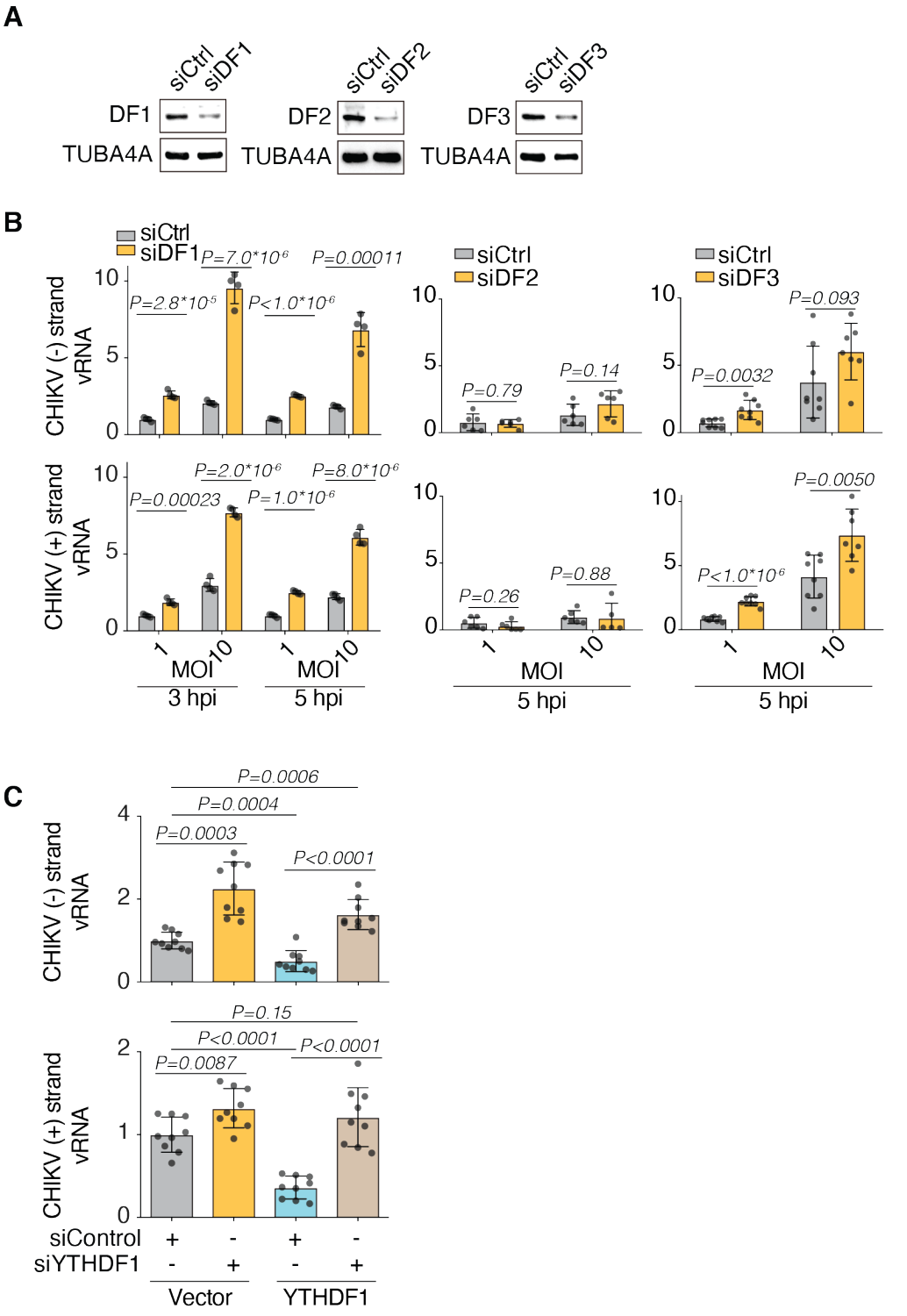


Figure 3.12 Knockdown of YTHDF proteins regulates CHIKV replication. A) Immunoblots of YTHDF protein expression in U2OS cells treated with siRNA targeting GFP (Ctrl), YTHDF1, YTHDF2, or YTHDF3. B) Strand-specific qRT-PCR of CHIKV RNA levels in U2OS cells treated with siRNAs. Cells were infected with CHIKV 72 h after siRNA transfection. n = 4. C) Rescue of CHIKV RNA levels by overexpression of YTHDF1. n = 9. Figure adapted from (B. Kim et al., 2020).

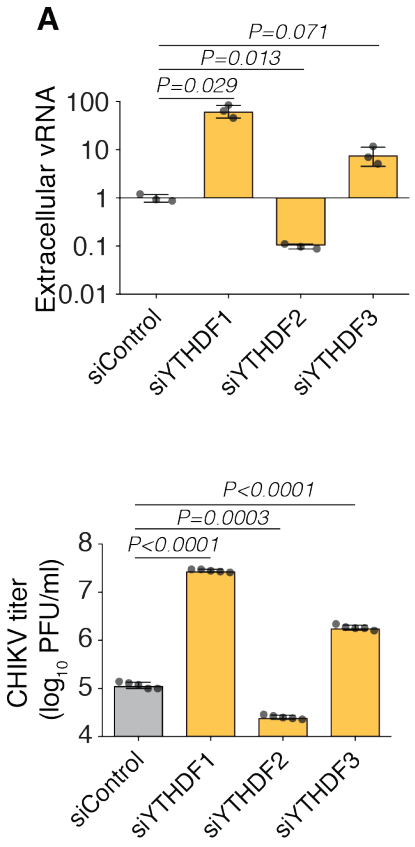


Figure 3.13 YTHDF proteins regulate production of infectious CHIKV particles. A) Top: qRT-PCT of extracellular CHIKV RNA from U2OS cells 24 hpi. MOI = 10-4. Cells were transfected with siRNAs 72 h before infection with CHIKV. Bottom: CHIKV titer measured using a plaque assay with Vero cells. n = 3. In A – E: error bars are mean ± SD; p-values calculated using two-tailed Student's t test. Figure adapted from (B. Kim et al., 2020).

The YTHDF proteins bind to m^{6}A, an RNA modification found in mRNA and lncRNA (Jia et al., 2013) and within viral genomes from the *Retro-*, *Orthomyxo-*, and *Flaviviridae* families (Courtney et al., 2017; Gokhale et al., 2016; Kennedy et al., 2017; Krug et al., 1976; Lichinchi et al., 2016a; Tirumuru et al., 2016). To explore whether CHIKV contains m^{6}A, we performed m^{6}A RNA-immunoprecipitation (MeRIP) on RNA from purified CHIKV virions. CHIKV virion RNA immunoprecipitated with an anti-m^{6}A antibody, but not with an IgG control (Figure 3.14A). We then performed Me-RIP-RT-qPCR with primers tiled along the CHIKV genome every ~1000 nt. We observed a significant enrichment within 2000 nt at the 5' end of the genome (Figure 3.14B). We then examined whether the m^{6}A sites on the CHIKV genome are bound by YTHDF1 using RIP-qRT-PCR. First, we confirmed that YTHDF1 is not present in purified CHIKV virions, despite being expressed in both the propagating and host cell lines (Figure 3.14C). RIP-qRT-PCR revealed that YTHDF1 is highly enriched along the entire CHIKV genome (Figure 3.15D). This is inconsistent with previous *in vitro* data demonstrating that the YTH domain (Figure 1.5) has ~10x higher affinity for m^{6}A-modified RNA versus unmodified RNA. However, recent reports have shown that the YTHDF proteins can bind unmodified viral RNA (Gokhale et al., 2016) and host mRNA (Yuan Zhang et al., 2019) *in vivo*. The binding determinants of YTHDF proteins on unmodified RNAs *in vivo* will be explored in Chapter IV.

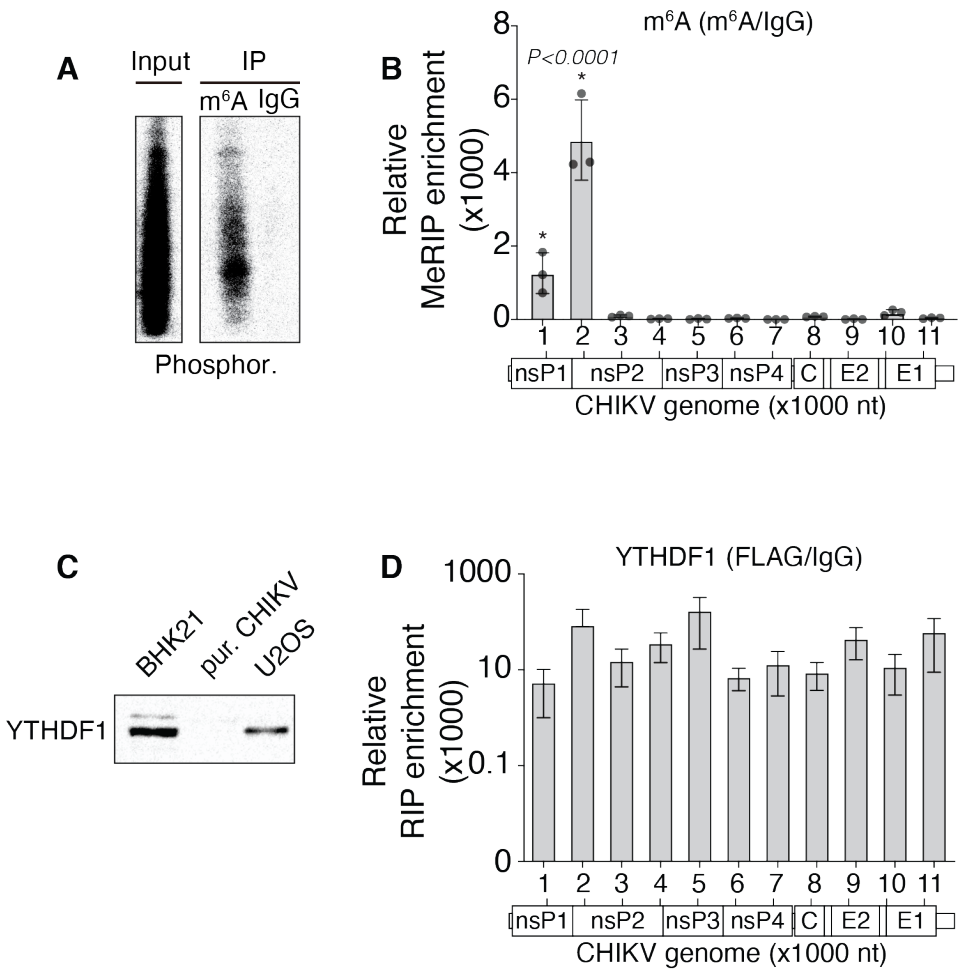


Figure 3.14 The CHIKV genome contains N^6-methyladenosine modifications. A) Phosphorimage of 32P-labeled RNA purified using meRIP with CHIKV virion RNA. CHIKV virions were purified twice by 30% - 60% discontinuous sucrose gradient, followed by concentration with a 30% sucrose cushion. n = 2. B) MeRIP-qRT-PCR of CHIKV virion RNA purified as in (A). Primers were tiled every 1000nt over the CHIKV genome. Enrichment shows m^{6}A IP over IgG IP. n = 3, error bars show mean ± SD. P-values were calculated by one-way ANOVA. C) Immunoblot for YTHDF1 in lysate from BHK21 cells, CHIKV virions purified as in (A), or U2OS cells. D). RIP-qRT-PCR of CHIKV RNA from U2OS cells stably expressing Flag-HA-YTHDF1. U2OS cells were infected with 4SU-labeled CHIKV, and cells were crosslinked with 365nm light prior to lysis. Primers were the same as in (A). Enrichment shows Flag IP over IgG IP. n = 2, error bars show mean ± SD. Panels (A, B) adapted from (B. Kim et al., 2020)

Conclusions

While PRRs and traditional antiviral proteins are important in innate immunity, proteins not considered traditional pro- or antiviral factors also contribute to the outcome of viral infection. "Intrinsic immunity" (Sheehy et al., 2002) factors aid normal cellular function but are co-opted for other processes during infection. An RNA virus can be a substrate for host RBPs, thus providing a mechanism for hijacking host RNA metabolism. RBPs can also aid cellular defense, as they can bind viral RNA before upregulation of traditional antiviral proteins. The earliest timepoints of infection are likely when these roles of RBPs are critical. Thus, VIR-CLASP aims to identify intrinsic and innate immune RBPs that drive initial cellular responses to infection and viral replication.

FASN represents an ideal target for co-option by viruses because it is required for cellular survival. A previous siRNA screen to identify regulators of CHIKV replication examined FASN enzymatic activity (Karlas et al., 2016). VIR-CLASP expands on these observations to show that FASN interacts with CHIKV RNA; whether interaction inhibits or promotes FASN enzymatic function is unknown. We favor a model in which CHIKV RNA hijacks FASN to localize its enzymatic activity to sites of viral replication, where elevated concentrations of palmitic acid would benefit viral packaging (Heaton et al., 2010) and replication (Na Zhang et al., 2019). It is unknown if this RNA-binding localization scheme contributes to cellular FASN function, or if other post-translational modification enzymes use RNA-binding to localize to sites of protein synthesis.

m^{6}A RNA modification is necessary for cellular survival (Geula et al., 2015). Whether cells use m^{6}A to mark viruses, or whether RNA viruses adopt m^{6}A to promote their replication appears virus- and cell-type specific (Courtney et al., 2017; Gokhale et al., 2016; Kennedy et al., 2017; Lichinchi et al., 2016b; 2016a; Tirumuru et al., 2016). Recent screens identified more m^{6}A-binding proteins, and proteins repelled by m^{6}A (Edupuganti et al., 2017; H. Huang et al., 2018). Thus, the global role of m^{6}A involves combinatorial or competing activities of a growing list of m^{6}A-sensitive RBPs. The reports identifying novel m^{6}A-RBPs investigated steady-state conditions. As many ISGs are also RBPs, there are likely unidentified m^{6}A-sensitive RBPs relevant to innate immunity.

YTHDF1 has a distinct effect on early and late CHIKV replication compared to YTHDF2-3, even though they bind the same RNA modification. YTHDF1 has a more persistent interaction with CHIKV than YTHDF2-3 by immunoblot, which may explain its antiviral potency (see Figure 3.10B). These observations are consistent with the cellular roles of YTHDF1-3, which also differ: YTHDF1 promotes translation of mRNAs (X. Wang et al., 2015a), while YTHDF2 destabilizes mRNA (X. Wang et al., 2014). Whether YTHDF1-3 have similar functions on viral RNAs remains unclear. m^{6}A modifications are present on mRNA and viral RNA, so our observations with CHIKV may be an indirect consequence of YTHDF1-3 acting on mRNA.

Previous efforts to identify novel RBPs used steady-state conditions (Baltz et al., 2012; Castello et al., 2012). VIR-CLASP uses biologically relevant conditions to identify RBPs that may be inactive unless stimulated. One example is that VIR-CLASP identified a novel RNA-binding function of an innate immune PRR. While IFI16 is known to recognize pathogen-associated DNA, our results demonstrate that it can also bind viral RNA. It is unclear how IFI16 distinguishes

between viral and host nucleic acids and whether its functions are distinct when binding foreign DNA versus foreign RNA.

The primary CHIKV interactome contains ~600 distinct proteins across 6 conditions, and the primary IAV interactome at 1 hpi contains ~300 proteins. They include known pro- and antiviral proteins and host proteins involved in cellular homeostasis. Though the high viral inoculum (Kummer et al., 2014) used in VIR-CLASP may raise false-positives, all validation experiments used MOI 10 or less. We cannot rule out the possibility that for some viruses, cellular RNA can be packaged stochastically with viral RNA. However, for IAV (Noda et al., 2018) and CHIKV (this report), little cellular RNA is packaged (IAV: $< 3\%$ (Noda et al., 2018), CHIKV: $< 1.3\%$ (this report)) (Figure 3.3B). For CHIKV, our data also indicate that few cellular RBPs exist in the viral particle (< 140 proteins with an intensity value > 0 in one or both biological replicates); only 40 RBPs in one or both replicates of the virion mass-spec had identical peptides found by VIR-CLASP post-infection (Figure 3.3A). For viruses that are known to package significant amounts of cellular RNA, use of VIR-CLASP may need additional optimization.

Few treatments or vaccines exist for emerging RNA viruses, in part due to their high mutation rates (Steinhauer and Holland, 1987). The identification of RBPs from VIR-CLASP can provide insights into RNA biology, improve our understanding of post-transcriptional gene regulation in host and pathogen, and lead to identification of new targets for therapeutic intervention.

Table 3.1 Enriched GO terms and KEGG pathways for VIR-CLASP with CHIKV. GO terms were collapsed by semantic similarity. Terms are grouped by VIR-CLASP condition. For brevity, only top 20 terms by statistical significance (adjusted p-value) are displayed. Full list of enriched GO terms can be accessed in (B. Kim et al., 2020).

Molecular Function			
Condition	**GO term**	**Gene ratio**	**p.adjust**
0 hpi, -IFN	mRNA binding	33/388	1.49E-12
0 hpi, -IFN	ATP-dependent helicase activity	20/388	1.49E-12
0 hpi, -IFN	purine NTP-dependent helicase activity	20/388	1.49E-12
0 hpi, -IFN	RNA-dependent ATPase activity	14/388	1.86E-10
0 hpi, -IFN	double-stranded RNA binding	17/388	2.12E-09
0 hpi, -IFN	structural constituent of ribosome	19/388	5E-08
0 hpi, -IFN	actin filament binding	21/388	5.56E-06
0 hpi, -IFN	ligase activity	17/388	8.35E-06
0 hpi, -IFN	ribonucleoprotein complex binding	15/388	3.06E-05
0 hpi, -IFN	telomeric DNA binding	8/388	0.000125
0 hpi, -IFN	ADP binding	8/388	0.000235
0 hpi, -IFN	aminoacyl-tRNA editing activity	4/388	0.002785
0 hpi, -IFN	structural constituent of nuclear pore	5/388	0.00343
0 hpi, -IFN	importin-alpha family protein binding	3/388	0.005429
0 hpi, -IFN	S100 protein binding	4/388	0.009278
1 hpi, -IFN	mRNA binding	31/357	1.26E-11
1 hpi, -IFN	helicase activity	24/357	3.79E-11
1 hpi, -IFN	ATP-dependent helicase activity	18/357	3.79E-11
1 hpi, -IFN	purine NTP-dependent helicase activity	18/357	3.79E-11
1 hpi, -IFN	RNA-dependent ATPase activity	13/357	1.05E-09
1 hpi, -IFN	structural constituent of ribosome	20/357	1.72E-09
1 hpi, -IFN	double-stranded RNA binding	16/357	5.12E-09
1 hpi, -IFN	ribonucleoprotein complex binding	16/357	2.83E-06
1 hpi, -IFN	ligase activity	16/357	1.35E-05
1 hpi, -IFN	actin filament binding	19/357	2.12E-05
1 hpi, -IFN	dolichyl-diphosphooligosaccharide-protein glycotransferase activity	4/357	0.000506
1 hpi, -IFN	ADP binding	7/357	0.000902
1 hpi, -IFN	structural constituent of nuclear pore	5/357	0.002146
1 hpi, -IFN	importin-alpha family protein binding	3/357	0.00381
1 hpi, -IFN	telomeric DNA binding	6/357	0.003916

