Emerging RNA-centric technologies to probe RNA-protein interactions: importance in decoding the life cycle of positive sense single strand RNA viruses and antiviral discovery

Sreemoyee Ghosh[1†], Shivam Kumar[2†], Rohit Verma[1], Shabnam Ansari[1], Samrat Chatterjee[2] and Milan Surjit[1*]

[1]Virology Laboratory, Centre for Virus Research, Therapeutics and Vaccines, Translational Health Science and Technology Institute, NCR Biotech Science Cluster, Faridabad, Haryana, India, [2]Complex Analysis Group, Computational and Mathematical Biology Centre, Translational Health Science and Technology Institute, NCR Biotech Science Cluster, Faridabad, Haryana, India

Positive sense single strand RNA (+ssRNA) viruses are one of the evolutionary successful organisms and many of them pose a significant threat to human health. Diseases caused by +ssRNA viruses such as COVID-19, Flu and acute viral hepatitis are major public health concern worldwide. Therefore, a lot of research is focused at decoding the life cycle of +ssRNA viruses and develop specific antiviral therapeutics against them. Interaction of the viral RNA with virus-encoded proteins and host proteins drives the lifecycle and pathogenesis of +ssRNA viruses. Recent developments in computational and high-throughput omics-based experimental technologies offer the sensitivity and specificity for molecular characterization of these RNA-protein complexes. These are promising tools to revolutionize the field of +ssRNA virus research and pave the way for antiviral discovery. This review summarizes the current scientific resources available to characterize the RNA-protein interactome of +ssRNA viruses and provides an overview of the drug discovery pipeline for developing antivirals against pathogenic +ssRNA viruses.

KEYWORDS

RNA-protein interactions, RNA binding protein, positive strand RNA viruses, RaPID assay, RAP-MS

1 Introduction

The central dogma of molecular biology signifies the importance of flow of genetic information from DNA to RNA to protein. Decades of research have further uncovered multiple layers of complex mechanisms by which biological systems accurately process the flow of information and maintain homeostasis. Such precision and specificity of the

biological systems are mostly attributed to close interaction between different components of the system.

RNA and proteins are two fundamental components of living organisms, required for their survival and propagation. Ribosomal-RNA, transfer-RNA and messenger-RNA work in a coordinated fashion to generate proteins, which perform major cellular function to maintain homeostasis. Interaction between RNA and proteins (RNA-binding proteins, denoted as "RBPs" hereafter) plays a major role in mediating the function of both and such interactions are indispensable for many essential processes in living organisms. RBPs serve diverse cellular functions: for example, RBP-RNA interacts to form the ribonucleoprotein particles (RNPs), dynamic complexes, involved in different steps of gene expression, intracellular trafficking of RNA, decay of RNA and control of protein turnover etc (Dreyfuss et al., 2002; Gerstberger et al., 2014). The RBPs function by synergistically interacting with structurally well-defined binding domains. Although these domains are limited in number, they are tailored to perform specific function (Lunde et al., 2007). The major RBP binding domains with over 100 PDB structures are Zinc Finger, Helicase, RNA Recognition Motif, PUA domain, and KH domain (Corley et al., 2020). RBPs are consistent with their frequent housekeeping roles, widely distributed across tissues, and more evolutionarily conserved than standard regulators like transcription factors (Gerstberger et al., 2014).

In addition to endogenous cellular regulations, RBPs play a pivotal role in determining the fate of pathogens, such as viruses, within our bodies. Pathogenic +ssRNA viruses are a major human health concern. Owing to simple organization and high mutation rate of their genome, they generate a number of distinct variants in a short span of time, making it more difficult to control their spread. Notably, the central dogma of flow of genetic information in +ssRNA viruses rely only on two components, that is, from RNA to protein. Viral RNA serves as the genetic material and with the help of virus-encoded proteins and host proteins, it plays a central role in transmission, spread and maintenance of genomic integrity of the virus. Knowledge gained from research on many +ssRNA viruses suggest that specific and spatio-temporally controlled RNA-protein interactions among viral RNA and proteins as well as viral RNA/proteins and host RNA/proteins enable these viruses to hijack the host cellular machineries in order to survive and proliferate inside the host and maintain their genomic integrity through generations (Nagy and Pogany, 2012; Robinson et al., 2018). Therefore, molecular dissection of these RNA-protein interactions is key to understanding the mechanistic details of survival and spread of the pathogenic +ssRNA viruses as well as designing specific antiviral therapeutics against them.

Due to the unstable nature and crucial role of secondary and tertiary structures of RNA in dictating its function, it is not easy to characterize RNA-protein interactions. However, with the development of more sensitive proteomics techniques and computational methods, it is now possible to construct the RNA-protein interactome of +ssRNA viruses. RNA-protein interactome of few +ssRNA viruses such as SARS-CoV-2 and Zika virus have been generated, which helped in understanding the life cycle of the virus and identification of putative antiviral targets

(Flynn et al., 2021; Kamel et al., 2021a; Schmidt et al., 2021; Verma et al., 2021; Zhang et al., 2022). In this review, we focus on available RNA-centric techniques to construct the RNA-protein interactome and discuss the functional significance of the data in understanding the life cycle of +ssRNA viruses and antiviral target discovery.