3 hpi, -IFN	mRNA binding	30/360	1.02E-10
3 hpi, -IFN	RNA helicase activity	14/360	5.35E-10
3 hpi, -IFN	ATP-dependent helicase activity	17/360	5.35E-10
3 hpi, -IFN	purine NTP-dependent helicase activity	17/360	5.35E-10
3 hpi, -IFN	RNA-dependent ATPase activity	13/360	1.21E-09
3 hpi, -IFN	double-stranded RNA binding	16/360	6.72E-09
3 hpi, -IFN	structural constituent of ribosome	19/360	1.4E-08
3 hpi, -IFN	actin filament binding	22/360	3.66E-07
3 hpi, -IFN	ADP binding	10/360	2.03E-06
3 hpi, -IFN	ribonucleoprotein complex binding	15/360	1.31E-05
3 hpi, -IFN	ligase activity	16/360	1.31E-05
3 hpi, -IFN	importin-alpha family protein binding	4/360	0.000128
3 hpi, -IFN	disordered domain specific binding	7/360	0.002062
3 hpi, -IFN	structural constituent of nuclear pore	5/360	0.002379
3 hpi, -IFN	telomeric DNA binding	6/360	0.004627
3 hpi, -IFN	steroid hormone receptor binding	9/360	0.007477
3 hpi, -IFN	protein C-terminus binding	14/360	0.007477
3 hpi, -IFN	dolichyl-diphosphooligosaccharide-protein glycotransferase activity	3/360	0.009688
0 hpi, +IFN	double-stranded RNA binding	20/370	5.38E-12
0 hpi, +IFN	helicase activity	23/370	1.05E-09
0 hpi, +IFN	mRNA binding	27/370	7.66E-09
0 hpi, +IFN	actin filament binding	23/370	1.73E-07
0 hpi, +IFN	RNA-dependent ATPase activity	11/370	3.18E-07
0 hpi, +IFN	structural constituent of ribosome	16/370	5.01E-06
0 hpi, +IFN	protein C-terminus binding	19/370	3.68E-05
0 hpi, +IFN	ligase activity	15/370	0.000111
0 hpi, +IFN	ADP binding	8/370	0.000228
0 hpi, +IFN	structural constituent of nuclear pore	6/370	0.000311
0 hpi, +IFN	ribonucleoprotein complex binding	13/370	0.00045
0 hpi, +IFN	S100 protein binding	5/370	0.000764
0 hpi, +IFN	importin-alpha family protein binding	3/370	0.005194
0 hpi, +IFN	telomeric DNA binding	6/370	0.005799
0 hpi, +IFN	regulatory RNA binding	6/370	0.007066
0 hpi, +IFN	carbon-oxygen lyase activity	8/370	0.008225
0 hpi, +IFN	phosphatidylinositol-3,4,5-trisphosphate binding	6/370	0.008626
0 hpi, +IFN	RNA polymerase core enzyme binding	6/370	0.008747
0 hpi, +IFN	catalytic activity, acting on DNA	13/370	0.009318
0 hpi, +IFN	protein kinase C binding	6/370	0.009663
1 hpi, +IFN	double-stranded RNA binding	20/400	2.36E-11
1 hpi, +IFN	mRNA binding	28/400	2.24E-08
1 hpi, +IFN	helicase activity	22/400	2.28E-08
1 hpi, +IFN	structural constituent of ribosome	19/400	1.3E-07

1 hpi, +IFN	DNA-dependent ATPase activity	16/400	1.81E-07
1 hpi, +IFN	ligase activity	18/400	4.26E-06
1 hpi, +IFN	aminoacyl-tRNA ligase activity	10/400	4.26E-06
1 hpi, +IFN	ligase activity, forming carbon-oxygen bonds	10/400	4.26E-06
1 hpi, +IFN	actin filament binding	20/400	3.3E-05
1 hpi, +IFN	ADP binding	9/400	4.44E-05
1 hpi, +IFN	ribonucleoprotein complex binding	15/400	4.91E-05
1 hpi, +IFN	structural constituent of nuclear pore	6/400	0.000376
1 hpi, +IFN	RNA polymerase core enzyme binding	7/400	0.00226
1 hpi, +IFN	protein C-terminus binding	16/400	0.002709
1 hpi, +IFN	carbon-oxygen lyase activity	9/400	0.002853
1 hpi, +IFN	aminoacyl-tRNA editing activity	4/400	0.002969
1 hpi, +IFN	importin-alpha family protein binding	3/400	0.005829
1 hpi, +IFN	catalytic activity, acting on DNA	14/400	0.006728
1 hpi, +IFN	telomeric DNA binding	6/400	0.007442
1 hpi, +IFN	regulatory RNA binding	6/400	0.009091
3 hpi, +IFN	double-stranded RNA binding	20/401	2.51E-11
3 hpi, +IFN	helicase activity	24/401	8.01E-10
3 hpi, +IFN	mRNA binding	27/401	4.76E-08
3 hpi, +IFN	structural constituent of ribosome	18/401	6.63E-07
3 hpi, +IFN	actin filament binding	23/401	6.75E-07
3 hpi, +IFN	RNA-dependent ATPase activity	11/401	6.75E-07
3 hpi, +IFN	ligase activity	19/401	6.99E-07
3 hpi, +IFN	ribonucleoprotein complex binding	17/401	2.68E-06
3 hpi, +IFN	ADP binding	10/401	4.67E-06
3 hpi, +IFN	structural constituent of nuclear pore	7/401	3.39E-05
3 hpi, +IFN	RNA polymerase core enzyme binding	7/401	0.002411
3 hpi, +IFN	carbon-oxygen lyase activity	9/401	0.003047
3 hpi, +IFN	aminoacyl-tRNA editing activity	4/401	0.00314
3 hpi, +IFN	importin-alpha family protein binding	3/401	0.005964
3 hpi, +IFN	catalytic activity, acting on DNA	14/401	0.007006
3 hpi, +IFN	protein C-terminus binding	15/401	0.007458
3 hpi, +IFN	telomeric DNA binding	6/401	0.007458
3 hpi, +IFN	S100 protein binding	4/401	0.009985
Biological Process			
Condition	**GO term**	**GeneRatio**	**p.adjust**
0 hpi, -IFN	ribosome biogenesis	41/389	2.25E-19
0 hpi, -IFN	rRNA processing	32/389	5.82E-16
0 hpi, -IFN	translational initiation	32/389	4.2E-15
0 hpi, -IFN	protein localization to endoplasmic reticulum	23/389	2.26E-11
0 hpi, -IFN	SRP-dependent cotranslational protein targeting to membrane	20/389	2.26E-11
0 hpi, -IFN	nucleic acid transport	27/389	2.26E-11

0 hpi, -IFN	RNA transport	27/389	2.26E-11
0 hpi, -IFN	establishment of RNA localization	27/389	2.3E-11
0 hpi, -IFN	RNA localization	29/389	2.3E-11
0 hpi, -IFN	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	22/389	2.39E-11
0 hpi, -IFN	regulation of mRNA processing	20/389	1.07E-07
0 hpi, -IFN	ribonucleoprotein complex subunit organization	25/389	3.46E-07
0 hpi, -IFN	nuclear export	22/389	8.98E-07
0 hpi, -IFN	ribonucleoprotein complex localization	17/389	9.43E-07
0 hpi, -IFN	ribonucleoprotein complex assembly	23/389	1.86E-06
0 hpi, -IFN	DNA geometric change	13/389	2.12E-05
0 hpi, -IFN	retrograde vesicle-mediated transport, Golgi to ER	13/389	2.19E-05
0 hpi, -IFN	protein-containing complex localization	23/389	2.19E-05
0 hpi, -IFN	ribose phosphate biosynthetic process	23/389	3.16E-05
0 hpi, -IFN	negative regulation of translation	16/389	3.16E-05
1 hpi, -IFN	ribosome biogenesis	41/356	6.98E-21
1 hpi, -IFN	rRNA processing	32/356	3.84E-17
1 hpi, -IFN	translational initiation	33/356	4E-17
1 hpi, -IFN	SRP-dependent cotranslational protein targeting to membrane	21/356	4.92E-13
1 hpi, -IFN	establishment of protein localization to endoplasmic reticulum	22/356	5.46E-13
1 hpi, -IFN	nucleic acid transport	27/356	2.02E-12
1 hpi, -IFN	RNA transport	27/356	2.02E-12
1 hpi, -IFN	RNA localization	29/356	2.02E-12
1 hpi, -IFN	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	22/356	3.17E-12
1 hpi, -IFN	ribonucleoprotein complex subunit organization	29/356	6.42E-11
1 hpi, -IFN	ribonucleoprotein complex assembly	27/356	4.8E-10
1 hpi, -IFN	nuclear export	21/356	9.41E-07
1 hpi, -IFN	ribonucleoprotein complex localization	16/356	1.76E-06
1 hpi, -IFN	regulation of mRNA processing	17/356	4.44E-06
1 hpi, -IFN	mitotic spindle organization	14/356	1.4E-05
1 hpi, -IFN	telomere maintenance	16/356	2.36E-05
1 hpi, -IFN	telomere organization	16/356	2.71E-05
1 hpi, -IFN	ribose phosphate biosynthetic process	22/356	2.73E-05
1 hpi, -IFN	negative regulation of translation	15/356	4.78E-05
1 hpi, -IFN	ribonucleotide biosynthetic process	21/356	6.27E-05
3 hpi, -IFN	ribosome biogenesis	37/362	9.23E-17
3 hpi, -IFN	RNA localization	33/362	5.94E-15
3 hpi, -IFN	translational initiation	31/362	6.28E-15
3 hpi, -IFN	rRNA metabolic process	32/362	9.58E-15
3 hpi, -IFN	nucleic acid transport	29/362	6.91E-14

3 hpi, -IFN	RNA transport	29/362	6.91E-14
3 hpi, -IFN	protein localization to endoplasmic reticulum	23/362	3.34E-12
3 hpi, -IFN	SRP-dependent cotranslational protein targeting to membrane	20/362	4.01E-12
3 hpi, -IFN	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	22/362	4.93E-12
3 hpi, -IFN	ribonucleoprotein complex export from nucleus	20/362	8.75E-10
3 hpi, -IFN	ribonucleoprotein complex localization	20/362	9.02E-10
3 hpi, -IFN	ribonucleoprotein complex subunit organization	27/362	2.84E-09
3 hpi, -IFN	ribonucleoprotein complex assembly	25/362	2.04E-08
3 hpi, -IFN	protein-containing complex localization	25/362	3.26E-07
3 hpi, -IFN	ribose phosphate biosynthetic process	24/362	2.62E-06
3 hpi, -IFN	telomere maintenance via telomere lengthening	13/362	4.23E-06
3 hpi, -IFN	telomere organization	17/362	6.5E-06
3 hpi, -IFN	ribonucleotide biosynthetic process	23/362	6.92E-06
3 hpi, -IFN	regulation of RNA splicing	16/362	8.23E-06
3 hpi, -IFN	mitotic spindle organization	14/362	1.48E-05
0 hpi, +IFN	ribosome biogenesis	37/369	1.75E-16
0 hpi, +IFN	rRNA metabolic process	32/369	3.32E-14
0 hpi, +IFN	translational initiation	28/369	4.44E-12
0 hpi, +IFN	ribonucleoprotein complex subunit organization	27/369	1.91E-08
0 hpi, +IFN	RNA localization	25/369	1.91E-08
0 hpi, +IFN	nuclear-transcribed mRNA catabolic process	23/369	8.55E-08
0 hpi, +IFN	ribonucleoprotein complex assembly	25/369	8.55E-08
0 hpi, +IFN	establishment of protein localization to endoplasmic reticulum	17/369	8.55E-08
0 hpi, +IFN	SRP-dependent cotranslational protein targeting to membrane	16/369	8.81E-08
0 hpi, +IFN	nucleic acid transport	21/369	3.26E-07
0 hpi, +IFN	RNA transport	21/369	3.26E-07
0 hpi, +IFN	positive regulation of mRNA metabolic process	13/369	4.48E-06
0 hpi, +IFN	negative regulation of translation	17/369	4.76E-06
0 hpi, +IFN	spindle organization	18/369	6.35E-06
0 hpi, +IFN	negative regulation of cellular amide metabolic process	18/369	8.73E-06
0 hpi, +IFN	telomere maintenance	17/369	9.29E-06
0 hpi, +IFN	regulation of viral life cycle	16/369	1.07E-05
0 hpi, +IFN	telomere organization	17/369	1.07E-05
0 hpi, +IFN	vesicle budding from membrane	14/369	1.64E-05
0 hpi, +IFN	viral genome replication	15/369	1.85E-05
1 hpi, +IFN	rRNA metabolic process	36/399	1.66E-16
1 hpi, +IFN	ribosome biogenesis	38/399	1.68E-16
1 hpi, +IFN	translational initiation	31/399	7.7E-14
1 hpi, +IFN	protein localization to endoplasmic reticulum	22/399	4.57E-10

1 hpi, +IFN	nuclear-transcribed mRNA catabolic process	27/399	4.57E-10
1 hpi, +IFN	SRP-dependent cotranslational protein targeting to membrane	19/399	4.75E-10
1 hpi, +IFN	RNA localization	26/399	9.46E-09
1 hpi, +IFN	nucleic acid transport	22/399	2.66E-07
1 hpi, +IFN	RNA transport	22/399	2.66E-07
1 hpi, +IFN	ribonucleoprotein complex subunit organization	25/399	8.02E-07
1 hpi, +IFN	positive regulation of mRNA metabolic process	14/399	1.39E-06
1 hpi, +IFN	ribonucleoprotein complex assembly	23/399	3.87E-06
1 hpi, +IFN	spindle organization	19/399	3.87E-06
1 hpi, +IFN	ER to Golgi vesicle-mediated transport	21/399	1.1E-05
1 hpi, +IFN	negative regulation of translation	17/399	1.15E-05
1 hpi, +IFN	negative regulation of cellular amide metabolic process	18/399	2.25E-05
1 hpi, +IFN	telomere maintenance	17/399	2.31E-05
1 hpi, +IFN	regulation of RNA stability	17/399	2.49E-05
1 hpi, +IFN	telomere organization	17/399	2.62E-05
1 hpi, +IFN	regulation of viral process	18/399	2.7E-05
3 hpi, +IFN	ribosome biogenesis	39/400	4.69E-17
3 hpi, +IFN	rRNA metabolic process	36/400	9.05E-17
3 hpi, +IFN	translational initiation	31/400	8.29E-14
3 hpi, +IFN	SRP-dependent cotranslational protein targeting to membrane	19/400	6.97E-10
3 hpi, +IFN	establishment of protein localization to endoplasmic reticulum	20/400	7E-10
3 hpi, +IFN	nuclear-transcribed mRNA catabolic process	26/400	2.31E-09
3 hpi, +IFN	RNA localization	26/400	1E-08
3 hpi, +IFN	regulation of viral process	22/400	9.1E-08
3 hpi, +IFN	ribonucleoprotein complex subunit organization	26/400	2.18E-07
3 hpi, +IFN	nucleic acid transport	22/400	2.3E-07
3 hpi, +IFN	RNA transport	22/400	2.3E-07
3 hpi, +IFN	viral genome replication	18/400	3.86E-07
3 hpi, +IFN	negative regulation of translation	19/400	4.49E-07
3 hpi, +IFN	ribonucleoprotein complex assembly	24/400	8.53E-07
3 hpi, +IFN	negative regulation of cellular amide metabolic process	19/400	5.04E-06
3 hpi, +IFN	telomere maintenance	18/400	5.04E-06
3 hpi, +IFN	regulation of RNA stability	18/400	5.45E-06
3 hpi, +IFN	telomere organization	18/400	5.9E-06
3 hpi, +IFN	spindle organization	18/400	1.38E-05
3 hpi, +IFN	cell junction assembly	23/400	2.47E-05
Cellular Component			
Condition	**GO term**	**GeneRatio**	**p.adjust**
0 hpi, -IFN	cytosolic ribosome	23/391	1.06E-13
0 hpi, -IFN	spliceosomal complex	25/391	1.05E-10

0 hpi, -IFN	cytosolic part	27/391	1.08E-09
0 hpi, -IFN	preribosome	14/391	3.28E-09
0 hpi, -IFN	vesicle coat	14/391	1.32E-08
0 hpi, -IFN	membrane coat	16/391	2.1E-07
0 hpi, -IFN	coated membrane	16/391	2.1E-07
0 hpi, -IFN	myelin sheath	19/391	1.71E-06
0 hpi, -IFN	brush border	14/391	2.74E-06
0 hpi, -IFN	actomyosin	13/391	2.74E-06
0 hpi, -IFN	chromosome, telomeric region	16/391	4.42E-06
0 hpi, -IFN	nuclear chromosome, telomeric region	14/391	6.45E-06
0 hpi, -IFN	aminoacyl-tRNA synthetase multienzyme complex	6/391	6.45E-06
0 hpi, -IFN	melanosome	14/391	1.25E-05
0 hpi, -IFN	pigment granule	14/391	1.25E-05
0 hpi, -IFN	endoplasmic reticulum-Golgi intermediate compartment	14/391	3.29E-05
0 hpi, -IFN	cell cortex part	16/391	0.000111
0 hpi, -IFN	nuclear periphery	14/391	0.000192
0 hpi, -IFN	nuclear pore	9/391	0.001017
0 hpi, -IFN	midbody	13/391	0.002276
1 hpi, -IFN	cytosolic ribosome	23/360	1.64E-14
1 hpi, -IFN	cytosolic part	26/360	1.13E-09
1 hpi, -IFN	preribosome	14/360	1.23E-09
1 hpi, -IFN	spliceosomal complex	21/360	1.71E-08
1 hpi, -IFN	vesicle coat	13/360	5.09E-08
1 hpi, -IFN	membrane coat	15/360	5.11E-07
1 hpi, -IFN	coated membrane	15/360	5.11E-07
1 hpi, -IFN	stress fiber	12/360	1.77E-06
1 hpi, -IFN	contractile actin filament bundle	12/360	1.77E-06
1 hpi, -IFN	myelin sheath	17/360	8.22E-06
1 hpi, -IFN	chromosome, telomeric region	14/360	3.74E-05
1 hpi, -IFN	nuclear chromosome, telomeric region	12/360	7.86E-05
1 hpi, -IFN	aminoacyl-tRNA synthetase multienzyme complex	5/360	9.06E-05
1 hpi, -IFN	melanosome	12/360	0.000138
1 hpi, -IFN	brush border	11/360	0.000149
1 hpi, -IFN	oligosaccharyltransferase complex	4/360	0.000494
1 hpi, -IFN	cell cortex part	14/360	0.000625
1 hpi, -IFN	midbody	13/360	0.00109
1 hpi, -IFN	endoplasmic reticulum-Golgi intermediate compartment	11/360	0.001181
1 hpi, -IFN	nucleoid	6/360	0.00411
3 hpi, -IFN	cytosolic ribosome	22/364	2.75E-13
3 hpi, -IFN	cytosolic part	28/364	5.26E-11
3 hpi, -IFN	vesicle coat	14/364	9.03E-09

3 hpi, -IFN	spliceosomal complex	21/364	2.17E-08
3 hpi, -IFN	myelin sheath	21/364	3.33E-08
3 hpi, -IFN	small-subunit processome	9/364	6.38E-07
3 hpi, -IFN	membrane coat	15/364	6.38E-07
3 hpi, -IFN	coated membrane	15/364	6.38E-07
3 hpi, -IFN	actomyosin	13/364	1.45E-06
3 hpi, -IFN	aminoacyl-tRNA synthetase multienzyme complex	6/364	5.24E-06
3 hpi, -IFN	melanosome	13/364	3.17E-05
3 hpi, -IFN	brush border	12/364	3.17E-05
3 hpi, -IFN	chromosome, telomeric region	14/364	3.92E-05
3 hpi, -IFN	cell cortex part	16/364	4.98E-05
3 hpi, -IFN	nuclear chromosome, telomeric region	12/364	7.65E-05
3 hpi, -IFN	endoplasmic reticulum-Golgi intermediate compartment	12/364	0.000331
3 hpi, -IFN	ruffle	14/364	0.000646
3 hpi, -IFN	midbody	13/364	0.001186
3 hpi, -IFN	costamere	5/364	0.001514
3 hpi, -IFN	invadopodium	4/364	0.002898
0 hpi, +IFN	cytosolic ribosome	20/375	6.86E-11
0 hpi, +IFN	myelin sheath	24/375	5.93E-10
0 hpi, +IFN	vesicle coat	15/375	1.37E-09
0 hpi, +IFN	spliceosomal complex	21/375	5.44E-08
0 hpi, +IFN	membrane coat	16/375	1.52E-07
0 hpi, +IFN	coated membrane	16/375	1.52E-07
0 hpi, +IFN	cytosolic part	23/375	1.71E-07
0 hpi, +IFN	melanosome	16/375	4.11E-07
0 hpi, +IFN	pigment granule	16/375	4.11E-07
0 hpi, +IFN	preribosome	11/375	1.46E-06
0 hpi, +IFN	stress fiber	12/375	2.68E-06
0 hpi, +IFN	contractile actin filament bundle	12/375	2.68E-06
0 hpi, +IFN	brush border	12/375	4.79E-05
0 hpi, +IFN	endoplasmic reticulum-Golgi intermediate compartment	13/375	0.00012
0 hpi, +IFN	aminoacyl-tRNA synthetase multienzyme complex	5/375	0.00012
0 hpi, +IFN	midbody	15/375	0.000144
0 hpi, +IFN	costamere	6/375	0.000196
0 hpi, +IFN	chromosome, telomeric region	13/375	0.000251
0 hpi, +IFN	cell cortex part	15/375	0.000286
0 hpi, +IFN	cell-cell adherens junction	12/375	0.000407
1 hpi, +IFN	cytosolic ribosome	22/405	2.58E-12
1 hpi, +IFN	vesicle coat	15/405	4.01E-09
1 hpi, +IFN	preribosome	14/405	6.27E-09
1 hpi, +IFN	cytosolic part	26/405	9.29E-09