2 RNA-protein interactions help the +ssRNA viruses escape the host innate immune response and complete their life cycle

2.1 RNA-protein interactions in viral evasion of the host innate immune response

Host innate immune effectors differentiate between self and non-self RNAs. After entry of an +ssRNA virus into the host cell, viral RNA is released from the capsid, which may be recognized by the host antiviral immune effectors such as Toll like receptor 7/8 (TLR7/8), 2′-5′-oligoadenylate synthetase (OAS)/RNase L and targeted for degradation (Figure 1) (Chan and Gack, 2016). Further, during replication of the viral genome, double-strand RNA is generated, which is recognized by host antiviral immune effectors such as Toll-like receptors (TLRs) and RIG-I-like receptors (RLRs). Among the TLRs, TLR3 and RLR family proteins like, retinoic acid inducible gene-I (RIG-I), melanoma differentiation-associated gene 5 (MDA5) and laboratory of genetics and physiology 2 (LGP2) channelize the viral RNA for degradation (Figure 1) (Ma and Suthar, 2015; Chan and Gack, 2016). In many cases, +ssRNA virus infection also causes mitochondrial damage, resulting in the release of mitochondrial DNA, which is sensed by the cyclic GMP-AMP (cGMP) synthase (cGAS), leading to the induction of type I interferon and interferon-stimulated genes (ISGs), thereby mounting a strong antiviral response. RNA viruses also activate NOD-like receptor thermal protein domain associated protein 3 (NLRP3), activating inflammasomes and/or pyroptosis (Choudhury et al., 2021). Viral RNA may also modulate cellular autophagy machinery and components of the stress granule, RNA granule or P bodies due to their link with the host's innate immune response (White and Lloyd, 2012; Tsai and Lloyd, 2014). Interaction between the viral RNA and host proteins mediate the above-mentioned processes. For example, RNA-protein interactome of the SARS-CoV-2 5'- and 3'-UTR regions identified DDX24 and ABCE1 as interaction partners of the viral 3'-UTR and 5'-UTR, respectively (Verma et al., 2021). DDX24 associates with RNA and negatively regulates RIG-I-like receptor signaling, inhibiting the host antiviral response (Ma et al., 2013). ABCE1 (RNase L inhibitor) inhibits the activity of RNase L, which is activated by the host in response to RNA virus infection or interferon alpha/beta (IFN-α/β) stimulation (Tian et al., 2012). Active RNase L cleaves the viral RNA, which is prevented in the presence of ABCE1. Hence, DDX24-3'-UTR and ABCE1-5'-UTR interactions appear to be immune evasion strategies of the SARS-CoV-2. A phylogenetically conserved RNA structure within the 3C

FIGURE 1

Recognition of +ssRNA viruses by the host innate immune pathways and generation of antiviral response. RIG-I and MDA5 recognize dsRNA, TLR3 and TLR7/8 sense dsRNA and ssRNA, respectively, and activate the indicated pathways to express type I and type III interferons and proinflammatory cytokines. dsRNA also activates RNase L, which cleaves the former. Viral proteins can damage mitochondria and/or activate the inflammasome. RIG-I, Retinoic acid-inducible gene; MDA5, Melanoma differentiation-associated gene 5; LGP2, Laboratory of genetics and physiology 2; MAVS, Mitochondrial antiviral signaling protein; TRAF3, TNF receptor associated factor 3; TBK1, TANK-binding kinase 1; IKKε, IκB kinase ε; IRF3/7, Interferon regulatory factor 3 or 7; TLR3, toll-like receptor 3; TLR7/8, toll-like receptor 7 or 8; TRIF, TIR-domain containing adaptor inducing interferon-β; RIP-1, receptor-interacting protein 1; TRAF6, TNF receptor associated factor 6; TAK1, TGFβ-activated kinase 1; IKKα/β, IκB kinase α/β; MyD88, Myeloid differentiation primary response 88; IRAK1,4, interleukin-1 receptor-associated kinase 1,4; OAS, oligoadenylate synthetase; 2′-5′ A, 2′-5′ oligoadenylate; NLRP3, NOD-like receptor thermal protein domain associated protein 3; cGAS, cyclic GMP-AMP synthase; cGAMP, cyclic GMP-AMP; STING, Stimulator of interferon genes,IL-1β, Interleukin-1β; IFN-α/β, interferon α/β; IFNαR1/2, interferon α/β receptor 1/2; JAK1, janus kinase 1; TYK2, tyrosine kinase 2; STAT1/2, signal transducer and activator of transcription ½; IRF9, interferon regulatory factor 9; ISRE, interferon stimulated response element; ISGs, interferon stimulatory genes. The figure is made in Microsoft PowerPoint and BioRender.