1 hpi, +IFN	membrane coat	16/405	4.07E-07
1 hpi, +IFN	coated membrane	16/405	4.07E-07
1 hpi, +IFN	spliceosomal complex	20/405	4.93E-07
1 hpi, +IFN	stress fiber	13/405	9.72E-07
1 hpi, +IFN	contractile actin filament bundle	13/405	9.72E-07
1 hpi, +IFN	aminoacyl-tRNA synthetase multienzyme complex	6/405	9.25E-06
1 hpi, +IFN	myelin sheath	18/405	9.72E-06
1 hpi, +IFN	melanosome	14/405	2.17E-05
1 hpi, +IFN	endoplasmic reticulum-Golgi intermediate compartment	14/405	5.58E-05
1 hpi, +IFN	costamere	6/405	0.000325
1 hpi, +IFN	chromosome, telomeric region	13/405	0.00058
1 hpi, +IFN	cell cortex part	15/405	0.0007
1 hpi, +IFN	nuclear chromosome, telomeric region	11/405	0.001047
1 hpi, +IFN	brush border	10/405	0.001979
1 hpi, +IFN	midbody	13/405	0.003522
1 hpi, +IFN	host	5/405	0.007632
3 hpi, +IFN	cytosolic ribosome	22/406	2.74E-12
3 hpi, +IFN	vesicle coat	15/406	5.6E-09
3 hpi, +IFN	cytosolic part	26/406	1.73E-08
3 hpi, +IFN	preribosome	13/406	5.11E-08
3 hpi, +IFN	membrane coat	17/406	6.34E-08
3 hpi, +IFN	coated membrane	17/406	6.34E-08
3 hpi, +IFN	catalytic step 2 spliceosome	15/406	9.39E-08
3 hpi, +IFN	spliceosomal complex	21/406	9.39E-08
3 hpi, +IFN	myelin sheath	21/406	1.5E-07
3 hpi, +IFN	aminoacyl-tRNA synthetase multienzyme complex	7/406	4.83E-07
3 hpi, +IFN	cell cortex part	19/406	4.44E-06
3 hpi, +IFN	melanosome	15/406	5.12E-06
3 hpi, +IFN	stress fiber	12/406	5.23E-06
3 hpi, +IFN	contractile actin filament bundle	12/406	5.23E-06
3 hpi, +IFN	chromosome, telomeric region	16/406	7.33E-06
3 hpi, +IFN	nuclear chromosome, telomeric region	13/406	4.82E-05
3 hpi, +IFN	midbody	16/406	8.24E-05
3 hpi, +IFN	cell-cell adherens junction	13/406	0.000198
3 hpi, +IFN	brush border	11/406	0.000391
3 hpi, +IFN	endoplasmic reticulum-Golgi intermediate compartment	12/406	0.000878
KEGG Pathway			
Condition	**KEGG term**	**GeneRatio**	**p.adjust**
0 hpi, -IFN	Ribosome	19/252	6.15E-05
0 hpi, -IFN	Spliceosome	18/252	9.68E-05
0 hpi, -IFN	Ribosome biogenesis in eukaryotes	14/252	0.000341

0 hpi, -IFN	Protein processing in endoplasmic reticulum	17/252	0.000937
0 hpi, -IFN	HIF-1 signaling pathway	12/252	0.004913
0 hpi, -IFN	Fatty acid elongation	6/252	0.004913
0 hpi, -IFN	Biosynthesis of unsaturated fatty acids	6/252	0.004913
0 hpi, -IFN	Pathogenic Escherichia coli infection	17/252	0.004977
0 hpi, -IFN	Aminoacyl-tRNA biosynthesis	9/252	0.004977
0 hpi, -IFN	Salmonella infection	17/252	0.00885
1 hpi, -IFN	Ribosome	20/224	1.57E-06
1 hpi, -IFN	Ribosome biogenesis in eukaryotes	13/224	0.000672
1 hpi, -IFN	Protein processing in endoplasmic reticulum	16/224	0.000856
1 hpi, -IFN	Spliceosome	15/224	0.000856
1 hpi, -IFN	Fatty acid elongation	6/224	0.002803
1 hpi, -IFN	Biosynthesis of unsaturated fatty acids	6/224	0.002803
1 hpi, -IFN	RNA transport	15/224	0.004133
1 hpi, -IFN	Salmonella infection	16/224	0.008139
3 hpi, -IFN	Ribosome	19/233	1.86E-05
3 hpi, -IFN	Spliceosome	17/233	0.000155
3 hpi, -IFN	HIF-1 signaling pathway	14/233	0.000222
3 hpi, -IFN	Ribosome biogenesis in eukaryotes	13/233	0.000573
3 hpi, -IFN	Protein processing in endoplasmic reticulum	16/233	0.001124
3 hpi, -IFN	Biosynthesis of amino acids	10/233	0.00217
3 hpi, -IFN	Pathogenic Escherichia coli infection	17/233	0.002552
3 hpi, -IFN	Glycolysis / Gluconeogenesis	9/233	0.004038
3 hpi, -IFN	Salmonella infection	17/233	0.004038
3 hpi, -IFN	Tight junction	14/233	0.009055
0 hpi, +IFN	Pathogenic Escherichia coli infection	21/233	7.99E-05
0 hpi, +IFN	Protein processing in endoplasmic reticulum	18/233	0.000136
0 hpi, +IFN	Ribosome biogenesis in eukaryotes	14/233	0.000136
0 hpi, +IFN	Ribosome	16/233	0.000482
0 hpi, +IFN	Biosynthesis of unsaturated fatty acids	6/233	0.003605
0 hpi, +IFN	Spliceosome	14/233	0.003605
0 hpi, +IFN	Tight junction	15/233	0.003605
0 hpi, +IFN	Salmonella infection	17/233	0.004386
0 hpi, +IFN	RNA transport	15/233	0.005322
1 hpi, +IFN	Ribosome	19/250	6.22E-05
1 hpi, +IFN	Ribosome biogenesis in eukaryotes	15/250	9.8E-05
1 hpi, +IFN	Protein processing in endoplasmic reticulum	18/250	0.00032
1 hpi, +IFN	Pathogenic Escherichia coli infection	19/250	0.001037
1 hpi, +IFN	Aminoacyl-tRNA biosynthesis	10/250	0.001683
1 hpi, +IFN	Salmonella infection	18/250	0.004708
1 hpi, +IFN	HIF-1 signaling pathway	12/250	0.004708
1 hpi, +IFN	Biosynthesis of unsaturated fatty acids	6/250	0.004708
1 hpi, +IFN	Spliceosome	14/250	0.006177

1 hpi, +IFN	Legionellosis	8/250	0.008979
3 hpi, +IFN	Ribosome	18/255	0.000187
3 hpi, +IFN	Pathogenic Escherichia coli infection	21/255	0.000187
3 hpi, +IFN	Ribosome biogenesis in eukaryotes	14/255	0.000414
3 hpi, +IFN	Protein processing in endoplasmic reticulum	17/255	0.001157
3 hpi, +IFN	Aminoacyl-tRNA biosynthesis	10/255	0.001844
3 hpi, +IFN	Salmonella infection	18/255	0.004946
3 hpi, +IFN	Biosynthesis of unsaturated fatty acids	6/255	0.004946
3 hpi, +IFN	HIF-1 signaling pathway	12/255	0.004946
3 hpi, +IFN	Spliceosome	14/255	0.007013
3 hpi, +IFN	Legionellosis	8/255	0.009491

Table 3.2 Enriched GO terms and KEGG pathways for VIR-CLASP with IAV. GO terms were collapsed by semantic similarity. For brevity, only top 20 terms by statistical significance (adjusted p-value) are displayed. Full list of enriched GO terms can be accessed in (B. Kim et al., 2020).

Molecular Function		
GO term	**GeneRatio**	**p.adjust**
DNA-dependent ATPase activity	17/305	1.82E-09
ATP-dependent helicase activity	15/305	7.18E-09
purine NTP-dependent helicase activity	15/305	7.18E-09
double-stranded RNA binding	12/305	4.58E-06
ligase activity	15/305	1.1E-05
catalytic activity, acting on DNA	15/305	0.00016
mRNA 5'-UTR binding	6/305	0.000211
DNA topoisomerase activity	4/305	0.000311
microtubule binding	16/305	0.000414
single-stranded DNA binding	10/305	0.001097
structural constituent of nuclear pore	5/305	0.001114
calmodulin binding	14/305	0.001114
ribonucleoprotein complex binding	10/305	0.004762
RNA polymerase binding	7/305	0.005938
actin filament binding	13/305	0.005938

Biological Process		
GO term	**GeneRatio**	**p.adjust**
RNA localization	25/302	1.3E-09
nucleic acid transport	20/302	2.01E-07
RNA transport	20/302	2.01E-07
nucleobase-containing compound transport	22/302	2.01E-07
telomere maintenance	17/302	1.86E-06
telomere organization	17/302	2E-06
mitotic cytokinesis	12/302	2.21E-05
positive regulation of viral process	11/302	0.000102
viral genome replication	13/302	0.00011
apical junction assembly	9/302	0.000316
spindle organization	14/302	0.000321
positive regulation of chromosome organization	15/302	0.000321
establishment of cell polarity	13/302	0.000321
ribonucleoprotein complex localization	12/302	0.000366
protein-containing complex localization	18/302	0.000371
tight junction assembly	8/302	0.000377
tight junction organization	8/302	0.000648
cytoskeleton-dependent intracellular transport	14/302	0.000942
cortical actin cytoskeleton organization	7/302	0.000978
regulation of gene expression, epigenetic	17/302	0.000978
Cellular Component		
GO term	**GeneRatio**	**p.adjust**
ribonucleoprotein granule	19/304	1.29E-05
actomyosin	12/304	1.29E-05
midbody	16/304	1.57E-05
cell-cell adherens junction	13/304	6.21E-05
brush border	11/304	0.000122
aminoacyl-tRNA synthetase multienzyme complex	5/304	0.000134
cell cortex part	14/304	0.000345
nuclear pore	9/304	0.000518
lateral plasma membrane	8/304	0.000663
cell division site part	8/304	0.000948
melanosome	10/304	0.001607
pigment granule	10/304	0.001607
small-subunit processome	5/304	0.003494
chromosome, telomeric region	10/304	0.00424
myelin sheath	11/304	0.006619
KEGG Pathways		
KEGG Term	**GeneRatio**	**p.adjust**
Spliceosome	12/170	0.006346

Adherens junction	8/170	0.013444
Protein processing in endoplasmic reticulum	12/170	0.015657
Aminoacyl-tRNA biosynthesis	7/170	0.026032
RNA transport	11/170	0.034062
Tight junction	11/170	0.034599
Pathogenic Escherichia coli infection	12/170	0.037471

Table 3.3 Hits from VIR-CLASP with CHIKV that were identified in screens for CHIKV pro- and anti-viral factors. Full list of overlapping hits from VIR-CLASP for CHIKV and other screens (Karlas et al., 2016; Ooi et al., 2013; Radoshitzky et al., 2016; Schoggins et al., 2011). Screen results are colored by "pro-viral" (green) and "anti-viral" (yellow), and VIR-CLASP conditions in which they were identified are listed.

Gene Symbol	Karlas	Ooi	Radoshitzky	Schoggins
CLTC	All conditions		All conditions	
CNOT1	1hpi, -IFN			
KDM1A	0hpi, -IFN; 3hpi, -IFN			
EEF2	All conditions			
IQGAP1	All conditions			
FASN	All conditions			
ACLY	All conditions			
RACK1	All conditions			
PPIL2	0hpi, +IFN			
BOP1	1hpi, -IFN; 1hpi, +IFN; 3hpi, +IFN			
GCN1	All conditions			
DHX38	All hpi, -IFN			
AP2M1	All conditions			
PDLIM5	All conditions			
SIN3A	All conditions			
TECR	All hpi, -IFN			
RRBP1	All conditions			

COPG1	All conditions			
STK10		1hpi, +IFN		
ARCN1		All conditions		
ADAR				All conditions
GALNT2				3hpi, -IFN; 3hpi, +IFN
OAS3				0hpi, -IFN; 0hpi, +IFN; 1hpi, +IFN; 3hpi, +IFN
STAU1			1hpi, +IFN	
COPA			All conditions	
VCP			All conditions	
RAC1			0hpi, -IFN	
PTBP1	All hpi, -IFN			
VARS		All conditions		
YARS		0hpi, -IFN; 1hpi, +IFN; 3hpi, -IFN; 3hpi, +IFN		
OGDH		1hpi, +IFN		
STT3B	All hpi, -IFN			
LIMA1		All conditions		

CHAPTER IV

YTHDF2 BINDS TO NON-M^{6}A-MODIFIED RNA IN CELLS TO REGULATE TRANSLATION EFFICIENCY

Introduction

YTHDF proteins are thought to bind m^{6}A-modified RNA to regulate various aspects of RNA metabolism including stability (X. Wang et al., 2014), translation (X. Wang et al., 2015a), and localization (Ries et al., 2019). Through studies of RNA viruses, it is clear that these ubiquitous RNA metabolism roles for YTHDF proteins just scratch the surface of their biological potential. Through binding to specific m^{6}A sites on viral RNAs, YTHDFs can also inhibit innate immune recognition (G.-W. Kim et al., 2020; M. Lu et al., 2020), promote nuclear export of viral transcripts (Lichinchi et al., 2016a), engage in viral packaging processes (W. Lu et al., 2018), and interact with viral proteins to either promote or inhibit viral replication (Jurczyszak et al., 2020; W. Lu et al., 2018). As is the case with other studies of host protein roles in viral replication, it is likely that some of these functions of YTHDFs on viral RNA are also at play in host processes.

Perhaps the most intriguing observations of YTHDF function come from the technological advances of methylated RNA immunoprecipitation and sequencing (meRIP-seq or m^{6}A-seq) (Dominissini et al., 2012; Meyer et al., 2012) and cross-linking and immunopurification (CLIP-seq) (Sugimoto et al., 2012) assays. Briefly, meRIP-seq is used to profile m^{6}A sites across the transcriptome in cells; CLIP-seq is used to identify the RNA targets of an RNA-binding protein

of interest. Through combined CLIP-seq for YTHDFs and meRIP-seq, it has become clear that YTHDF proteins likely bind non-m^{6}A-modified (hereafter simplified as "unmodified") RNA sites in cells.

Efforts to map both YTHDF and m^{6}A sites in cellular or viral RNAs has revealed that between 40% - 80% of identified YTHDF sites do not overlap with an m^{6}A site (Gokhale et al., 2016; X. Wang et al., 2015a; 2014). This observation has held true for human cells (X. Wang et al., 2015a), viral RNAs (Gokhale et al., 2016), and mouse cells (Yuan Zhang et al., 2019). While it is possible that the lack of overlap may be an artefact of the peak calling algorithms, the consistency of this observation across multiple biological contexts and biochemical techniques suggests it may have functional importance.

Preliminary *in vitro* data of YTHDF binding preferences has confirmed that the YTH domain as a ~10 fold preference for m^{6}A-modified RNA probes over unmodified RNA (Zhu et al., 2014). Crystal structures have been solved for the YTH domain bound to methylated oligonucleotides (Zhu et al., 2014). Yet, much of the remainder of the YTHDF proteins remains poorly characterized. This is predominantly due to the inability of the N-terminal region to form crystal structures, likely indicating unstructured or disordered domains (Fu and Zhuang, 2020). Furthermore, "intrinsically disordered regions" are known to participate in RNA-binding (Calabretta and Richard, 2015). Thus, if YTHDF proteins indeed bind unmodified RNAs, one possible mechanism is through interactions between the uncharacterized regions of the proteins.

An alternative possibility to explain how YTHDF proteins may bind unmodified RNAs is through interactions with other protein binding partners. Multiple reports have shown that YTHDF proteins interact with various cellular partners to effect changes in RNA metabolism (Du et al., 2016; X. Wang et al., 2015a). Whether these binding partners may also contribute to the

binding specificity of YTHDF protein is unknown. Importantly, CLIP assays utilize photo-crosslinking of RNAs to proteins followed by stringent high-salt wash steps (Ascano et al., 2012; Licatalosi et al., 2008). Therefore, the YTHDF binding sites captured by CLIP assays should indicate a direct interaction between YTHDFs and RNA; indirect associations between the protein of interest and RNA through a multi-protein complex are not captured.

In this report, I will demonstrate that YTHDF2 indeed binds unmodified RNAs in cells through careful examination of potential artefacts due to peak calling. Next, I show that unmodified YTHDF2 peaks have distinct sequence motifs and have increased density at start and stop codons compared to m^{6}A-modified peaks. I then demonstrate that YTHDF2 has distinct roles in mRNA metabolism based on the m^{6}A status of its binding site: mRNAs with mostly unmodified YTHDF2 peaks have increased translation efficiency, while mRNAs with mostly m^{6}A-modified peaks have decreased stability. Lastly, I show that the proteins encoded by YTHDF2 targets are distinct in biological function and subcellular localization based upon m^{6}A status. These findings have important relevance for our understanding of m^{6}A and YTHDF function.

Results

YTHDF2 binds unmodified RNAs in cells

Recent work has shown that many YTHDF2 peaks do not overlap with m^{6}A sites in cells. This observation has at least three possible explanations. First, the YTHDF peaks that do not overlap with m^{6}A sites may be false positives; this would suggest that increased stringency in YTHDF2 peak calling would increase the overlap between YTHDF2 and m^{6}A sites. Second, the

unmodified YTHDF2 peaks may be false negative m^{6}A peaks; this suggests that relaxing the stringency of m^{6}A peak calling could rescue the overlap between YTHDF2 and m^{6}A. The third possibility is that the unmodified peaks actually represent a biologically relevant category of YTHDF2 targets. To investigate which of these three possibilities is most consistent with the data, I first generated sets of YTHDF2 and m^{6}A peaks from published raw PAR-CLIP and m^{6}A-seq data from HeLa cells (X. Wang et al., 2014). Using default parameters for both PAR-CLIP and m^{6}A-seq peak callers (PARalyzer V1.5 and MACS2) led to identification of 13127 peaks for YTHDF2 and 12260 peaks for m^{6}A (Figure 4.1A). Overlap of these two sets of peaks revealed that $< 50\%$ of the YTHDF2 peaks overlapped with an m^{6}A peak by at least one nucleotide.

To test the first possibility, that the unmodified YTHDF2 peaks are in fact false-positives, I examined the fraction of T-C mutations in YTHDF2 peaks, a key scoring metric indicating biochemical confidence in the validity of the peak. In PAR-CLIP, the diagnostic T-C mutation is introduced into cDNA during the reverse-transcription step at the site of the protein-4SU crosslink (Hafner et al., 2010). The frequency of this mutation among reads comprising a "peak" is an indicator of biochemical confidence. False positives are known to be more common in lower-scoring peaks (Ascano et al., 2012; Hafner et al., 2010), thus if the unmodified YTHDF2 peaks are false-positives, one would expect to see that the T-C fraction is lower on average than in the m^{6}A-modified peaks. In fact, the opposite was observed: while the distribution of T-C fraction is similar between unmodified and m^{6}A-modified YTHDF2 peaks, the unmodified peaks have significantly higher T-C fractions on average (Figure 4.1B). This demonstrates that the unmodified YTHDF2 are unlikely to be false positives.