region of Polio virus ORF actively inhibits the endoribonuclease activity of RNase L (Han et al., 2007). RNA-protein interactome of the SARS-CoV-2 5'- and 3'-UTR regions also identified the antiviral role of LAMP2a, which is the receptor for chaperone-mediated autophagy (Verma et al., 2021). DENV-2 PR-2B sfRNA (sub-genomic RNA fragments) interacts with TRIM25, interferes with its deubiquitylation and inhibits RIG-I signaling (Manokaran et al., 2015). DENV-2 non-coding sfRNA interacts with G3BP1, G3BP2 and CAPRIN1 and inactivates them to suppress the expression of ISGs (Bidet et al., 2017). N6-methyladenosine (m6A) modification of HCV and SARS-CoV-2 RNA helps them in evading recognition by RIG-I (Kim et al., 2020; Li et al., 2021). MRM2/FTSJ2, a

mitochondrial 2'-O-methyltransferase interacts with the SARS-CoV-2 RNA, which might shield the viral RNA from recognition by MDA5 (Flynn et al., 2021). NSP15 of coronaviruses (CoVs) encode endoribonuclease EndoU, which cleaves the viral polyuridine sequence, inhibiting the activation of host immune sensors. The viral 5'-polyuridine from negative-sense viral RNA, termed PUN RNA is the product of polyA-templated RNA synthesis and is an MDA5-dependent pathogen-associated molecular pattern (PAMP) (Hackbart et al., 2020).

Notably, viral proteases also play a key role in inhibiting the host innate immune components. For example, Picornavirus 2A^pro disrupts MDA5-MAVS mediated antiviral innate immune response,

Coxsackievirus B3 (CVB3) 2A^{pro} cleaves MDA5 and MAVS by caspase-proteasome independent pathway, while poliovirus (PV) 2A^{pro} cleaves MDA5 via caspase-proteosome dependent pathway, CVB3 2A^{pro} cleaves TRIF, thus antagonizing type-I and type-III interferon production (Lind et al., 2016). The RLR signaling pathway is disrupted by the 3C^{pro} of picornavirus. 3C^{pro} of EV-A71 binds to the N-terminal CARDs of RIG-I, inhibiting its interaction with MAVS, and thus disrupting activation of type-I IFN response and 3C^{pro} of encephalomyocarditis virus (EMCV) cleaves RIG-I *in vitro*, promoting its degradation by the caspase pathway (Papon et al., 2009; Lei et al., 2010). EMCV 3C^{pro} also disrupts the TANK–TBK1–IKKε–IRF3 complex by cleaving TANK, thus decreasing type-I IFN production (Huang et al., 2017). FMDV 3C^{pro} disrupts NF-κB and IRF3 signaling pathway by cleaving the C-terminal zinc finger domain of IKKγ (Wang et al., 2012). FMDV 3C^{pro} and 2B proteins inhibit LGP2 expression (Zhu et al., 2017).

Proteases of coronaviridae also interferes with innate immune response (Lei and Hilgenfeld, 2017). SARS-CoV PL^{pro} reduces the ubiquitination of STING, TRAF3 and TBK1, thus prohibiting their activation (Chen et al., 2014). It also stabilizes the IκBα and inhibits NF-κB signaling pathway (Frieman et al., 2009). 3CL^{pro} (Also known as M^{pro}) of porcine deltacoronavirus (PDCoV) and porcine epidemic diarrhea virus (PEDV) cleaves IKKγ, thereby abrogating NF-κB signaling (Wang et al., 2016; Zhu et al., 2017a). 3CL^{pro} of PDCoV cleaves STAT2, 2A of EV71, 3C of EMCV, and 3C of FMDV cleave STAT1, and 3C, 3D proteases of EV71 cleave IRF9 and disrupt the JAK-STAT pathway (Du et al., 2014; Wang et al., 2015; Huang et al., 2017; Zhu et al., 2017b). Further, leader protease (L^{pro}), found in many picornaviruses targets multiple host innate immune factors to promote survival of the virus. The FMDV-L^{pro} cleaves LGP2, inhibiting the type I IFN response (Rodríguez Pulido et al., 2018). FMDV-L^{pro} also induces the degradation of p65/RelA subunit of NF-κB and decreases the expression of IRF3 and IRF7, leading to inhibition of the NF-κB activity and IFN-α/β expression, respectively (de Los Santos et al., 2007; Wang et al., 2010). A shorter form of FMDV-L^{pro}, known as Lb^{pro}, inhibits the ubiquitination of RIG-I, TBK1, TRAF6, and TRAF3, thereby inhibiting the secretion of type I IFNs (Wang et al., 2011). The L^{pro} of Theiler's murine encephalomyelitis virus (TMEV) and Mengovirus inhibits IRF3 activity and blocks IFN-β transcription (Hato et al., 2007; Stavrou et al., 2010). Mengovirus-L^{pro} also inhibits NF-κB activity, leading to inhibition of IFN-α/β expression in virus-infected cells (Zoll et al., 2002).