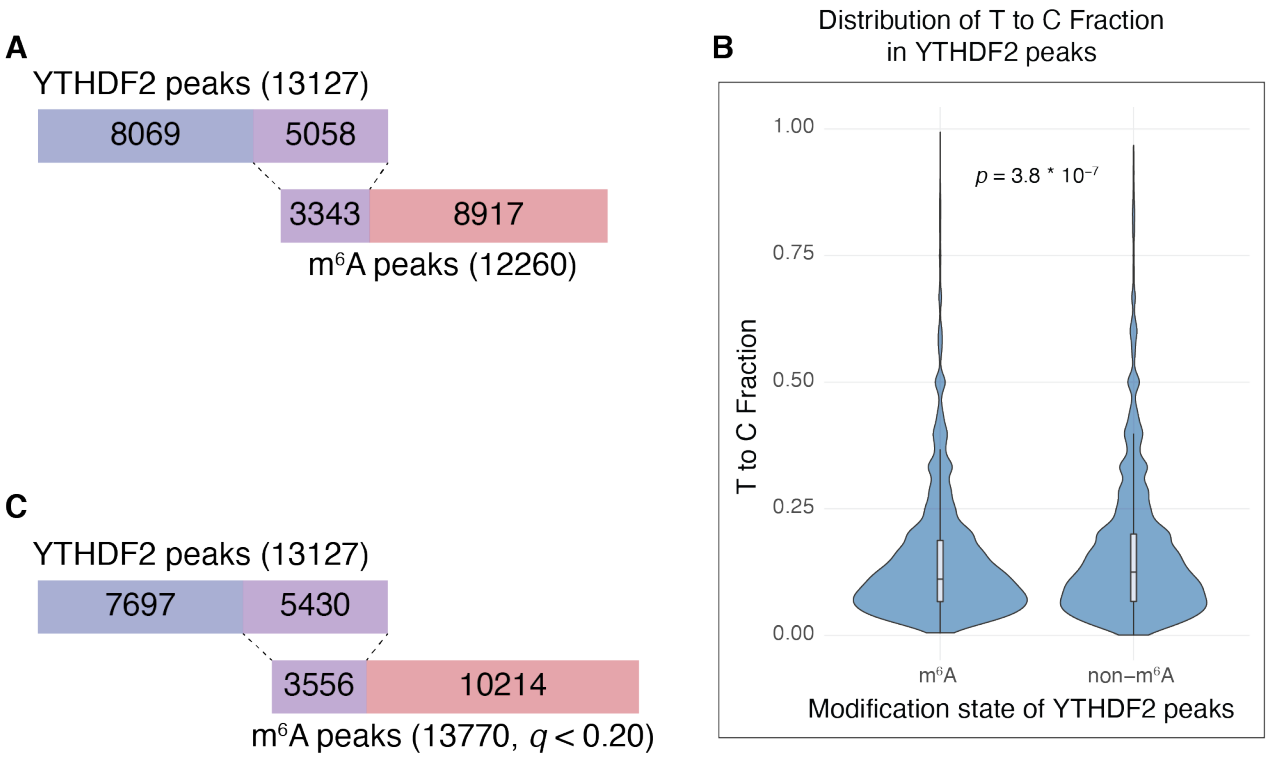


Figure 4.1 Overlap between YTHDF2 peaks and m^{6}A peaks in HeLa cells. A) Schematic showing overlapped peaks (minimum 1 nucleotide) between YTHDF2 PAR-CLIP and meRIP-seq in HeLa cells. B) Distribution of T-C fraction in YTHDF2 peaks categorized by overlap with m^{6}A. C) Schematic showing overlapped peaks as in (A), but with a relaxed significance threshold (q < 0.20).

To examine whether the unmodified YTHDF2 peaks are false-negatives for m^{6}A, I called peaks again on the m^{6}A-seq data using a relaxed cut-off for statistical significance (FDR < 0.20). This led to identification of 13700 peaks for m^{6}A (~ 1500 more than with FDR < 0.05). However, the overlap of YTHDF2 peaks with m^{6}A only increased by ~ 400 peaks, and the ratio of unmodified to m^{6}A-modified YTHDF2 peaks remained roughly the same (63% versus 70%). Thus, the category of unmodified YTHDF2 peaks is not an artefact of the stringency of m^{6}A peak calling. Together, these experiments support the model that unmodified YTHDF2 binding sites are a *bona fide* category of YTHDF2 targets in cells.

Unmodified YTHDF2 peaks have distinct binding determinants

I next examined whether the unmodified YTHDF2 sites are distinct in either sequence motif or location compared to the m^{6}A-modified sites. *De novo* motif finding revealed that m^{6}A-modified YTHDF2 peaks are highly enriched for the canonical m^{6}A motif, RRACH: three out of four of the top motifs contain GAC or CAC. This motif was first defined via chromatographic methods in the late 20$^{\text{th}}$ century for *in vitro* (Harper et al., 1990), mouse (Schibler et al., 1977), virus (Kane and Beemon, 1985), and bovine (Horowitz et al., 1984) RNAs; it was confirmed more recently by m^{6}A-seq on various human and viral RNAs. However, none of the top four motifs identified for unmodified YTHDF2 peaks contained an RRACH sequence. Instead, the unmodified YTHDF2 peaks are enriched for C-rich sequences, in particular CUC or CUUC. Interestingly, the CUC or CUUC motif was also identified for the m^{6}A-modified YTHDF2 peaks.

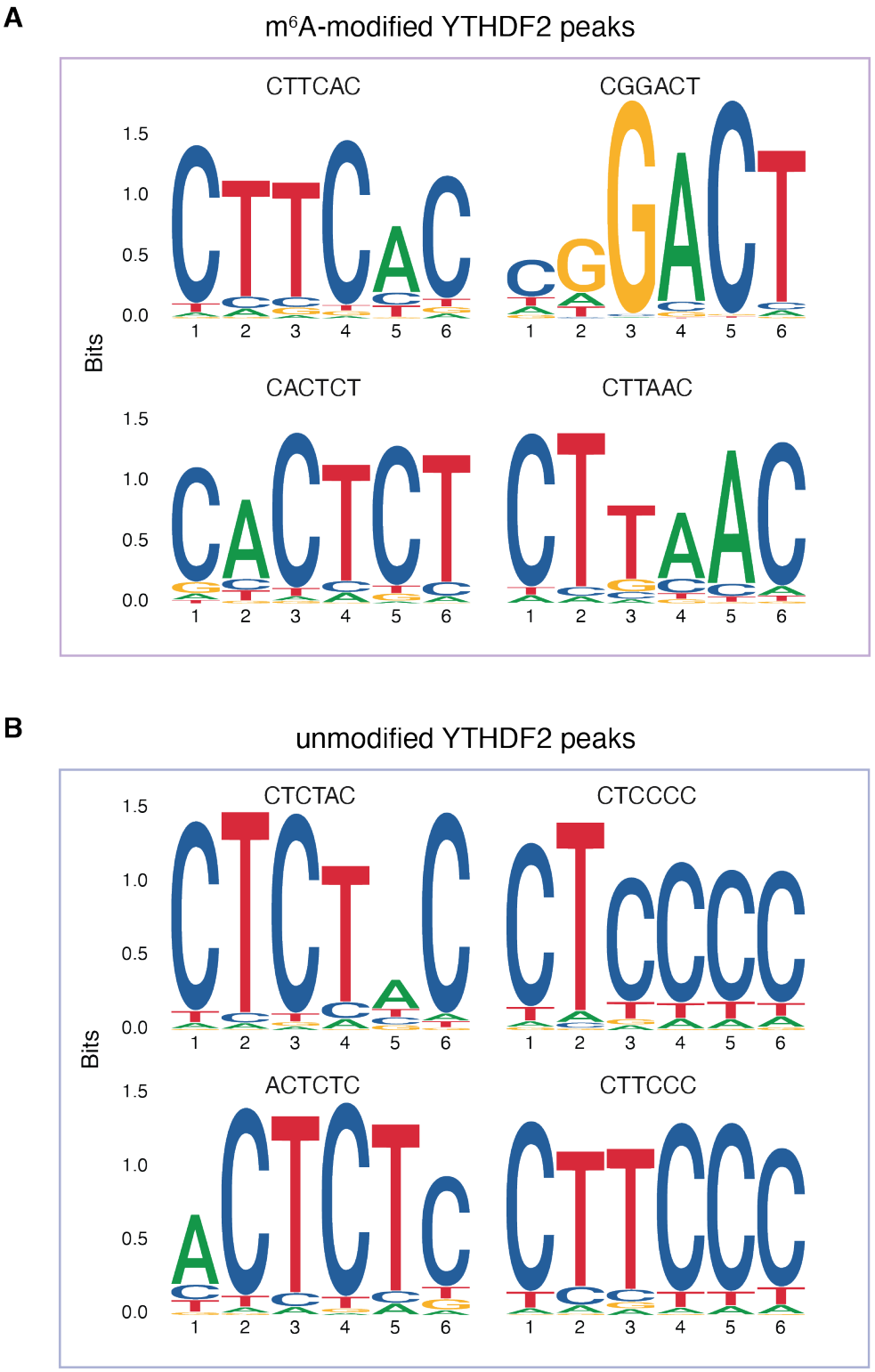


Figure 4.2 Enriched motifs in YTHDF2 peaks by m^{6}A status. A) Top 4 enriched motifs in mRNA binding sites of YTHDF2 that overlap with an m^{6}A site. B) Top 4 enriched motifs in mRNA binding sites of YTHDF2 that do not overlap with an m^{6}A site.

The unmodified and m^{6}A-modified peaks have distinct preferences for binding different transcript regions. When looking at transcript regions from a global perspective, unmodified YTHDF2 sites have a slight preference for 3' and 5' untranslated regions (UTRs) compared to m^{6}A-modified regions (Figure 4.3A). Both categories of YTHDF2 peaks bind to exonic sequences about three times as much as intronic sequences (Figure 4.3B). However, zooming in to the 3' UTR and 5' UTR boundaries with the coding sequence revealed that unmodified YTHDF2 sites have high density around the start and stop codons, while m^{6}A-modified YTHDF2 sites have much lower density at these regions (Figure 4.3C). In particular, the region around the stop codon has a distinct signal of high enrichment for unmodified peaks, while the m^{6}A-modified peaks show more general enrichment around the end of the coding sequence and 3' UTR. Consistent with Figure 4.3A, there is a higher density of both unmodified and m^{6}A-modified peaks around the 3' UTR than the 5' UTR. Together, the distinct motif and location signatures between unmodified and m^{6}A-modified YTHDF2 peaks suggests that these two categories of YTHDF2 sites may have different functions in mRNA metabolism.

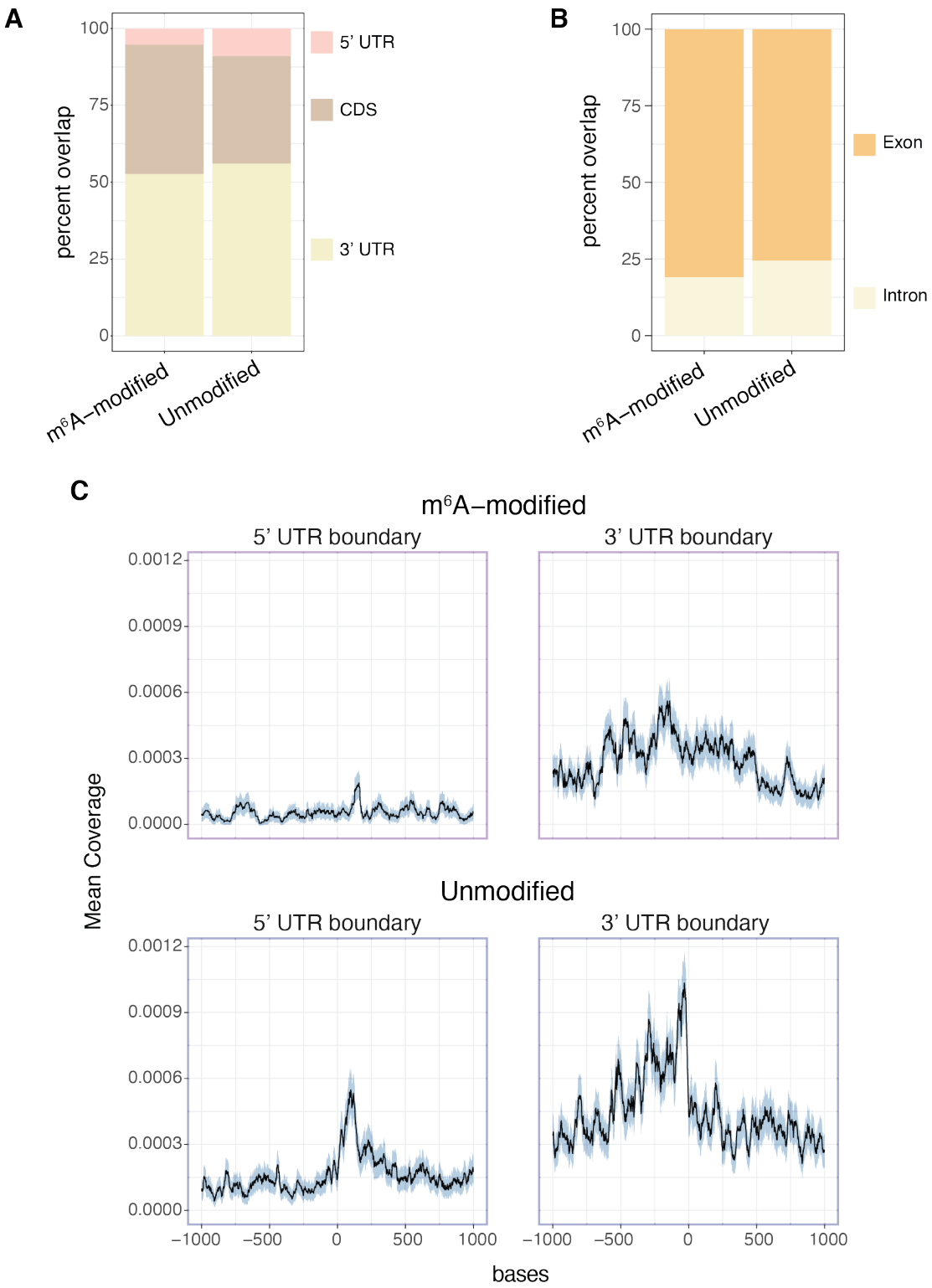


Figure 4.3 Transcript regions bound by YTHDF2 by m^{6}A status . A) Location of YTHDF2 peaks in major transcript features (3' UTR, CDS, 5' UTR) by m6A status. B) Location of YTHDF2 peaks in exons or introns by m^{6}A status. C) Mean peak coverage by YTHDF2 at 5' and 3' CDS boundaries. 0 indicates start (5') or end (3') of CDS, with 5' UTR indicated by negative numbers (right) and 3' UTR indicated by positive numbers (left).

YTHDF2 alters mRNA metabolism based on m^{6}A status

The assortment of YTHDF2 **peaks** into unmodified and m^{6}A-modified categories is straightforward, based on the overlap of YTHDF2 peaks and m^{6}A peaks described above (Figure 4.1A, 4.4A). However, when examining whether YTHDF2 can alter the metabolism of **whole target mRNAs**, classifying targets based on overlap between peaks can be done in many ways. To classify target genes based on YTHDF2 and m^{6}A peaks, I sorted them into two categories (Figure 4.4B). Targets upon which $> 50\%$ of YTHDF2 peaks overlap with an m^{6}A site are considered "mostly m^{6}A modified". Target genes upon which $< 50\%$ of YTHDF2 peaks overlap with an m^{6}A site are classified as "mostly unmodified".

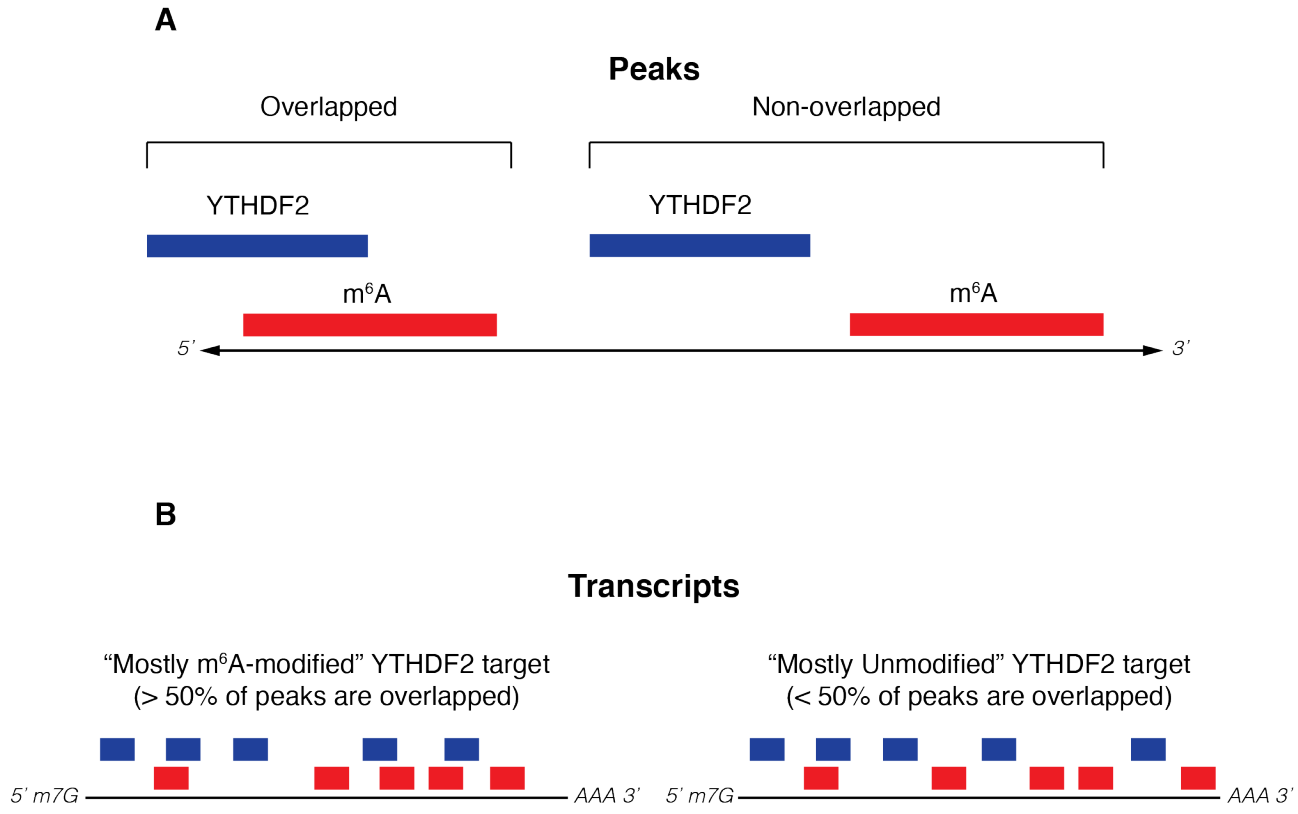


Figure 4.4 Schematic of categorization of YTHDF2 transcript targets by m^{6}A status of the binding sites. A) Schematic of overlapping peaks from YTHDF2 PAR-CLIP or m^{6}A-seq assays. Overlap between YTHDF2 peaks and m6A peaks can be easily calculated by detecting overlapping nucleotides (minimum 1nt). B) Schematic demonstrating how YTHDF2 target transcripts were categorized based on the fraction of YTHDF2 peaks that overlap with m^{6}A peaks. Transcripts bound by YTHDF2 were categorized into "Mostly m^{6}A-modified" if greater than 50% of YTHDF2 peaks on that transcript were m^{6}A modified. Transcripts were categorized into "Mostly unmodified" if less than 50% of the YTHDF2 peaks on that transcript were m^{6}A modified.