2.2 RNA-protein interactions drive the progress through different stages in the life cycle of +ssRNA viruses

The life cycle of a +ssRNA virus starts with the entry of the virus into the host cell. Post uncoating, the viral genome is released to the cytoplasm, where it serves as the template for translation of the non-structural and/or structural polyprotein, followed by their cleavage through autolysis and/or with the help of virus-encoded and/or host

proteases. Translation of proteins in +ssRNA viruses may be mediated via cap-dependent, cap-independent or a combination of both mechanisms. The presence of the 5'- end cap stabilizes the viral RNA and protects it from getting degraded by the host nucleases. The 5'-cap also enables cap-dependent translation of the viral RNA, using the host translation machinery. Both cap-dependent and cap-independent translation is driven by the interaction of viral genomic RNA with a temporally regulated complex of host translation factors. For example, RNA-protein interactome of the SARS-CoV-2-5'- and 3'-UTR RNAs show enrichment of host translation factors (Verma et al., 2021). Note that SARS-CoV-2 translation is a cap-dependent process. RNA-protein interactome of the Hepatitis E virus internal ribosome entry site (HEV-IRES), which drives cap-independent translation of the viral ORF4 protein, also shows enrichment of host translation factors (Kumar et al., 2023b). Poly(rC) binding proteins1 and 2 (also known as PCBP1 and PCBP2) enhance Polio virus translation by forming RNP complex with stem loop IV of the viral IRES (Blyn et al., 1996). The PTB-associated splicing factor (PSF) interacts with the cloverleaf structure in the IRES of coxsackievirus B3 (CVB3) and this interaction plays important role in viral translation (Dave et al., 2017). Another host protein, RNA helicase A (RHA) interacts with S fragment in the 5'-UTR of Foot-and-mouth disease virus (FMDV) RNA (Lawrence and Rieder, 2009). In coronaviruses, the cap and the poly (A) tail of the viral genomic RNA recruit initiation factor(s) that support the formation of a closed loop RNA conformation, which favors efficient translation initiation (Figure 2) (Walsh and Mohr, 2011; Lo et al., 2019; Stern-Ginossar et al., 2019; Sorokin et al., 2021).

2.2.1 RNA-protein interaction during the replication of +ssRNA viruses

Replication of the viral genome is central to the life cycle of a virus, which generates multiple copies of the viral genome to assemble progeny viruses. In the case of +ssRNA viruses, viral genomic RNA acts as the template and with the help of viral RNA-dependent RNA polymerase (RdRp) and many other viral and host proteins, viral genome is copied. RdRp usually binds to the +ssRNA virus genome at the 3'-end. Multiple host factors bind to the genome at the 5'- and 3'-ends, leading to the assembly of a RNA-protein complex, which facilitates circularization of the genome and formation of negative-strand RNA, sub-genomic RNAs and positive-strand genomic RNA (Figure 2). For example, genome circularization is important for replication of Flaviviruses (Villordo and Gamarnik, 2009). Nucleocapsid (N) protein of the Bovine Coronavirus (BCoV) interacts with both 5'- and 3'-ends of the viral genome, resulting in circularization of the viral genome, which is important for the synthesis of the negative strand RNA (Lo et al., 2019). Analysis of RNA-protein interactome of the SARS-CoV-2-5'- and 3'-UTR RNAs suggests PPI-mediated bridging of the 5'- and 3'- ends of the viral genome during replication (Verma et al., 2021). In the case of Zika virus (an enveloped positive strand RNA virus), interaction of the viral envelope (E) protein with multiple regions of Zika genomic RNA, [which includes two regions at the 5'- end

FIGURE 2
Simplified illustration of life cycle of +ssRNA viruses. The RNA virus life cycle has four major steps- entry, replication, assembly, and egress. After entry into the host and uncoating of the viral capsid, viral genomic RNA is translated to produce the non-structural polyprotein (NSP), which is subsequently processed into individual subunits. Viral RdRp assembles a RNA-protein complex, which interacts with the RNA-protein complexes assembled at 5'- and 3'- termini of the viral genomic RNA to form the viral replication complex. Viral genomic RNA likely forms a closed loop structure during replication. Antisense strand (-) as well as sub-genomic (sg) and genomic (g) RNA strands are synthesized by replication. Sub-genomic RNA is translated to produce the structural proteins (SP) that assembles the viral capsid, which encapsulates the genomic RNA. Progeny virions are subsequently released outside. A B C D illustrate 5'-UTR-interacting host proteins; a b c d illustrate 3'-UTR-interacting host proteins; h* illustrate host proteins interacting with the RdRp-bound RNA-protein complex.