To examine whether targets with "mostly unmodified" or "mostly m^6A-modified" YTHDF2 peaks have distinct functions in mRNA metabolism I analyzed published RNA half-life and ribosome profiling sequencing datasets (X. Wang et al., 2014). These datasets were generated by transcription inhibition and RNA-seq to measure RNA half-lives, and ribosome profiling to measure translation efficiency. The initial report found that knockdown of YTHDF2 alters mRNA half-life but has no effect on mRNA translation efficiency (X. Wang et al., 2014). To determine whether YTHDF2 may have different functions based on m^6A occupancy, I first assorted YTHDF2 target genes based on m^6A occupancy as described above (Figure 4.4B), then plotted the cumulative distribution functions for mRNA half-life or translation efficiency of both categories as well as non-target genes. YTHDF2 targets that are "mostly m^6A-modified" had shorter mRNA half-lives (median 3.35 hrs) compared to and non-targets (median 5.31 hrs) as described in previous reports. However, the "mostly unmodified" targets had half-lives similar to non-targets (median 5.86 hrs). Upon YTHDF2 knockdown, the half-lives of "mostly m6A-modified" targets increased close to non-target levels (median 4.41 hrs vs. 5.43 hrs), while "mostly unmodified" targets remained similar to the control (median 6.19 hrs) (Figure 4.5A). The difference in mRNA half-life change between "mostly m^6A-modified" and "mostly unmodified" YTHDF2 targets upon YTHDF2 knockdown was statistically significant ($p = 4.534 * 10^{-7}$).

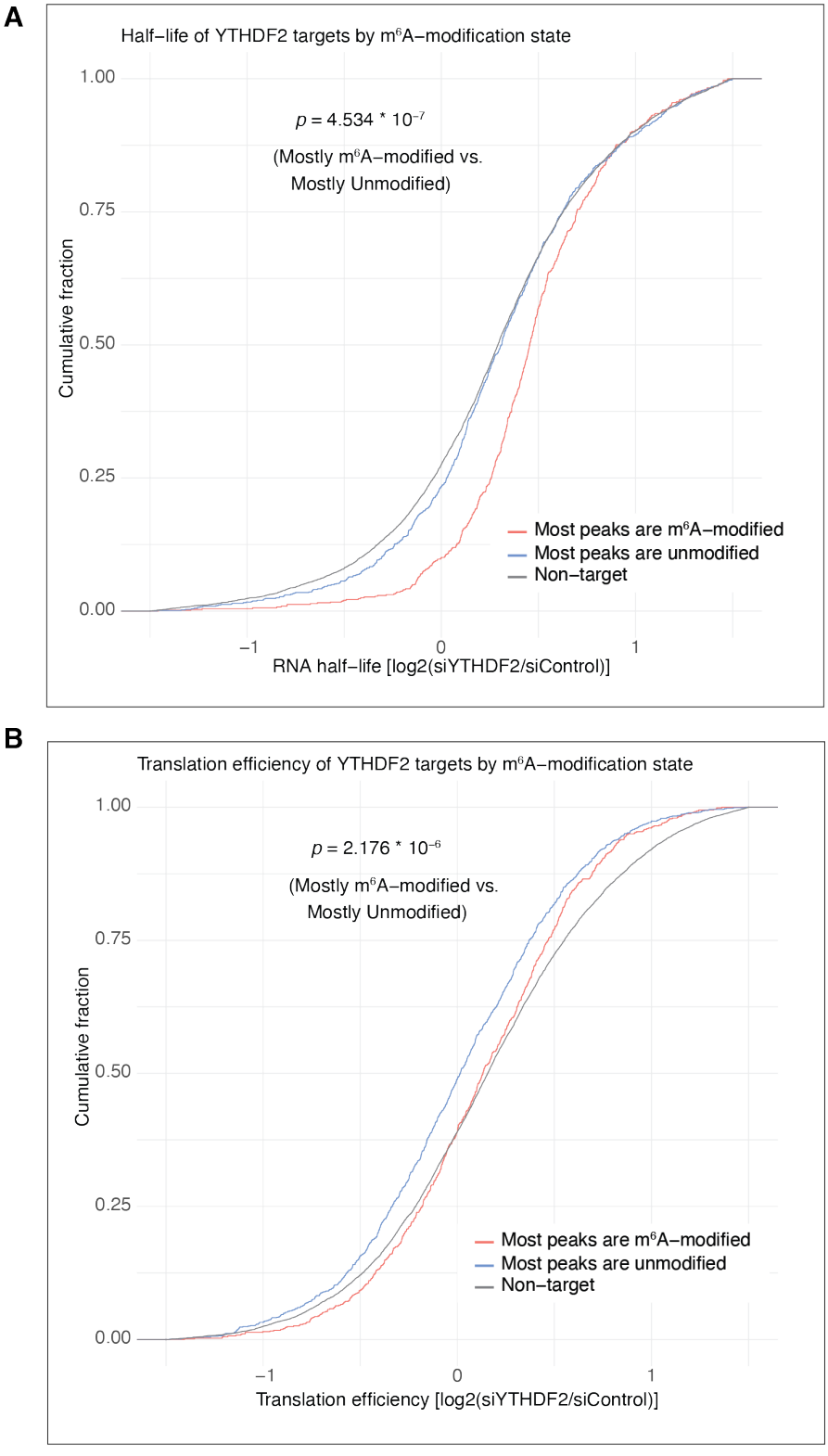


Figure 4.5 mRNA half-life and translation efficiency of YTHDF2 targets. (A-B) Cumulative distribution function plots of the change in mRNA half-life (A) or translation efficiency (B) upon knockdown of YTHDF2. Targets were categorized as described in Figure 4.4.

For translation efficiency, I observed the opposite effect as with mRNA half-life: "mostly unmodified" YTHDF2 targets had decreased translation efficiency upon YTHDF2 knockdown (median log2-change: 0.0361), while "mostly m^{6}A-modified" targets had a similar difference in translation efficiency compared to non-targets (median log2-change: 0.153 vs. 0.130) (Figure 4.5B). The difference in translation efficiency between "mostly unmodified" and "mostly m^{6}A-modified" targets was statistically significant ($p = 2.176 * 10^{-6}$). In all, these analyses show that YTHDF2 can alter either mRNA stability or mRNA translation, based on the m^{6}A status of the binding site.

Unmodified YTHDF2 targets encode distinct protein functions

I next hypothesized that the unique translation and stability effects on YTHDF2 targets based on m^{6}A status results in regulation of specific protein types. Gene Ontology (GO) analysis for Biological Process uncovered unique protein functions among YTHDF2 targets in which YTHDF2 peaks are "mostly unmodified" or "mostly m^{6}A-modified" (Figure 4.6 and Table 4.1). Specifically, "mostly unmodified" targets are enriched for proteins involved in RNA metabolism and protein folding. "Mostly m^{6}A-modified" targets are enriched for proteins involved in chromatin and histone modifications. GO analysis for Cellular Component yielded enrichment for cell adhesion structures in "mostly unmodified" targets, and ubiquitin ligase complex or histone methyltransferase complex for "mostly m^{6}A-modified" targets. Lastly, GO analysis for Molecular Function showed both unique and similar functions for both categories of YTHDF2 target, with enrichment for ubiquitin-related functions in both categories. Interestingly, "mostly unmodified" targets were enriched for transcription coactivator activity (including FUS, KMT2D, SOX4, and

components of the Mediator complex), while "mostly m^{6}A-modified" targets were enriched for transcription repressor activity (including YY1, JDP2, ETV3, and many zinc-finger and forkhead box proteins). And consistent with the results for Cellular Component, "mostly unmodified" targets were also enriched for cell adhesion-related terms. In summary, YTHDF2 targets that are "mostly unmodified" or "mostly m^{6}A-modified" encode proteins with distinct functions in the cell: "mostly unmodified" targets encode adhesion-related proteins, RNA-binding proteins, and transcriptional co-activators; "mostly m^{6}A-modified targets" encode chromatin modification enzymes and transcriptional repressors (Figure 4.6). Both categories encode ubiquitin ligases or ligase-binding proteins. Together with Figure 4.5, this data suggests that YTHDF2 participates in two modes of RNA regulation in the cell: decreasing stability of chromatin modification enzymes and transcriptional repressors or increasing translation of adhesion proteins and regulators of RNA metabolism. What is still unclear however, is how YTHDF2 is able to achieve specific binding to m^{6}A on some targets but not others.

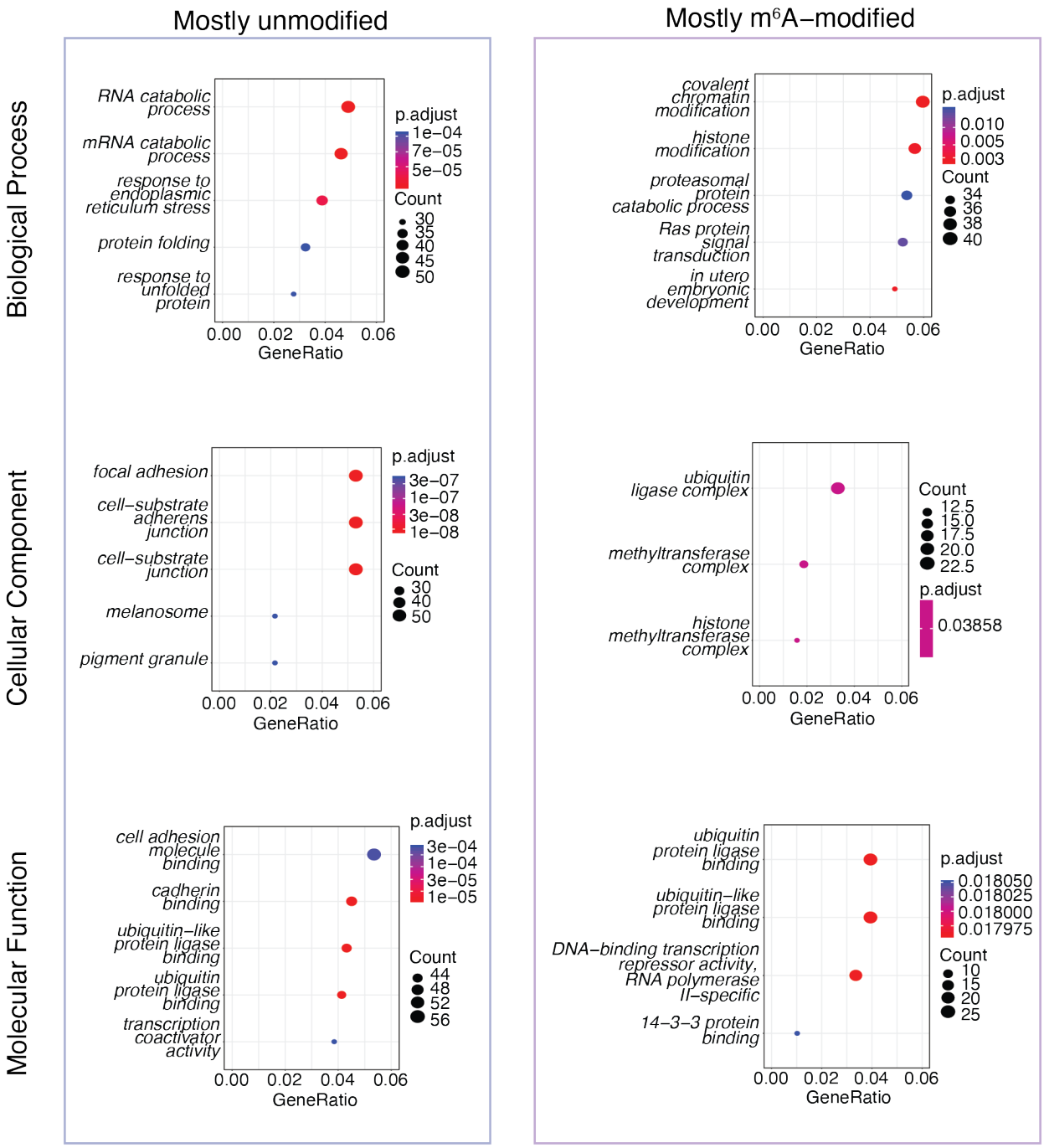


Figure 4.6 GO analysis of YTHDF2 targets by m^{6}A status. Dotplots showing enriched GO terms for Biological Process (top), Cellular Component (middle), or Molecular Function (bottom) for YTHDF2 targets categorized as in Figure 4.4. Dot size indicates number of targets captured by GO term, dot color indicates significance value. Top 5 terms by statistical significance for each category are shown.

The translation efficiency of mRNAs encoding highly structured protein domains is known to be decreased in order to promote folding fidelity (Rodnina, 2016). Thus, and given the unique proteins encoded by the two categories of YTHDF2 targets, I next hypothesized that YTHDF2 may be regulating translation of specific protein domains or secondary structures. To test this hypothesis, I utilized databases of annotated protein domains or protein secondary structures (Karolchik et al., 2004; UniProt Consortium, 2019) encoded within the human genome, and tested the overlap between the RNA sequences encoding these regions and YTHDF2 peaks that are unmodified or m^{6}A-modified. Both categories of YTHDF2 peaks predominantly overlap with zinc finger and protein kinase domains, with m^{6}A-modified peaks having a slight preference for zinc fingers compared to unmodified peaks (Figure 4.7A). Interestingly, two m^{6}A-modified peaks overlapped with the mRNA encoding the YTH-domain of the YTHDF proteins. Both categories of YTHDF2 peaks also overlap with largely the same protein secondary structures, mostly alpha helices and beta sheets (Figure 4.7B).

It is important to note that overall, the frequency of overlaps between YTHDF2 peaks with any protein domain or secondary structure is low, indicating two possibilities: first, that YTHDF2 binds mostly unstructured protein regions (which is consistent with the density of YTHDF2 binding sites at the 5' and 3' UTRs see in Figure 4.3); second, that the annotation database is incomplete. Thus, there is currently insufficient data to conclude whether YTHDF2 peaks encode similar or unique protein structures based on m^{6}A status. But given the prevalence of sites in the 3' and 5' UTRs, it is possible that the unique functions of YTHDF2 based on m^{6}A status are achieved through interactions in non-coding regions.

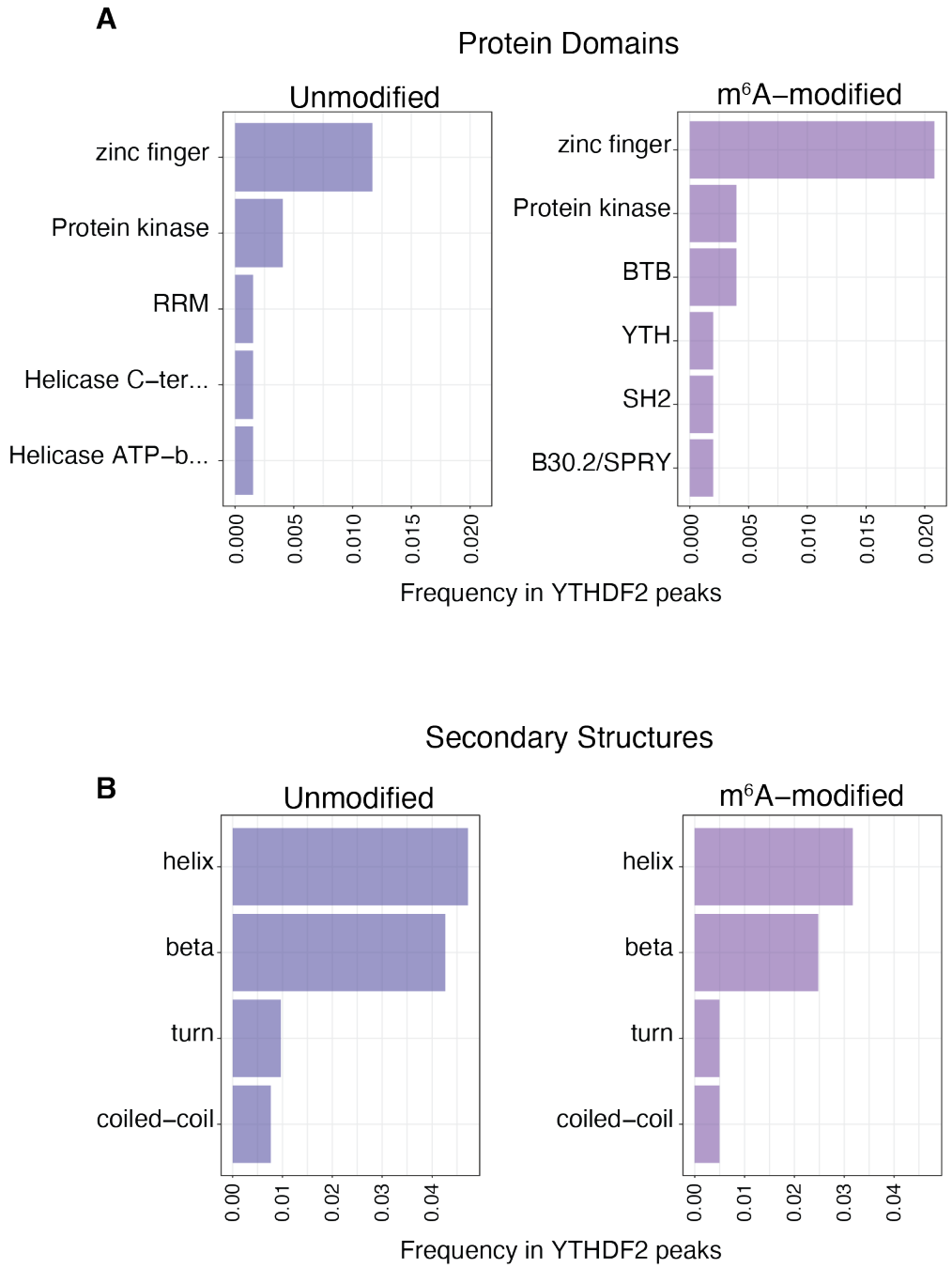


Figure 4.7 Protein domains and secondary structures encoded by YTHDF2-bound RNA. A) Bar chart showing the frequency of overlap (minimum 1 nucleotide) between YTHDF2 peaks and the mRNA encoding the indicated protein domain. Frequency indicates the number of YTHDF2 peaks overlapping that domain out of the total number of YTHDF2 peaks. B) Bar chart as in (A), but showing the frequency of overlap between YTHDF2 peaks and mRNA encoding protein secondary structures.

Conclusions

This chapter has investigated and confirmed the binding of YTHDF2 to unmodified RNAs in cells through careful examination of both statistical and biological evidence. Specifically, I demonstrated that unmodified YTHDF2 peaks are neither false-positive YTHDF2 peaks nor false-negative m^{6}A peaks. Further study of the sequence characteristics of these peaks revealed unique binding motifs. Notably, YTHDF2 peaks with m^{6}A are enriched for sequences matching the canonical m^{6}A RRACH motif, while unmodified YTHDF2 peaks are not. Unmodified YTHDF2 peaks were enriched at 3' and 5' UTRs compared to m^{6}A-modified peaks; these transcript regions are known to be important in regulating mRNA stability and translation. YTHDF2 can regulate distinct subsets of cellular transcripts encoding proteins with different functions and subcellular locations, and this subset specificity is due in part to recognition of m^{6}A. However, the mechanism by which YTHDF2 can regulate either translation or stability due to m^{6}A is likely achieved through interactions in the 3' or 5' UTR, as both unmodified and m^{6}A-modified YTHDF2 peaks showed similar preferences for RNA encoding specific protein domains or secondary structures.

The mechanism YTHDF2 uses to recognize unmodified RNAs is unclear. As discussed above, the uncharacterized N-terminal regions of YTHDF proteins may participate in binding to specific RNAs, thus explaining the discrepancy between the preference of the YTH domain for m^{6}A *in vitro* (Zhu et al., 2014), and the binding of full-length YTHDF2 to unmodified RNAs in cells. A recent report found that YTHDF3 can bind to FOXO3 during vesicular stomatitis virus (VSV) infection to suppress translation of ISGs (Yuan Zhang et al., 2019). However, this interaction was not due to m^{6}A or to the activity of METTL3 (Yuan Zhang et al., 2019). Binding of YTHDF3 to FOXO3 required both the YTH domain and the P/Q/N-rich region in the N-

terminal portion of the protein (Yuan Zhang et al., 2019). This study did not distinguish whether the contribution of the P/Q/N-region to RNA binding was direct or indirect. Therefore, it is possible that YTHDF3 bound unmodified RNA either through a novel function of its N-terminal region, or through recruitment of other protein binding partners to the N-terminal region.