(nt 135–294 and nt 734-899) and one region at the 3'- end (nt-10474-10644)] is important for viral replication (Hou et al., 2017). The stem loop I (SL-I) in the 5'-UTR of Polio viruses interacts with host PCBP2 and viral proteinase-polymerase precursor protein 3CD to form a ternary complex that is important for viral RNA replication (Gamarnik and Andino, 2000). 5'-UTR of the Enterovirus 71 RNA interacts with the hnRNP K and hnRNP A1, which is important for viral translation and replication (Lin et al., 2008; Levengood et al., 2013). La protein interacts with both the 3'- and 5'-UTRs of CVB3 independently of the poly(A) tail, and seems to play a role in mediating cross-talk between the 5'- and 3'-ends of the CVB3 genomic RNA, facilitating viral RNA replication (Cheung et al., 2007). In coronaviruses, genomic and sub-genomic RNAs consist of 5'- and 3'-UTRs at their terminals and a transcriptional regulatory sequence (TRS) within the 5'-UTR. TRS helps in template switching during the synthesis of the negative-strand RNA by base pairing

between the TRS-L and nascent TRS-BS by the viral transcriptase/replicase complex (Yang and Leibowitz, 2015).

2.2.2 RNA-protein interactions during progeny virus assembly and release

Translation of genomic and sub-genomic RNAs produce non-structural and structural proteins, necessary for replication and progeny virus assembly, respectively. Replication of the viral genome produces multiple copies of itself, which need to be protected from host endonucleases and thus are compactly packaged inside the viral nucleocapsid shell. The capsid protein of the virus directly interacts with the viral genomic RNA and on its own or with the help of M protein (Membrane/Matrix protein in many RNA viruses), genomic RNA is packaged into the nucleocapsid shell. Progeny viruses are subsequently released out by exploiting the host cellular transport machinery. Host RBPs are

FIGURE 3
Schematic of the RNA regulatory elements present in the genome of HCV and SARS-CoV-2. Stem loops in the 5' and 3' UTR (untranslated region) regions have been indicated in black color. Stem loops present in internal region are represented against corresponding proteins. Top schematic is for Hepatitis C virus which includes - C, Core/capsid protein; E1 and E2, Envelope glycoproteins; p7, Viroporin; NS, non-structural proteins. The bottom schematic is for SARS-CoV-2 which includes – ORF 1a to ORF14, Open reading frame; S, Spike protein; E, Envelope protein; M, Membrane protein; N, Nucleocapsid protein. S, M, E, N are the structural proteins.

involved in these steps as illustrated in the case of Flaviviruses (Diosa-Toro et al., 2020).

2.3 Impact of spatial and temporal binding of RBPs to the viral RNA

Localization of RBPs may be spatially restricted to specific intracellular organelles such as the ER, Golgi, lysosomes, recycling endosomes and autophagosomes. RBPs are also abundant in P-bodies and stress granules. RBPs may also be enriched at an intracellular site via RNA-protein/protein-protein interactions and liquid-liquid phase separation. Viruses may modulate the localization of RBPs or benefit from the presence of the RBP at a particular site. The ER and Golgi apparatus are essential for forming viral replication complexes and the biogenesis of viral membranes in many cases. HCV and DENV exploit ER-derived membranous webs, where RBPs like PTB stabilize viral RNA for replication (Anwar et al., 2009; Chatel-Chaix and Bartenschlager, 2014). Further, DENV 3'-UTR interacts with G3BP1/2 and DDX6, proteins found in stress granules and P-bodies, suggesting viral replication complexes localize between these granules (Ward et al., 2011). On the other hand, WNV disrupts P-body formation by recruiting DDX6 and other mRNA silencing components to viral replication sites, where they promote viral replication (Chahar et al., 2013). Lysosomes, endosomes, and autophagosomes are also key in viral entry, trafficking, replication, and survival. During SARS-CoV-2 infections, the autophagy receptor

SQSTM1 (p62) interacts with the viral RNA, inhibiting autophagy and generating autophagosomes that serve as replication platforms (Kamel et al., 2021a). A 2021 study highlighted how autophagosomes containing DENV proteins and genomic RNA evade immune detection (Wu et al., 2021). Temporal regulation further complicates this process, with RBPs binding viral RNA at distinct stages of infection. For example, ChIRP-MS and qTUX-MS using SILAC labeling have provided insights into temporal changes in RNA interactions during SARS-CoV-2 and DENV infections (Viktorovskaya et al., 2016; Flynn et al., 2021).

3 Regulatory elements in the +ssRNA virus genome mediate its interaction with viral proteins and host proteins

In contrast to DNA, RNA-RBP interaction is not necessarily sequence driven. Although there are well defined sequence motifs for recognition by specific RBPs, in many cases, RNA folds into secondary and tertiary structures generating specific conformations necessary for recognition by RBPs. Therefore, both sequence and structure of RNA regulatory elements are important for binding with RBPs. Regulatory elements are present at 5'-end, 3'-end and internal regions of the genome in +ssRNA viruses, schematically shown with examples of HCV and SARS-CoV-2 (Figure 3) (Tavares et al., 2021).