Identification of YTHDF2 protein binding partners in different contexts will be critical to further understanding of the function of YTHDF2 in RNA metabolism. In particular, experiments which distinguish between YTHDF2 binding partners on unmodified and m^6A-modified RNAs will yield insight into how YTHDF2 achieves this binding and functional specificity. Such assays could include an adapted affinity purification assay using first unmodified or m^6A-modified bait, followed by a second purification using an antibody to YTHDF2 or a protein tag, then finally mass-spec. Alternatively, RAP-MS could be used to profile all proteins bound to a specific YTHDF2 target, with comparisons made between mostly unmodified and mostly m^6A-modified targets. The location and prevalence of m^6A on RNA is known to change in response to various cellular stimuli (Aguilo et al., 2015; Geula et al., 2015; Zhou et al., 2015); thus, another future experiment could examine how YTHDF2 binding changes in response to increased or decreased m^6A on target transcripts.

A major limitation of this work is that it involved few datasets. Future studies should incorporate meRIP-seq and CLIP assays from a variety of cell types to examine any consistency between the binding determinants and function of YTHDF2 on unmodified RNAs in different contexts. Comparisons between experiments performed with different CLIP assay varieties (iCLIP, PAR-CLIP, HITs-CLIP, etc.) will also reveal any biases inherent to a particular method. Technological advances in peak calling algorithms for PAR-CLIP assays will be particularly valuable. Though it is clear that the unmodified YTHDF2 peaks are bona fide peaks (Figure 4.1),

the current lack of PAR-CLIP peak callers that utilize biological replicate information limits the conclusions that can be drawn from these assays, and their reproducibility. Biochemical validation of the individual hits from the sequencing-based results will also be critical for determining the mechanism of YTHDF2 in regulating translation. For example, tethering assays used by other groups can confirm that YTHDF2 is sufficient to increase translation of reporter RNAs lacking m^6A.

This work provides important insights into a previously ignored aspect of YTHDF function. Until now, the discrepancy between YTHDF binding *in vitro* and in cells has remained unexplored. Upon secondary analysis of published datasets, it is clear that YTHDF2 not only binds unmodified RNAs, but that the presence or absence of m^6A can inform the function of YTHDF2 on these RNAs. In all, this report identifies an important role of YTHDF2 in regulating translation through binding unmodified RNA sites in mammalian cells. These findings have important implications in our understanding of both host and virus biology.

Table 4.1 Enriched GO terms for unmodified and m^{6}A-modified YTHDF2 peaks. GO terms enriched in YTHDF2 targets. Table is displayed by simplified target category: Unmodified = "mostly unmodified"; m^{6}a-modified = "mostly m^{6}A-modified". For brevity, only top 20 terms by statistical significance (adjusted p-value) are displayed.

Molecular Function			
Target Category	**GO term**	**Gene ratio**	**p.adjust**
Unmodified	ubiquitin-like protein ligase binding	46/1065	4.16E-06
Unmodified	cadherin binding	48/1065	4.16E-06
Unmodified	ubiquitin protein ligase binding	44/1065	4.16E-06
Unmodified	cell adhesion molecule binding	57/1065	5.57E-04
Unmodified	transcription coactivator activity	41/1065	6.79E-04
Unmodified	disordered domain specific binding	10/1065	0.00242685
Unmodified	unfolded protein binding	20/1065	0.01459396
Unmodified	translation factor activity, RNA binding	15/1065	0.01773122
Unmodified	translation initiation factor activity	11/1065	0.01913464
m^{6}A-modified	ubiquitin protein ligase binding	27/685	0.01795714
m^{6}A-modified	ubiquitin-like protein ligase binding	27/685	0.01795714
m^{6}A-modified	DNA-binding transcription repressor activity, RNA polymerase II-specific	23/685	0.01795714
m^{6}A-modified	14-3-3 protein binding	7/685	0.01805115
Biological Process			
Target Category	**GO term**	**Gene ratio**	**p.adjust**
Unmodified	RNA catabolic process	53/1082	3.12E-05
Unmodified	mRNA catabolic process	50/1082	3.12E-05
Unmodified	response to endoplasmic reticulum stress	42/1082	4.08E-05
Unmodified	response to unfolded protein	30/1082	1.08E-04
Unmodified	protein folding	35/1082	1.08E-04
Unmodified	regulation of mRNA metabolic process	44/1082	1.08E-04
Unmodified	regulation of mRNA catabolic process	32/1082	1.08E-04
Unmodified	regulation of mitotic cell cycle phase transition	54/1082	1.31E-04
Unmodified	regulation of chromosome organization	45/1082	1.32E-04
Unmodified	regulation of transcription involved in G1/S transition of mitotic cell cycle	11/1082	1.63E-04
Unmodified	regulation of cell cycle phase transition	56/1082	2.25E-04
Unmodified	cell cycle G2/M phase transition	37/1082	2.75E-04

Unmodified	regulation of cell cycle G2/M phase transition	32/1082	2.75E-04
Unmodified	Golgi vesicle transport	46/1082	2.75E-04
Unmodified	G2/M transition of mitotic cell cycle	35/1082	2.75E-04
Unmodified	regulation of apoptotic signaling pathway	49/1082	2.75E-04
Unmodified	interleukin-1-mediated signaling pathway	20/1082	2.75E-04
Unmodified	regulation of DNA-templated transcription in response to stress	23/1082	2.75E-04
Unmodified	regulation of G2/M transition of mitotic cell cycle	30/1082	2.75E-04
Unmodified	negative regulation of cell cycle process	45/1082	2.75E-04
m^6A-modified	covalent chromatin modification	40/670	0.00222155
m^6A-modified	histone modification	38/670	0.00261261
m^6A-modified	in utero embryonic development	33/670	0.00264745
m^6A-modified	Ras protein signal transduction	35/670	0.01544132
m^6A-modified	proteasomal protein catabolic process	36/670	0.01967553
m^6A-modified	gland development	33/670	0.02082036
m^6A-modified	Rho protein signal transduction	20/670	0.02082036
m^6A-modified	cell cycle arrest	22/670	0.02082036
m^6A-modified	mRNA 3'-end processing	13/670	0.02082036
m^6A-modified	regulation of Ras protein signal transduction	22/670	0.02082036
m^6A-modified	negative regulation of kinase activity	23/670	0.02082036
m^6A-modified	positive regulation of catabolic process	32/670	0.02082036
m^6A-modified	regulation of cyclin-dependent protein serine/threonine kinase activity	13/670	0.02082036
m^6A-modified	retinoic acid receptor signaling pathway	7/670	0.02082036
m^6A-modified	nuclear export	19/670	0.02082036
m^6A-modified	negative regulation of transferase activity	24/670	0.02449578
m^6A-modified	cell cycle checkpoint	20/670	0.02449578
m^6A-modified	regulation of cyclin-dependent protein kinase activity	13/670	0.02449578
m^6A-modified	regulation of chromosome organization	27/670	0.02449578
m^6A-modified	negative regulation of protein kinase activity	21/670	0.02449578
Cellular Component			
Target Category	**GO term**	**Gene ratio**	**p.adjust**
Unmodified	focal adhesion	59/1111	8.23E-09
Unmodified	cell-substrate adherens junction	59/1111	8.23E-09
Unmodified	cell-substrate junction	59/1111	8.23E-09
Unmodified	melanosome	24/1111	5.01E-07
Unmodified	pigment granule	24/1111	5.01E-07
Unmodified	nuclear envelope	57/1111	2.82E-06
Unmodified	endoplasmic reticulum-Golgi intermediate compartment	25/1111	2.93E-06
Unmodified	spindle	46/1111	5.00E-06

Unmodified	ficolin-1-rich granule lumen	22/1111	1.00E-04
Unmodified	endoplasmic reticulum-Golgi intermediate compartment membrane	16/1111	1.00E-04
Unmodified	spindle pole	26/1111	1.00E-04
Unmodified	microtubule	45/1111	0.00115079
Unmodified	lysosomal membrane	40/1111	0.00115079
Unmodified	lytic vacuole membrane	40/1111	0.00115079
Unmodified	nuclear membrane	35/1111	0.00122906
Unmodified	secretory granule lumen	37/1111	0.00122906
Unmodified	endosome membrane	49/1111	0.00160477
Unmodified	ficolin-1-rich granule	25/1111	0.00163225
Unmodified	host	14/1111	0.00163225
Unmodified	host cell	14/1111	0.00163225
m^6A-modified	methyltransferase complex	13/696	0.03857893
m^6A-modified	ubiquitin ligase complex	23/696	0.03857893
m^6A-modified	histone methyltransferase complex	11/696	0.03857893

CHAPTER V

DISCUSSION AND FUTURE DIRECTIONS

RNA viruses impact lives and livelihoods across the globe, from the seasonal Influenza pandemics, to the threat of emerging or re-emerging tropical diseases such as Zika, West Nile, and Dengue viruses, to the current SARS-CoV-2 catastrophe. The lack of vaccines and therapeutics to prevent or treat these diseases is as much a function of funding limitations as technological ones. In this dissertation, I have presented VIR-CLASP, a technological innovation developed with Dr. Byungil Kim and advised by Dr. Manny Ascano to identify new therapeutic targets for RNA viruses (B. Kim et al., 2020). VIR-CLASP is designed to identify host proteins that interact with viral genomes in the most vulnerable stage of infection for both host and virus. While the innate immune system is not yet primed, and viral copy numbers are low, both entities have an outsized opportunity to gain an advantage. Thus, these interactions represent a critical aspect of host and virus biology, and powerful targets for therapeutic intervention.

Hundreds of host proteins interact with pre-replicated CHIKV genomes within minutes of infection. We identified proteins already known to be important in innate immunity or CHIKV replication, including IFI16, SAMD9, PNPT1, EIF4G, and RAC1. We also identified factors with no previously defined roles during CHIKV infection, including the YTHDF family of m^6A-binding proteins. From this finding, we hypothesized and then confirmed that CHIKV genomes contain the m^6A modification. We then showed that the YTHDF proteins have distinct roles in CHIKV replication: YTHDF1 and YTHDF3 inhibit CHIKV replication and infectious particle production,

while YTHDF2 has the opposite effect. The opposing roles of YTHDFs in CHIKV replication fit into a growing picture of YTHDF function in RNA virus replication. Recent studies of HIV, HCV, IAV, and others, have shown that genetic manipulation of YTHDF1-3 can have strain-, cell-line, and stage-specific results (Courtney et al., 2017; Gokhale et al., 2016; Kennedy et al., 2017; G.-W. Kim et al., 2020; Lichinchi et al., 2016a; Tirumuru et al., 2016; Winkler et al., 2019). It is unclear how YTHDFs achieve such differences in function, but it may be due to other effector proteins that interact with the YTHDFs. Future work to investigate the binding partners of YTHDFs during viral infection in different cellular contexts will be crucial to understanding the diverse functions of these proteins in RNA metabolism.

To illustrate the versatility of VIR-CLASP we compared the CHIKV pre-replicated interactome to that of IAV. We found that many shared proteins are known regulators of viral infection, including STAU1, ZC3HAV1, LARP1, and multiple RNA helicases. FXR1, which was recently identified as a putative m^6A "reader" protein (Edupuganti et al., 2017), bound to pre-replicated CHIKV but not IAV, despite both viruses containing m^6A modifications. These experiments demonstrate one aspect of how VIR-CLASP can be used to understand RNA virus biology: through comparisons of proteins bound to different RNA viruses pre-replication. Another use of VIR-CLASP is to investigate differences in tropism of a given RNA virus between two different cell lines. In future work, VIR-CLASP could be used to understand why RNA viruses can often replicate more efficiently in certain cell types. Additionally, VIR-CLASP can be used to uncover host interactors that are unique or shared between different strains of the same virus. For example, the m^6A modification has distinct locations on three different strains of Zika virus (Figure 1.8B). Thus, profiling the interacting proteins on these genomes could inform how m^6A achieves this site-specificity, and which m^6A reader proteins may be involved. Lastly, VIR-CLASP can be

used over a time-course, as was done with CHIKV, to profile time-dependent changes in the pre-replicated interactome of different RNA viruses.

The m^{6}A RNA modification has an increasingly large list of functions in practically every aspect of cell biology (Yang et al., 2018). A key component of many studies of m^{6}A is profiling the individual and combinatorial roles of the YTHDF reader proteins. Yet discrepancies have arisen between the purported m^{6}A-binding activity of the YTHDFs *in vitro* and their validated mRNA binding sites in cells. Specifically, numerous studies have found that up to 80% of YTHDF1-3 peaks do not overlap with an m^{6}A peak (Gokhale et al., 2016; X. Wang et al., 2015a; 2014). Thus, Chapter IV of this dissertation sought to investigate this discrepancy to first determine whether technological artefacts are to blame, and second to understand whether this non-m^{6}A binding is of functional relevance in mRNA metabolism.

The possibility of false positives and false negatives in peak-calling algorithms requires careful consideration in the case of YTHDF and m^{6}A peaks: the non-overlap could simply be due to a large number of false positive YTHDF2 sites or a large number false negative m^{6}A sites. Analysis of peak scores for YTHDF2 demonstrated that the non-m^{6}A-overlapped peaks are unlikely to be false positives, as they had on average higher scores than the m^{6}A-overlapped peaks. Likewise, relaxation of the statistical stringency of the m^{6}A peaks did not rescue the overlap with YTHDF2, therefore the non-m^{6}A-overlapped YTHDF2 peaks are unlikely false negatives for m^{6}A. This leaves the possibility that the non-m^{6}A YTHDF2 peaks have a legitimate biological function. I first showed that YTHDF2 binds to distinct transcript sequences and locations based on m^{6}A status; next, YTHDF2 either alters mRNA stability or translation depending on the m^{6}A status of its binding sites; lastly, these specific activities of YTHDF2 are likely due to binding at the 3' and 5' ends of the coding sequence, as there was no specific enrichment for sequences encoding

protein domains or secondary structures. In all, these experiments demonstrate that YTHDF2 regulates different aspects of mRNA metabolism through recognition of m^{6}A- and non-m^{6}A-modified mRNAs. Future work should examine how and in which cellular contexts YTHDF2 binds non-m^{6}A-modified RNA, and the mechanisms by which YTHDF2 switches between regulating either mRNA stability or translation.

The uncharacterized N-terminal region of YTHDF2 likely contributes to its binding affinity for non-m^{6}A-modified RNAs. This is consistent with recent work showing that both the YTH and P/Q/N regions of the related protein YTHDF3 are necessary for binding to FOXO3 mRNA in an m^{6}A-independent manner (Yuan Zhang et al., 2019). To test this hypothesis, mutations to the YTH domain that are known to specifically affect binding to m^{6}A-modified RNAs but not unmodified RNAs can be used (Zhu et al., 2014). This experiment would reveal whether the N-terminal region is sufficient to bind to non-m^{6}A-modified targets, but it would be difficult to interpret if both YTH and N-terminal regions are important. It is also possible that other protein binding partners of YTHDF2 later its preference for different RNAs. Of key importance will be identifying the YTHDF2 binding partners on m^{6}A-modified and non-m^{6}A-modified RNA. To do this, one could first use an antibody to YTHDF2 to pull down protein complexes, followed by a second pull-down using bait RNA with or without m^{6}A. Alternatively, one could use known YTHD2 targets to design primer libraries for RAP-MS to compare all RBPs bound to those transcripts. Lastly, competition for m^{6}A binding sites by a growing list of m^{6}A readers (Edupuganti et al., 2017) could explain the discrepancies between YTHDF2 binding *in vitro* and in cells. By comparing CLIP assays for different m^{6}A readers, one can identify whether shared targets are more likely to be bound at m^{6}A or non-m^{6}A sites.

An important validation experiment for this work would be to knockdown or knockout the METTL3 methyltransferase to decrease total m^{6}A, then to perform RIP-qPCR to test whether YTHDF2 still interacts with the non-m^{6}A-modified targets identified in this study. This validation experiment will provide biochemical insight into whether YTHDF2 is truly binding non-m^{6}A-modified RNA in cells, or whether the results outlined in this study are due to an unknown technological bias in PAR-CLIP or m^{6}A-seq. Future validation work should also involve comparisons to other datasets. This would inform whether the biological consequences of YTHDF2 binding based on m^{6}A are also at play in other cell lines or under other cellular perturbations (such as virus infection).

In conclusion, this dissertation outlines VIR-CLASP, a novel technique to identify RBPs important in RNA virus infection. Next, the validation of hits from VIR-CLASP identified diverse roles for m^{6}A and the YTHDF proteins in CHIKV infection. These experiments illuminated a discrepancy in this work and other published research between YTHDF activity *in vitro* and in cells. Exploration of the binding of YTHDF2 to non-m^{6}A-modified RNAs in cells illuminated a previously unknown role for YTHDF2 in regulating mRNA translation. These results suggest that other YTHDF proteins or m^{6}A readers may have different binding preferences based on cellular context. Together, this work furthers our understanding of RNA virus biology and post-transcriptional gene regulation. VIR-CLASP can be used with other viruses to identify common and specific host factors that bind to RNA virus genomes. Understanding the function of RNA modifications like m^{6}A in RNA virus or host cell biology will have relevance to rational vaccine design efforts. Lastly, both m^{6}A and host RBPs constrain RNA virus evolution: understanding how these systems interact to favor certain mutations may aid in prediction or prevention of future global pandemics.

REFERENCES

Aguilo, F., Zhang, F., Sancho, A., Fidalgo, M., Di Cecilia, S., Vashisht, A., Lee, D.-F., Chen, C.-H., Rengasamy, M., Andino, B., Jahouh, F., Roman, A., Krig, S.R., Wang, R., Zhang, W., Wohlschlegel, J.A., Wang, J., Walsh, M.J., 2015. Coordination of m6A mRNA Methylation and Gene Transcription by ZFP217 Regulates Pluripotency and Reprogramming. Cell Stem Cell 17, 689–704. doi:10.1016/j.stem.2015.09.005

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol. 11, R106. doi:10.1186/gb-2010-11-10-r106

Area, E., Martín-Benito, J., Gastaminza, P., Torreira, E., Valpuesta, J.M., Carrascosa, J.L., Ortín, J., 2004. 3D structure of the influenza virus polymerase complex: localization of subunit domains. Proc Natl Acad Sci USA 101, 308–313. doi:10.1073/pnas.0307127101

Ascano, M., Hafner, M., Cekan, P., Gerstberger, S., Tuschl, T., 2012. Identification of RNA-protein interaction networks using PAR-CLIP. Wiley Interdiscip Rev RNA 3, 159–177. doi:10.1002/wrna.1103

Backlund, P.S., Carotti, D., Cantoni, G.L., 1986. Effects of the S-adenosylhomocysteine hydrolase inhibitors 3-deazaadenosine and 3-deazaaristeromycin on RNA methylation and synthesis. Eur J Biochem 160, 245–251. doi:10.1111/j.1432-1033.1986.tb09963.x

Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., Landthaler, M., 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Molecular Cell 46, 674–690. doi:10.1016/j.molcel.2012.05.021

Barrows, N.J., Anglero-Rodriguez, Y., Kim, B., Jamison, S.F., Le Sommer, C., McGee, C.E., Pearson, J.L., Dimopoulos, G., Ascano, M., Bradrick, S.S., Garcia-Blanco, M.A., 2019. Dual roles for the ER membrane protein complex in flavivirus infection: viral entry and protein biogenesis. Scientific Reports 9, 9711. doi:10.1038/s41598-019-45910-9

Beachboard, D.C., Horner, S.M., 2016. Innate immune evasion strategies of DNA and RNA viruses. Curr Opin Microbiol 32, 113–119. doi:10.1016/j.mib.2016.05.015

Bieniasz, P.D., 2004. Intrinsic immunity: a front-line defense against viral attack. Nat Immunol 5, 1109–1115. doi:10.1038/ni1125

Bray, M., Driscoll, J., Huggins, J.W., 2000. Treatment of lethal Ebola virus infection in mice with a single dose of an S-adenosyl-L-homocysteine hydrolase inhibitor. Antiviral Research 45, 135–147. doi:10.1016/s0166-3542(00)00066-8

Calabretta, S., Richard, S., 2015. Emerging Roles of Disordered Sequences in RNA-Binding Proteins. Trends in Biochemical Sciences 40, 662–672. doi:10.1016/j.tibs.2015.08.012

Carey, D.E., 1971. Chikungunya and dengue: a case of mistaken identity? J Hist Med Allied Sci 26, 243–262. doi:10.1093/jhmas/xxvi.3.243

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., Krijgsveld, J., Hentze, M.W., 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149, 1393–1406. doi:10.1016/j.cell.2012.04.031

Chen, M., Urs, M.J., Sánchez-González, I., Olayioye, M.A., Herde, M., Witte, C.-P., 2018. m6A RNA Degradation Products Are Catabolized by an Evolutionarily Conserved N6-Methyl-AMP Deaminase in Plant and Mammalian Cells. Plant Cell 30, 1511–1522. doi:10.1105/tpc.18.00236

Chow, K.T., Gale, M., Loo, Y.-M., 2018. RIG-I and Other RNA Sensors in Antiviral Immunity. Annu. Rev. Immunol. 36, 667–694. doi:10.1146/annurev-immunol-042617-053309

Conway, J.R., Lex, A., Gehlenborg, N., 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics 33, 2938–2940. doi:10.1093/bioinformatics/btx364

Corcoran, D.L., Georgiev, S., Mukherjee, N., Gottwein, E., Skalsky, R.L., Keene, J.D., Ohler, U., 2011. PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. Genome Biol. 12, R79. doi:10.1186/gb-2011-12-8-r79

Couderc, T., Chrétien, F., Schilte, C., Disson, O., Brigitte, M., Guivel-Benhassine, F., Touret, Y., Barau, G., Cayet, N., Schuffenecker, I., Desprès, P., Arenzana-Seisdedos, F., Michault, A., Albert, M.L., Lecuit, M., 2008. A mouse model for Chikungunya: young age and inefficient type-I interferon signaling are risk factors for severe disease. PLoS Pathog 4, e29. doi:10.1371/journal.ppat.0040029

Courtney, D.G., Kennedy, E.M., Dumm, R.E., Bogerd, H.P., Tsai, K., Heaton, N.S., Cullen, B.R., 2017. Epitranscriptomic Enhancement of Influenza A Virus Gene Expression and Replication. Cell Host Microbe 22, 377–386.e5. doi:10.1016/j.chom.2017.08.004

Couturier, E., Guillemin, F., Mura, M., Léon, L., Virion, J.-M., Letort, M.-J., De Valk, H., Simon, F., Vaillant, V., 2012. Impaired quality of life after chikungunya virus infection: a 2-year follow-up study. Rheumatology (Oxford) 51, 1315–1322. doi:10.1093/rheumatology/kes015

Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology 26, 1367–1372. doi:10.1038/nbt.1511

Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.-Y., Schneller, S., Zust, R., Dong, H., Thiel, V., Sen, G.C., Fensterl, V., Klimstra, W.B., Pierson, T.C., Buller, R.M., Gale, M., Shi, P.-Y., Diamond, M.S., 2010. 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature Publishing Group 468, 452–456. doi:10.1038/nature09489

Das, K., Aramini, J.M., Ma, L.-C., Krug, R.M., Arnold, E., 2010. Structures of influenza A proteins and insights into antiviral drug targets. Nat. Struct. Mol. Biol. 17, 530–538. doi:10.1038/nsmb.1779

de Clercq, E., Montgomery, J.A., 1983. Broad-spectrum antiviral activity of the carbocyclic analog of 3-deazaadenosine. Antiviral Research 3, 17–24. doi:10.1016/0166-3542(83)90011-6

de Lucas, S., Peredo, J., Marión, R.M., Sánchez, C., Ortín, J., 2010. Human Staufen1 protein interacts with influenza virus ribonucleoproteins and is required for efficient virus multiplication. J. Virol. 84, 7603–7612. doi:10.1128/JVI.00504-10

De Vlugt, C., Sikora, D., Pelchat, M., 2018. Insight into Influenza: A Virus Cap-Snatching. Viruses 10, 641. doi:10.3390/v10110641

Dhir, A., Dhir, S., Borowski, L.S., Jimenez, L., Teitell, M., Rötig, A., Crow, Y.J., Rice, G.I., Duffy, D., Tamby, C., Nojima, T., Munnich, A., Schiff, M., de Almeida, C.R., Rehwinkel, J., Dziembowski, A., Szczesny, R.J., Proudfoot, N.J., 2018. Mitochondrial double-stranded RNA triggers antiviral signalling in humans. Nature 560, 238–242. doi:10.1038/s41586-018-0363-0

Dickson, A.M., Anderson, J.R., Barnhart, M.D., Sokoloski, K.J., Oko, L., Opyrchal, M., Galanis, E., Wilusz, C.J., Morrison, T.E., Wilusz, J., 2012. Dephosphorylation of HuR protein during alphavirus infection is associated with HuR relocalization to the cytoplasm. J. Biol. Chem. 287, 36229–36238. doi:10.1074/jbc.M112.371203

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. doi:10.1093/bioinformatics/bts635

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., Sorek, R., Rechavi, G., 2012. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature Publishing Group 485, 201–206. doi:10.1038/nature11112

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., Wu, L., 2016. YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat Commun 7, 12626. doi:10.1038/ncomms12626

Durbin, A.F., Wang, C., Marcotrigiano, J., Gehrke, L., 2016. RNAs Containing Modified Nucleotides Fail To Trigger RIG-I Conformational Changes for Innate Immune Signaling. MBio 7, e00833–16. doi:10.1128/mBio.00833-16

Edupuganti, R.R., Geiger, S., Lindeboom, R.G.H., Shi, H., Hsu, P.J., Lu, Z., Wang, S.-Y., Baltissen, M.P.A., Jansen, P.W.T.C., Rossa, M., Müller, M., Stunnenberg, H.G., He, C., Carell, T., Vermeulen, M., 2017. N6-methyladenosine (m6A) recruits and repels proteins to regulate mRNA homeostasis. Nat. Struct. Mol. Biol. 1–15. doi:10.1038/nsmb.3462

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E., Finn, R.D., 2019. The Pfam protein families database in 2019. Nucleic Acids Research 47, D427–D432. doi:10.1093/nar/gky995

Fischer, A.A., Müller, K., Scholtissek, C., 1990. Specific inhibition of the synthesis of influenza virus late proteins and stimulation of early, M2, and NS2 protein synthesis by 3-deazaadenosine. Virology 177, 523–531. doi:10.1016/0042-6822(90)90517-u

Fleming, A.M., Ding, Y., Alenko, A., Burrows, C.J., 2016. Zika Virus Genomic RNA Possesses Conserved G-Quadruplexes Characteristic of the Flaviviridae Family. ACS Infect Dis 2, 674–681. doi:10.1021/acsinfecdis.6b00109

Fleming, A.M., Nguyen, N.L.B., Burrows, C.J., 2019. Colocalization of m6A and G-Quadruplex-Forming Sequences in Viral RNA (HIV, Zika, Hepatitis B, and SV40) Suggests Topological Control of Adenosine N6-Methylation. ACS Cent Sci 5, 218–228. doi:10.1021/acscentsci.8b00963

Frolova, E.I., Gorchakov, R., Pereboeva, L., Atasheva, S., Frolov, I., 2010. Functional Sindbis virus replicative complexes are formed at the plasma membrane. J. Virol. 84, 11679–11695. doi:10.1128/JVI.01441-10

Fros, J.J., Pijlman, G.P., 2016. Alphavirus Infection: Host Cell Shut-Off and Inhibition of Antiviral Responses. Viruses 8, 166. doi:10.3390/v8060166

Fu, Y., Zhuang, X., 2020. m6A-binding YTHDF proteins promote stress granule formation. Nat. Chem. Biol. 16, 955–963. doi:10.1038/s41589-020-0524-y

Garcia-Moreno, M., Noerenberg, M., Ni, S., Järvelin, A.I., González-Almela, E., Lenz, C.E., Bach-Pages, M., Cox, V., Avolio, R., Davis, T., Hester, S., Sohier, T.J.M., Li, B., Heikel, G., Michlewski, G., Sanz, M.A., Carrasco, L., Ricci, E.P., Pelechano, V., Davis, I., Fischer, B., Mohammed, S., Castello, A., 2019. System-wide Profiling of RNA-Binding Proteins Uncovers

Key Regulators of Virus Infection. Molecular Cell 74, 196–211.e11. doi:10.1016/j.molcel.2019.01.017

Gerstberger, S., Hafner, M., Tuschl, T., 2014. A census of human RNA-binding proteins. Nature Publishing Group 15, 829–845. doi:10.1038/nrg3813

Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., Ben-Haim, M.S., Eyal, E., Yunger, S., Pinto, Y., Jaitin, D.A., Viukov, S., Rais, Y., Krupalnik, V., Chomsky, E., Zerbib, M., Maza, I., Rechavi, Y., Massarwa, R., Hanna, S., Amit, I., Levanon, E.Y., Amariglio, N., Stern-Ginossar, N., Novershtern, N., Rechavi, G., Hanna, J.H., 2015. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science 347, 1002–1006. doi:10.1126/science.1261417

Gokhale, N.S., Horner, S.M., 2017. RNA modifications go viral. PLoS Pathog 13, e1006188–6. doi:10.1371/journal.ppat.1006188

Gokhale, N.S., McIntyre, A.B.R., McFadden, M.J., Roder, A.E., Kennedy, E.M., Gandara, J.A., Hopcraft, S.E., Quicke, K.M., Vazquez, C., Willer, J., Ilkayeva, O.R., Law, B.A., Holley, C.L., Garcia-Blanco, M.A., Evans, M.J., Suthar, M.S., Bradrick, S.S., Mason, C.E., Horner, S.M., 2016. N6-Methyladenosine in Flaviviridae Viral RNA Genomes Regulates Infection. Cell Host Microbe 20, 654–665. doi:10.1016/j.chom.2016.09.015

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., Ulrich, A., Wardle, G.S., Dewell, S., Zavolan, M., Tuschl, T., 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141, 129–141. doi:10.1016/j.cell.2010.03.009

Halstead, S.B., 2015. Reappearance of chikungunya, formerly called dengue, in the Americas. Emerg Infect Dis 21, 557–561. doi:10.3201/eid2104.141723

Hao, H., Hao, S., Chen, H., Chen, Z., Zhang, Y., Wang, J., Wang, H., Zhang, B., Qiu, J., Deng, F., Guan, W., 2019. N6-methyladenosine modification and METTL3 modulate enterovirus 71 replication. Nucleic Acids Research 47, 362–374. doi:10.1093/nar/gky1007

Harper, J.E., Miceli, S.M., Roberts, R.J., Manley, J.L., 1990. Sequence specificity of the human mRNA N6-adenosine methylase in vitro. Nucleic Acids Research 18, 5735–5741. doi:10.1093/nar/18.19.5735

Hawkins, T.L., O'Connor-Morin, T., Roy, A., Santillan, C., 1994. DNA purification and isolation using a solid-phase. Nucleic Acids Research 22, 4543–4544.

Heaton, N.S., Perera, R., Berger, K.L., Khadka, S., Lacount, D.J., Kuhn, R.J., Randall, G., 2010. Dengue virus nonstructural protein 3 redistributes fatty acid synthase to sites of viral

replication and increases cellular fatty acid synthesis. Proc. Natl. Acad. Sci. U.S.A. 107, 17345–17350. doi:10.1073/pnas.1010811107

Horn, S.R.G.-V., Sarnow, P., 2017. Making the Mark: The Role of Adenosine Modifications in the Life Cycle of RNA Viruses. Cell Host Microbe 21, 661–669. doi:10.1016/j.chom.2017.05.008

Horowitz, S., Horowitz, A., Nilsen, T.W., Munns, T.W., Rottman, F.M., 1984. Mapping of N6-methyladenosine residues in bovine prolactin mRNA. Proc Natl Acad Sci USA 81, 5667–5671. doi:10.1073/pnas.81.18.5667

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research 37, 1–13. doi:10.1093/nar/gkn923

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44–57. doi:10.1038/nprot.2008.211

Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B.S., Mesquita, A., Liu, C., Yuan, C.L., Hu, Y.-C., Hüttelmaier, S., Skibbe, J.R., Su, R., Deng, X., Dong, L., Sun, M., Li, C., Nachtergaele, S., Wang, Y., Hu, C., Ferchen, K., Greis, K.D., Jiang, X., Wei, M., Qu, L., Guan, J.-L., He, C., Yang, J., Chen, J., 2018. Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nature Cell Biology 1–18. doi:10.1038/s41556-018-0045-z

Huang, R., Han, M., Meng, L., Chen, X., 2018. Transcriptome-wide discovery of coding and noncoding RNA-binding proteins. Proc. Natl. Acad. Sci. U.S.A. 115, E3879–E3887. doi:10.1073/pnas.1718406115

Jan, E., Mohr, I., Walsh, D., 2016. A Cap-to-Tail Guide to mRNA Translation Strategies in Virus-Infected Cells. Annu Rev Virol 3, 283–307. doi:10.1146/annurev-virology-100114-055014

Jaubert, C., Bedrat, A., Bartolucci, L., Di Primo, C., Ventura, M., Mergny, J.-L., Amrane, S., Andreola, M.-L., 2018. RNA synthesis is modulated by G-quadruplex formation in Hepatitis C virus negative RNA strand. Scientific Reports 8, 8120. doi:10.1038/s41598-018-26582-3

Jia, G., Fu, Y., He, C., 2013. Reversible RNA adenosine methylation in biological regulation. Trends in Genetics 29, 108–115. doi:10.1016/j.tig.2012.11.003

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G., He, C., 2011. N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat. Chem. Biol. 7, 885–887. doi:10.1038/nchembio.687

Jin, X., Lian, J.-S., Hu, J.-H., Gao, J., Zheng, L., Zhang, Y.-M., Hao, S.-R., Jia, H.-Y., Cai, H., Zhang, X.-L., Yu, G.-D., Xu, K.-J., Wang, X.-Y., Gu, J.-Q., Zhang, S.-Y., Ye, C.-Y., Jin, C.-

L., Lu, Y.-F., Yu, X., Yu, X.-P., Huang, J.-R., Xu, K.-L., Ni, Q., Yu, C.-B., Zhu, B., Li, Y.-T., Liu, J., Zhao, H., Zhang, X., Yu, L., Guo, Y.-Z., Su, J.-W., Tao, J.-J., Lang, G.-J., Wu, X.-X., Wu, W.-R., Qv, T.-T., Xiang, D.-R., Yi, P., Shi, D., Chen, Y., Ren, Y., Qiu, Y.-Q., Li, L.-J., Sheng, J., Yang, Y., 2020. Epidemiological, clinical and virological characteristics of 74 cases of coronavirus-infected disease 2019 (COVID-19) with gastrointestinal symptoms. Gut 69, 1002–1009. doi:10.1136/gutjnl-2020-320926

Jurczyszak, D., Zhang, W., Terry, S.N., Kehrer, T., Bermúdez González, M.C., McGregor, E., Mulder, L.C.F., Eckwahl, M.J., Pan, T., Simon, V., 2020. HIV protease cleaves the antiviral m6A reader protein YTHDF3 in the viral particle. PLoS Pathog 16, e1008305. doi:10.1371/journal.ppat.1008305

Kane, S.E., Beemon, K., 1985. Precise localization of m6A in Rous sarcoma virus RNA reveals clustering of methylation sites: implications for RNA processing. Molecular and Cellular Biology 5, 2298–2306. doi:10.1128/MCB.5.9.2298

Karlas, A., Berre, S., Couderc, T.E.R.E.S., Varjak, M., Braun, P., Meyer, M., Gangneux, N., Karo-Astover, L., Weege, F., Raftery, M., nrich, G.U.N.S.O., Klemm, U., Wurzlbauer, A., Bracher, F., Merits, A., Meyer, T.F., Lecuit, M., 2016. A human genome-wide loss-of-function screen identifies effective chikungunya antiviral drugs. Nat Commun 7, 1–14. doi:10.1038/ncomms11320

Karlas, A., Machuy, N., Shin, Y., Pleissner, K.-P., Artarini, A., Heuer, D., Becker, D., Khalil, H., Ogilvie, L.A., Hess, S., Mäurer, A.P., Müller, E., Wolff, T., Rudel, T., Meyer, T.F., 2010. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. Nature 463, 818–822. doi:10.1038/nature08760

Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., Kent, W.J., 2004. The UCSC Table Browser data retrieval tool. Nucleic Acids Research 32, D493–6. doi:10.1093/nar/gkh103

Kennedy, E.M., Bogerd, H.P., Kornepati, A.V.R., Kang, D., Ghoshal, D., Marshall, J.B., Poling, B.C., Tsai, K., Gokhale, N.S., Horner, S.M., Cullen, B.R., 2017. Posttranscriptional m6A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. Cell Host Microbe 22, 830. doi:10.1016/j.chom.2017.11.010

Khan, A.H., Morita, K., Parquet, M.D.C., Hasebe, F., Mathenge, E.G.M., Igarashi, A., 2002. Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. Journal of General Virology 83, 3075–3084. doi:10.1099/0022-1317-83-12-3075

Kim, B., Arcos, S., Rothamel, K., Jian, J., Rose, K.L., McDonald, W.H., Bian, Y., Reasoner, S., Barrows, N.J., Bradrick, S., Garcia-Blanco, M.A., Ascano, M., 2020. Discovery of Widespread Host Protein Interactions with the Pre-replicated Genome of CHIKV Using VIR-CLASP. Molecular Cell 78, 624–640.e7. doi:10.1016/j.molcel.2020.04.013

Kim, G.-W., Imam, H., Khan, M., Siddiqui, A., 2020. N6-Methyladenosine modification of hepatitis B and C viral RNAs attenuates host innate immunity via RIG-I signaling. J. Biol. Chem. 295, 13123–13133. doi:10.1074/jbc.RA120.014260

Krug, R.M., Morgan, M.A., Shatkin, A.J., 1976. Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. J. Virol. 20, 45–53.