3.1 Regulatory elements at the 5′-UTR

5'-UTR of the +ssRNA viruses may be capped or uncapped. In the case of capped-RNA, UTR contains multiple stem loop (SL) structures, followed by Kozak sequence and initiation codon for the non-structural protein. Stem loops present in the 5'-UTR are important for protecting the RNA, assembling the translation initiation complex, and packaging the viral genome. They also aid in viral transcription and replication process. For example, 5'-UTR of SARS-CoV-2 spans 265 nucleotides and consists of 5 stem-loop structures. The transcriptional regulatory sequence (TRS) is present in the SL-III, which controls the discontinuous transcription (Liu et al., 2007; Sola et al., 2015; Miao et al., 2021). The SL-V has been indicated to be involved in viral RNA packaging and translation of ORF1ab polyprotein (Miao et al., 2021). In addition, both SL-III and SL-IV are targets for the binding of viral and cellular proteins, thus may play a role in viral replication (Sola et al., 2011; Madhugiri et al., 2016).

In the case of uncapped RNA, UTR contains multiple stem loop (SL) structures, followed by internal ribosome entry site (IRES) and initiation codon for the non-structural protein. IRES is a stretch of highly structured RNA elements, which directly recruit the initiation factors and promote translation through a scanning independent process, except type I IRES, which depends on the ribosomal scanning process. There are 5 major types of IRES based on their RNA structure and mode of ribosome recruitment. Notably, type I and type II IRES are found in the Picorna viruses such as PV and the FMDV, respectively. The PV IRES harbors six stem loops named as domain I to VI. The Domain I forms unique clover leaf structure and plays a critical role in replication of both the positive and the negative sense RNA. The domains II to VI are responsible for the PV IRES function. During PV infections, viral 2Apro cleaves the eIF4E binding N-terminal domain of the eIF4G without affecting its eIF3/eIF4A binding property. Stable association of the eIF4G with the PV IRES domain V enables its association with other initiation factors, leading to formation of the 43S preinitiation complex. The FMDV IRES is a classic example of the type II IRES. The domain IV of the FMDV IRES binds with scaffold protein eIF4G. The 3Cpro and Lpro of FMDV cleave the eIF4G. Importantly, the FMDV IRES skips ribosomal scanning, instead, IRES proximal stem loop formation brings 84 nucleotides downstream AUG, close to the first AUG to start the translation by direct ribosome transfer (Lee et al., 2017). The type III IRES is found in the 5'-UTR of the Hepatitis A virus genomic RNA (Brown et al., 1994). It requires eIF4E binding for translation initiation (Ali et al., 2001). The type IV IRES have been reported in the HCV (Hepatitis C virus)/HCV-like IRES. The 5'-UTR of HCV contains four domains: the domain I and II plays important roles in the viral replication while the domains III and IV are involved in translation (Khawaja et al., 2015; Kerr and Jan, 2016). The domains II and III contain several subdomains for interaction with the 40S ribosomal subunit. The type V IRES includes the long intergenic region (IGR) IRES, found between two open reading frames in the viral genomes and conserved in the dicistroviridae family. IGR IRES elements directly binds to the ribosomes and initiates translation with the

alanine-tRNAi (ala-tRNAi) instead of the met-tRNAi, without involving the eIFs (Wilson et al., 2000; Pestova and Hellen, 2003). Thus, RNA-protein interactions play indispensable roles in the function of viral IRESs.

3.2 Regulatory elements at the 3′-UTR

3'-UTR of the +ssRNA virus genome usually contains a stretch of Adenine, followed by multiple SLs, which are important for binding of the viral RdRp and other virus-encoded and host factors as well as for RNA-RNA interactions. 3'-UTR is important for viral replication, translation and evasion of host antiviral response. For example, SARS-CoV-2-3'-UTR is 228 nucleotides long and contains 4 SLs. Pseudo-stem-loop (PK), bulge stem-loop (BSL), and S2M domain (HVR) in the 3'-UTR are supposed to be important for the life cycle of the virus (Rangan et al., 2020). The 3'-UTR carries distinct nucleotide combinations such as CTC, TGT, CGT for every group i.e., SARS-CoV-2, SARS-CoV, and Bat-CoVs, respectively. These nucleotide combinations overlap with S2m, a highly conserved RNA motif, which likely have a role in viral pathogenesis (Kelly et al., 2021). These positions were also found to overlap with BSL and PK regions of the 3'-UTR among all βCoVs. The hypervariable region consists of an octa-nucleotide sequence (5'-GGA AGA GG-3') that is conserved among coronaviruses (Sola et al., 2011).

The coordinated interaction between 5'- and 3'-UTR through host and viral proteins forms the foundation for efficient replication of the virus (Nicholson and White, 2014). These long-range interactions create functional ribonucleoprotein complexes that enable three fundamental processes: genome cyclization, replication initiation and host immune modulation. The process begins with genome circularization, as exemplified by flaviviruses like DENV, where complementary sequences in the UTRs form panhandle structures that bring the RNA ends into proximity (Khromykh et al., 2001; Liu et al., 2020). This structural rearrangement is facilitated by host RNA-binding proteins such as DDX6, which specifically recognizes and stabilizes pseudoknot formations in the 3'-UTR to enhance both translation and replication efficiency (Liu et al., 2020). Similarly, in Hepatitis E virus, the 3'-UTR stem-loops SL1 and SL2 directly interact with the viral RdRp to initiate replication (Agrawal et al., 2001), while the 5'-UTR hairpin recruits the structural protein ORF2, likely for virion assembly (Surjit et al., 2004). Beyond structural roles, these terminal interactions serve as regulatory hubs. The polyadenylated 3'-UTR of HEV performs dual functions, serving as both a replication element and a potent activator of RIG-I-mediated innate immunity through its U-rich region (Sooryanarain et al., 2020). This exemplifies how viral RNA termini have evolved to balance replication needs with immune evasion strategies. The importance of host RBPs in maintaining these functional interactions is evident across virus families. Poliovirus employs hnRNP C as an RNA chaperone to keep its 3'-UTR in a single-stranded conformation optimal for replication initiation (Brunner et al., 2005; Ertel et al., 2010), while mouse hepatitis coronavirus utilizes PTB and hnRNP A1 to