Kummer, S., Flöttmann, M., Schwanhäusser, B., Sieben, C., Veit, M., Selbach, M., Klipp, E., Herrmann, A., 2014. Alteration of protein levels during influenza virus H1N1 infection in host cells: a proteomic survey of host and virus reveals differential dynamics. PLoS ONE 9, e94257. doi:10.1371/journal.pone.0094257

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25. doi:10.1186/gb-2009-10-3-r25

Lenarcic, E.M., Landry, D.M., Greco, T.M., Cristea, I.M., Thompson, S.R., 2013. Thiouracil cross-linking mass spectrometry: a cell-based method to identify host factors involved in viral amplification. J. Virol. 87, 8697–8712. doi:10.1128/JVI.00950-13

Li, M.M.H., Lau, Z., Cheung, P., Aguilar, E.G., Schneider, W.M., Bozzacco, L., Molina, H., Buehler, E., Takaoka, A., Rice, C.M., Felsenfeld, D.P., MacDonald, M.R., 2017. TRIM25 Enhances the Antiviral Action of Zinc-Finger Antiviral Protein (ZAP). PLoS Pathog 13, e1006145. doi:10.1371/journal.ppat.1006145

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., Darnell, J.C., Darnell, R.B., 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456, 464–469. doi:10.1038/nature07488

Lichinchi, G., Gao, S., Saletore, Y., Gonzalez, G.M., Bansal, V., Wang, Y., Mason, C.E., Rana, T.M., 2016a. Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. Nature Microbiology 1, 16011. doi:10.1038/nmicrobiol.2016.11

Lichinchi, G., Zhao, B.S., Wu, Y., Lu, Z., Qin, Y., He, C., Rana, T.M., 2016b. Dynamics of Human and Viral RNA Methylation during Zika Virus Infection. Cell Host Microbe 20, 666–673. doi:10.1016/j.chom.2016.10.002

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., Dai, Q., Chen, W., He, C., 2014. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat. Chem. Biol. 10, 93–95. doi:10.1038/nchembio.1432

Liu, J., McFadden, G., 2015. SAMD9 is an innate antiviral host factor with stress response properties that can be antagonized by poxviruses. J. Virol. 89, 1925–1931.

doi:10.1128/JVI.02262-14

Liu, Y., You, Y., Lu, Z., Yang, J., Li, P., Liu, L., Xu, H., Niu, Y., Cao, X., 2019. N6-methyladenosine RNA modification-mediated cellular metabolism rewiring inhibits viral replication. Science 365, 1171–1176. doi:10.1126/science.aax4468

Lu, M., Zhang, Z., Xue, M., Zhao, B.S., Harder, O., Li, A., Liang, X., Gao, T.Z., Xu, Y., Zhou, J., Feng, Z., Niewiesk, S., Peeples, M.E., He, C., Li, J., 2020. N6-methyladenosine modification enables viral RNA to escape recognition by RNA sensor RIG-I. Nature Microbiology 5, 584–598. doi:10.1038/s41564-019-0653-9

Lu, W., Tirumuru, N., St Gelais, C., Koneru, P.C., Liu, C., Kvaratskhelia, M., He, C., Wu, L., 2018. N6-methyladenosine-binding proteins suppress HIV-1 infectivity and viral production. J. Biol. Chem. doi:10.1074/jbc.RA118.004215

Lumsden, W.H., 1955. An epidemic of virus disease in Southern Province, Tanganyika Territory, in 1952-53. II. General description and epidemiology. Trans R Soc Trop Med Hyg 49, 33–57. doi:10.1016/0035-9203(55)90081-x

Luo, S., Tong, L., 2014. Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. Proc. Natl. Acad. Sci. U.S.A. 111, 13834–13839. doi:10.1073/pnas.1412742111

Ma, Z., Ni, G., Damania, B., 2018. Innate Sensing of DNA Virus Genomes. Annu Rev Virol 5, 341–362. doi:10.1146/annurev-virology-092917-043244

Meyer, K.D., 2019. DART-seq: an antibody-free method for global m6A detection. Nat Meth 16, 1275–1280. doi:10.1038/s41592-019-0570-0

Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.-B., Jaffrey, S.R., 2015. 5′ UTR m6A Promotes Cap-Independent Translation. Cell 163, 999–1010. doi:10.1016/j.cell.2015.10.012

Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., Jaffrey, S.R., 2012. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149, 1635–1646. doi:10.1016/j.cell.2012.05.003

Molinie, B., Wang, J., Lim, K.-S., Hillebrand, R., Lu, Z.-X., Van Wittenberghe, N., Howard, B.D., Daneshvar, K., Mullen, A.C., Dedon, P., Xing, Y., Giallourakis, C.C., 2016. m6A-LAIC-seq reveals the census and complexity of the m6A epitranscriptome. Nat Meth 13, 692–698. doi:10.1038/nmeth.3898

Montgomery, J.A., Clayton, S.J., Thomas, H.J., Shannon, W.M., Arnett, G., Bodner, A.J., Kion, I.K., Cantoni, G.L., Chiang, P.K., 1982. Carbocyclic analogue of 3-deazaadenosine: a novel

antiviral agent using S-adenosylhomocysteine hydrolase as a pharmacological target. J Med Chem 25, 626–629. doi:10.1021/jm00348a004

Morazzani, E.M., Compton, J.R., Leary, D.H., Berry, A.V., Hu, X., Marugan, J.J., Glass, P.J., Legler, P.M., 2019. Proteolytic cleavage of host proteins by the Group IV viral proteases of Venezuelan equine encephalitis virus and Zika virus. Antiviral Research 164, 106–122. doi:10.1016/j.antiviral.2019.02.001

Murphy, K., Travers, P., Walport, M., Janeway, C., 2017. The Induced Responses of Innate Immunity, in: Janeways Immunobiology. New York.

Nan, Y., Nan, G., Zhang, Y.-J., 2014. Interferon induction by RNA viruses and antagonism by viral pathogens. Viruses 6, 4999–5027. doi:10.3390/v6124999

Noda, T., Murakami, S., Nakatsu, S., Imai, H., Muramoto, Y., Shindo, K., Sagara, H., Kawaoka, Y., 2018. Importance of the 1+7 configuration of ribonucleoprotein complexes for influenza A virus genome packaging. Nat Commun 9, 54. doi:10.1038/s41467-017-02517-w

Ooi, Y.S., Stiles, K.M., Liu, C.Y., Taylor, G.M., Kielian, M., 2013. Genome-wide RNAi screen identifies novel host proteins required for alphavirus entry. PLoS Pathog 9, e1003835. doi:10.1371/journal.ppat.1003835

Patil, D.P., Chen, C.-K., Pickering, B.F., Chow, A., Jackson, C., Guttman, M., Jaffrey, S.R., 2016. m(6)A RNA methylation promotes XIST-mediated transcriptional repression. Nature 537, 369–373. doi:10.1038/nature19342

Pendleton, K.E., Chen, B., Liu, K., Hunter, O.V., Xie, Y., Tu, B.P., Conrad, N.K., 2017. The U6 snRNA m6A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. Cell 169, 824–829.e14. doi:10.1016/j.cell.2017.05.003

Perez-Perri, J.I., Rogell, B., Schwarzl, T., Stein, F., Zhou, Y., Rettel, M., Brosig, A., Hentze, M.W., 2018. Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. Nat Commun 9, 4408–13. doi:10.1038/s41467-018-06557-8

Phillips, S.L., Soderblom, E.J., Bradrick, S.S., Garcia-Blanco, M.A., 2016. Identification of Proteins Bound to Dengue Viral RNA In Vivo Reveals New Host Proteins Important for Virus Replication. MBio 7, e01865–15. doi:10.1128/mBio.01865-15

Ping, X.-L., Sun, B.-F., Wang, L., Xiao, W., Yang, X., Wang, W.-J., Adhikari, S., Shi, Y., Lv, Y., Chen, Y.-S., Zhao, X., Li, A., Yang, Y., Dahal, U., Lou, X.-M., Liu, X., Huang, J., Yuan, W.-P., Zhu, X.-F., Cheng, T., Zhao, Y.-L., Wang, X., Rendtlew Danielsen, J.M., Liu, F., Yang, Y.-G., 2014. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. 24, 177–189. doi:10.1038/cr.2014.3

Plaskon, N.E., Adelman, Z.N., Myles, K.M., 2009. Accurate strand-specific quantification of viral RNA. PLoS ONE 4, e7468. doi:10.1371/journal.pone.0007468

Radoshitzky, S.R., Pegoraro, G., Chī, X., Dǒng, L., Chiang, C.-Y., Jozwick, L., Clester, J.C., Cooper, C.L., Courier, D., Langan, D.P., Underwood, K., Kuehl, K.A., Sun, M.G., Caì, Y., Yú, S., Burk, R., Zamani, R., Kota, K., Kuhn, J.H., Bavari, S., 2016. siRNA Screen Identifies Trafficking Host Factors that Modulate Alphavirus Infection. PLoS Pathog 12, e1005466–30. doi:10.1371/journal.ppat.1005466

Renault, P., Josseran, L., Pierre, V., 2008. Chikungunya-related fatality rates, Mauritius, India, and Reunion Island. Emerg Infect Dis 14, 1327–1327. doi:10.3201/eid1408.080201

Ries, R.J., Zaccara, S., Klein, P., Olarerin-George, A., Namkoong, S., Pickering, B.F., Patil, D.P., Kwak, H., Lee, J.H., Jaffrey, S.R., 2019. m6A enhances the phase separation potential of mRNA. Nature 571, 424–428. doi:10.1038/s41586-019-1374-1

Ríos-Marco, P., Romero-López, C., Berzal-Herranz, A., 2016. The cis-acting replication element of the Hepatitis C virus genome recruits host factors that influence viral replication and translation. Scientific Reports 6, 25729. doi:10.1038/srep25729

Robinson, M.C., 1955. An epidemic of virus disease in Southern Province, Tanganyika Territory, in 1952-53. I. Clinical features. Trans R Soc Trop Med Hyg 49, 28–32. doi:10.1016/0035-9203(55)90080-8

Rodnina, M.V., 2016. The ribosome in action: Tuning of translational efficiency and protein folding. Protein Sci 25, 1390–1406. doi:10.1002/pro.2950

Rusinova, I., Forster, S., Yu, S., Kannan, A., Masse, M., Cumming, H., Chapman, R., Hertzog, P.J., 2013. Interferome v2.0: an updated database of annotated interferon-regulated genes. Nucleic Acids Research 41, D1040–6. doi:10.1093/nar/gks1215

Ryan, C., Ivanova, L., Schlesinger, M.J., 1998. Effects of site-directed mutations of transmembrane cysteines in sindbis virus E1 and E2 glycoproteins on palmitylation and virus replication. Virology 249, 62–67. doi:10.1006/viro.1998.9281

Samji, T., 2009. Influenza A: understanding the viral life cycle. Yale J Biol Med 82, 153–159.

Sandri-Goldin, R.M., 2004. Viral regulation of mRNA export. J. Virol. 78, 4389–4396. doi:10.1128/jvi.78.9.4389-4396.2004

Schibler, U., Kelley, D.E., Perry, R.P., 1977. Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. J Mol Biol 115, 695–714. doi:10.1016/0022-2836(77)90110-3

Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., Rice, C.M., 2011. A diverse range of gene products are effectors of the type I interferon antiviral response. Nature 472, 481–485. doi:10.1038/nature09907

Sheehy, A.M., Gaddis, N.C., Choi, J.D., Malim, M.H., 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature 418, 646–650. doi:10.1038/nature00939

Silva, L.A., Dermody, T.S., 2017. Chikungunya virus: epidemiology, replication, disease mechanisms, and prospective intervention strategies. Journal of Clinical Investigation 127, 737–749. doi:10.1172/JCI84417

Smyth, G.K., 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3, Article3. doi:10.2202/1544-6115.1027

Solignat, M., Gay, B., Higgs, S., Briant, L., Devaux, C., 2009. Replication cycle of chikungunya: A re-emerging arbovirus. Virology 393, 183–197. doi:10.1016/j.virol.2009.07.024

Steinhauer, D.A., Holland, J.J., 1987. Rapid evolution of RNA viruses. Annu. Rev. Microbiol. 41, 409–433. doi:10.1146/annurev.mi.41.100187.002205

Strauss, J.H., Strauss, E.G., 1994. The alphaviruses: gene expression, replication, and evolution. Microbiol Rev 58, 491–562.

Sugimoto, Y., König, J., Hussain, S., Zupan, B., Curk, T., Frye, M., Ule, J., 2012. Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. Genome Biol. 13, R67. doi:10.1186/gb-2012-13-8-r67

Suzuki, Y., Chin, W.-X., Han, Q., Ichiyama, K., Lee, C.H., Eyo, Z.W., Ebina, H., Takahashi, H., Takahashi, C., Tan, B.H., Hishiki, T., Ohba, K., Matsuyama, T., Koyanagi, Y., Tan, Y.-J., Sawasaki, T., Chu, J.J.H., Vasudevan, S.G., Sano, K., Yamamoto, N., 2016. Characterization of RyDEN (C19orf66) as an Interferon-Stimulated Cellular Inhibitor against Dengue Virus Replication. PLoS Pathog 12, e1005357. doi:10.1371/journal.ppat.1005357

Sysoev, V.O., Fischer, B., Frese, C.K., Gupta, I., Krijgsveld, J., Hentze, M.W., Castello, A., Ephrussi, A., 2016. Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in Drosophila. Nat Commun 7, 12128–11. doi:10.1038/ncomms12128

Thompson, M.R., Sharma, S., Atianand, M., Jensen, S.B., Carpenter, S., Knipe, D.M., Fitzgerald, K.A., Kurt-Jones, E.A., 2014. Interferon γ-inducible protein (IFI) 16 transcriptionally regulates type i interferons and other interferon-stimulated genes and controls the interferon response to both DNA and RNA viruses. J. Biol. Chem. 289, 23568–23581. doi:10.1074/jbc.M114.554147

Tirumuru, N., Wu, L., 2019. HIV-1 envelope proteins up-regulate N6-methyladenosine levels of cellular RNA independently of viral replication. J. Biol. Chem. 294, 3249–3260. doi:10.1074/jbc.RA118.005608

Tirumuru, N., Zhao, B.S., Lu, W., Lu, Z., He, C., Wu, L., 2016. N(6)-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. Elife 5, 165. doi:10.7554/eLife.15528

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., Pachter, L., 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nature Biotechnology 31, 46–53. doi:10.1038/nbt.2450

UniProt Consortium, 2019. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Research 47, D506–D515. doi:10.1093/nar/gky1049

Uyar, B., Yusuf, D., Wurmus, R., Rajewsky, N., Ohler, U., Akalin, A., 2017. RCAS: an RNA centric annotation system for transcriptome-wide regions of interest. Nucleic Acids Research 45, e91–e91. doi:10.1093/nar/gkx120

van Duijl-Richter, M.K.S., Hoornweg, T.E., Rodenhuis-Zybert, I.A., Smit, J.M., 2015. Early Events in Chikungunya Virus Infection-From Virus Cell Binding to Membrane Fusion. Viruses 7, 3647–3674. doi:10.3390/v7072792

Volk, S.M., Chen, R., Tsetsarkin, K.A., Adams, A.P., Garcia, T.I., Sall, A.A., Nasar, F., Schuh, A.J., Holmes, E.C., Higgs, S., Maharaj, P.D., Brault, A.C., Weaver, S.C., 2010. Genome-scale phylogenetic analyses of chikungunya virus reveal independent emergences of recent epidemics and various evolutionary rates. J. Virol. 84, 6497–6504. doi:10.1128/JVI.01603-09

Walsh, D., Mathews, M.B., Mohr, I., 2013. Tinkering with translation: protein synthesis in virus-infected cells. Cold Spring Harb Perspect Biol 5, a012351–a012351. doi:10.1101/cshperspect.a012351

Wang, S.-R., Min, Y.-Q., Wang, J.-Q., Liu, C.-X., Fu, B.-S., Wu, F., Wu, L.-Y., Qiao, Z.-X., Song, Y.-Y., Xu, G.-H., Wu, Z.-G., Huang, G., Peng, N.-F., Huang, R., Mao, W.-X., Peng, S., Chen, Y.-Q., Zhu, Y., Tian, T., Zhang, X.-L., Zhou, X., 2016. A highly conserved G-rich consensus sequence in hepatitis C virus core gene represents a new anti-hepatitis C target. Sci Adv 2, e1501535. doi:10.1126/sciadv.1501535

Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., Ren, B., Pan, T., He, C., 2014. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505, 117–120. doi:10.1038/nature12730

Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., He, C., 2015a. N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell 161, 1388–1399. doi:10.1016/j.cell.2015.05.014

Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., He, C., 2015b. N6-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell 161, 1388–1399. doi:10.1016/j.cell.2015.05.014

Weaver, S.C., Lecuit, M., 2015. Chikungunya virus and the global spread of a mosquito-borne disease. N. Engl. J. Med. 372, 1231–1239. doi:10.1056/NEJMra1406035

Wei, C.-M., Gershowitz, A., Moss, B., 1975. Methylated Nucleotides Block 5' Terminus of HeLa Cell Messenger RNA. Cell 4, 379–386.

White, L.K., Sali, T., Alvarado, D., Gatti, E., Pierre, P., Streblow, D., Defilippis, V.R., 2011. Chikungunya virus induces IPS-1-dependent innate immune activation and protein kinase R-independent translational shutoff. J. Virol. 85, 606–620. doi:10.1128/JVI.00767-10

Wickham, H., 2016. ggplot2. Springer, Cham.

Winkler, R., Gillis, E., Lasman, L., Safra, M., Geula, S., Soyris, C., Nachshon, A., Tai-Schmiedel, J., Friedman, N., Le-Trilling, V.T.K., Trilling, M., Mandelboim, M., Hanna, J.H., Schwartz, S., Stern-Ginossar, N., 2019. m6A modification controls the innate immune response to infection by targeting type I interferons. Nature Publishing Group 20, 173–182. doi:10.1038/s41590-018-0275-z

Wyde, P.R., Ambrose, M.W., Meyer, H.L., Zolinski, C.L., Gilbert, B.E., 1990. Evaluation of the toxicity and antiviral activity of carbocyclic 3-deazaadenosine against respiratory syncytial and parainfluenza type 3 viruses in tissue culture and in cotton rats. Antiviral Research 14, 215–225. doi:10.1016/0166-3542(90)90003-p

Xiao, W., Adhikari, S., Dahal, U., Chen, Y.-S., Hao, Y.-J., Sun, B.-F., Sun, H.-Y., Li, A., Ping, X.-L., Lai, W.-Y., Wang, X., Ma, H.-L., Huang, C.-M., Yang, Y., Huang, N., Jiang, G.-B., Wang, H.-L., Zhou, Q., Wang, X.-J., Zhao, Y.-L., Yang, Y.-G., 2016. Nuclear m6A Reader YTHDC1 Regulates mRNA Splicing. Molecular Cell 61, 507–519. doi:10.1016/j.molcel.2016.01.012

Xue, M., Zhao, B.S., Zhang, Z., Lu, M., Harder, O., Chen, P., Lu, Z., Li, A., Ma, Y., Xu, Y., Liang, X., Zhou, J., Niewiesk, S., Peeples, M.E., He, C., Li, J., 2019. Viral N6-methyladenosine upregulates replication and pathogenesis of human respiratory syncytial virus. Nat Commun 10, 4595. doi:10.1038/s41467-019-12504-y

Yang, Y., Hsu, P.J., Chen, Y.-S., Yang, Y.-G., 2018. Dynamic transcriptomic m6A decoration: writers, erasers, readers and functions in RNA metabolism. Cell Res. 28, 616–624. doi:10.1038/s41422-018-0040-8

Yao, M., Dong, Y., Wang, Y., Liu, H., Ma, H., Zhang, H., Zhang, L., Cheng, L., Lv, X., Xu, Z., Zhang, F., Lei, Y., Ye, W., 2020. N6-methyladenosine modifications enhance enterovirus 71

ORF translation through METTL3 cytoplasmic distribution. Biochem. Biophys. Res. Commun. 527, 297–304. doi:10.1016/j.bbrc.2020.04.088

Yoo, S., Dynan, W.S., 1998. Characterization of the RNA binding properties of Ku protein. Biochemistry 37, 1336–1343. doi:10.1021/bi972100w

Yu, G., Wang, L.-G., Han, Y., He, Q.-Y., 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16, 284–287. doi:10.1089/omi.2011.0118

Yu, G., Wang, L.-G., Yan, G.-R., He, Q.-Y., 2015. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. Bioinformatics 31, 608–609. doi:10.1093/bioinformatics/btu684

Yue, Y., Liu, J., He, C., 2015. RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. Genes Dev. 29, 1343–1355. doi:10.1101/gad.262766.115

Zhang, Na, Zhao, H., Zhang, L., 2019. Fatty Acid Synthase Promotes the Palmitoylation of Chikungunya Virus nsP1. J. Virol. 93, 662. doi:10.1128/JVI.01747-18

Zhang, Yong, Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S., 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137. doi:10.1186/gb-2008-9-9-r137

Zhang, Yuan, Wang, X., Zhang, X., Wang, J., Ma, Y., Zhang, L., Cao, X., 2019. RNA-binding protein YTHDF3 suppresses interferon-dependent antiviral responses by promoting FOXO3 translation. Proc. Natl. Acad. Sci. U.S.A. 116, 976–981. doi:10.1073/pnas.1812536116

Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C.J., Vågbø, C.B., Shi, Y., Wang, W.-L., Song, S.-H., Lu, Z., Bosmans, R.P.G., Dai, Q., Hao, Y.-J., Yang, X., Zhao, W.-M., Tong, W.-M., Wang, X.-J., Bogdan, F., Furu, K., Fu, Y., Jia, G., Zhao, X., Liu, J., Krokan, H.E., Klungland, A., Yang, Y.-G., He, C., 2013. ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. Molecular Cell 49, 18–29. doi:10.1016/j.molcel.2012.10.015

Zheng, Q., Hou, J., Zhou, Y., Li, Z., Cao, X., 2017. The RNA helicase DDX46 inhibits innate immunity by entrapping m6A-demethylated antiviral transcripts in the nucleus. Nature Publishing Group 18, 1–14. doi:10.1038/ni.3830

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., Qian, S.-B., 2015. Dynamic m6A mRNA methylation directs translational control of heat shock response. Nature 526, 591–594. doi:10.1038/nature15377

Zhu, T., Roundtree, I.A., Wang, P., Wang, X., Wang, L., Sun, C., Tian, Y., Li, J., He, C., Xu, Y., 2014. Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of N6-methyladenosine. Cell Res. 24, 1493–1496. doi:10.1038/cr.2014.152