physically bridge its 5' and 3' UTRs (Li et al., 1999; Barton et al., 2001; Huang and Lai, 2001). Even in bacteriophage systems like Qβ, conserved mechanisms exist where internal-3'-UTR base-pairing, mediated by host factors, facilitates replicase assembly (Wu et al., 2009). From flavivirus genome cyclization to coronavirus UTR bridging, the conserved requirement for 5'-3' communication mediated by specific RNA-protein interactions underscores their fundamental importance in the viral life cycle.

3.3 Cis-regulatory elements in the internal regions of the +ssRNA virus genome

Internal cis-regulatory elements refer to stable RNA secondary structures present in between the ORFs or in the coding region within the viral RNA. The presence of such cis-regulatory elements have been experimentally shown in the genome of +ssRNA viruses such as HCV and SARS-CoV-2 (Figure 3) (Tavares et al., 2021). For example, many cis-regulatory elements are found in the Core, NS4B and NS5B coding regions in the HCV genome (Tavares et al., 2021). Cis-regulatory elements are found in the ORF1a, ORF1b, S, ORF3a, E, M, ORF6, ORF7a/b, ORF8 and N coding region in the SARS-CoV-2 genome (Tavares et al., 2021). Further, nine TRS elements are present in the SARS-CoV-2 genome, which are important for sub-genomic RNA synthesis (Rangan et al., 2020). CRE is located in the 2C ORF of enteroviruses, the 2A ORF of species A rhinoviruses, the VP1 ORF of species B rhinoviruses, the VP2 ORF of species C rhinoviruses and cardioviruses, VP0 ORF of Parechovirus and upstream of the IRES in the FMDV (Mcknight and Lemon, 1998; Lobert et al., 1999; Goodfellow et al., 2000; Paul et al., 2000; Gerber et al., 2001; Mason et al., 2002; Al-Sunaidi et al., 2007; Cordey et al., 2008).

Another important internal regulatory element in the RNA virus genome is the frameshifting element. The frameshift element of SARS-CoV-1 has a pseudoknot (PK) structure. The dimerization domain of PK is critical for programmed ribosomal frameshifting (PRF), an essential event for forming ORF1a and ORF1b proteins from the same genomic region (Kelly et al., 2021). SARS-CoV-2 frameshifting element (FSE) is composed of a stem-loop attenuator, and a slippery sequence followed by a single-stranded spacer and an RNA pseudoknot. RNA-RNA interactions between the 3'-end of ORF1a and 5'-end of the ORF1b generates the FSE-arch, which is highly conserved among SARS-related coronaviruses and possess high folding stability *in vivo*. The FSE-arch likely controls the FSE activity (Ziv et al., 2020; Zhang et al., 2021).

4 Methods to generate RNA-protein interactome of +ssRNA viruses

4.1 Generation of RNA-protein interactome using biological samples

RNA-protein interactions can be experimentally demonstrated using either RNA-centric or protein-centric approaches. This review focuses on RNA-centric methods to identify the RBPs associated with the viral genomic or sub-genomic RNA. RNA-centric methods may be broadly classified into *in vitro* and *in vivo* methods. *In vivo* methods can detect the interaction between the whole viral genome or parts of viral genome and associated proteins whereas *in vitro* methods are generally used to detect the interaction between parts of viral genome and associated proteins.

4.1.1 *In vitro* methods to detect RNA-protein interactions

In-vitro methods such as pull down and microarray-based binding assays are used to detect the interaction between parts of a viral genome and associated proteins (Figure 4). These methods are beneficial when the test RNA or protein is unstable, not expressed well *in vivo* or the RNA binding proteins are less abundant *in vivo*. *In vitro* assays are also useful in characterizing the molecular details of a particular RNA-protein interaction at nucleotide and amino acid level.

In the case of pull down assay, the 5'- or 3'- end biotin-labeled RNA is synthesized *in vitro* and incubated with cellular extract, followed by isolation of the RNA-protein complex using streptavidin beads (Zheng et al., 2016). Alternatively, the RNA-protein complex may be isolated by using a biotin-labeled aptamer sequence against the test RNA (Srisawat and Engelke, 2001) (Figure 5A). In another study, Cys4 hairpin loop-tagged RNA has been used to select the test RNA bound complex, followed by elution of the test RNA-protein complex using imidazole, which activates the Cys4 endoribonuclease that cleaves the Cys4 RNA (Lee et al., 2013). The later technique may be useful in reducing the background signal as endogenously biotinylated proteins directly bind with the streptavidin beads irrespective of their RNA binding activity. RNA-protein complex may be UV crosslinked in an *in vitro* pull down assay. In a microarray-based binding assay, individual proteins are spotted on a microarray slide, followed by hybridization with a labeled test RNA (such as Cy5-labeled RNA) (Kretz et al., 2013) (Figure 5B). In both approaches, non-specific proteins are removed by multiple washing steps and interaction partners are detected by mass spectrometry or by fluorescence reading in the microarray scanner, respectively. Microarray-based binding assay detects direct interactions between the RNA and protein whereas pull down assay can detect both direct and indirect interaction partners. Although *in vitro* assays are simple and straight forward, it is limited by the fact that *in vitro* synthesized RNA may not fold properly or lack the native structure and modifications required for interaction with a particular interaction partner or protein complex. Further, High or low abundance of the protein(s) in the cellular extract may influence the result and proteins spotted on the microarray slides may not be properly folded or lack the required post-translational modification(s) or physiological environment required for interaction with the test RNA.

4.1.2 *In vivo* methods to detect RNA-protein interactions

The limitations of the *in vitro* assays are partly resolved through *in vivo* methods, which offer physiological and functional advantages.

FIGURE 4

In vitro and in vivo methods to detect RNA-Protein interactions in +ssRNA viruses. RaPID, RNA-Protein Interaction Detection Assay; RNA BioID- RNA proximity biotinylation; incPRINT- In-cell protein-RNA interaction; MS2-BioTRAP, MS2-*in vivo* Biotin Tagged RNA Affinity Purification; MTRAP-MS, MS2-tagged RNA affinity purification and Mass spectrometry; RAP-MS, RNA Antisense Purification followed by Mass Spectrometry; iDRiP, Identification of Direct RNA-interacting Proteins; TRIP, Tandem RNA Isolation Procedure; PAIR, Peptide-Nucleic acid Assisted Identification followed by Mass Spectrometry; vRIC-MS, Viral RNA Interactome Capture followed by Mass Spectrometry; VIR-CLASP, Viral Cross-linking And Solid-phase Purification; CHART-MS, Capture Hybridization Analysis of RNA Targets followed by Mass Spectrometry; ChIRP-MS, Comprehensive identification of RNA-binding proteins by mass spectrometry.

Various technologies to detect RNA-protein interactions *in vivo*, with or without crosslinking of the complex are summarized (Figure 4). Although phase separation-based techniques for isolating RNA-protein complex have emerged to be a powerful approach to identify RBPs (Queiroz et al., 2019; Trendel et al., 2019; Urdaneta and Beckmann, 2020), this review will focus on techniques relevant to detection of viral RNA binding proteins.

4.1.2.1 *In vivo* methods to detect RNA-protein interactions in non-crosslinked samples

Among the crosslinking-independent *in vivo* techniques, Yeast three hybrid is a classical genetics technique, useful in detecting direct interaction between a test RNA and protein(s) in a physiological environment (SenGupta et al., 1996). Here, host proteins are expressed in the yeast cells using a cDNA expression library of the host cell type of interest as a fusion protein with the GAL4-AD (activation domain of the GAL4 transcription factor) (Figure 6A). Viral RNA is expressed as a fusion with MS2-binding RNA element at the 5'- or 3'-end. Specific interaction of viral RNA with a protein activates the *HIS3* (imidazoleglycerol-phosphate dehydratase) and *lacZ*(β-galactosidase) reporter genes, allowing growth of the yeast transformants in histidine deficient medium and colorimetric scoring

by quantification of β-galactosidase activity, respectively. Interacting proteins are subsequently identified by isolating the cDNA clone and sequencing the plasmid DNA. Although the assay is conducted in a cellular milieu, which likely enables unbiased assessment of the interaction partners, screening of the cDNA library produces a lot of false positives and chances of misfolding of the test RNA and prey proteins cannot be ruled out.

Crosslinking-independent mammalian cell culture based proximity proteome labeling techniques such as RNA-protein interaction detection (RaPID) and the RNA proximity biotinylation (RNA BioID) have been reported (Ramanathan et al., 2018; Mukherjee et al., 2019; Verma et al., 2021; Kumar et al., 2023b). Proximity proteome labeling techniques rely on enzymes such as: Biotin ligases like BASU, BioID (mutant variant of the BirA enzyme) and its derivatives, which covalently attach biotin to proteins within 10-20nm radius; or ascorbic acid peroxidase (APEX) and its derivatives, which converts exogenously supplied biotin-phenol to biotin-phenoxyl radicals upon treatment with H_2O_2, resulting in covalent labeling of proteins. Both BASU and APEX label the proteins within 20nm radius, however APEX labeling is very fast (~1 min) compared to labeling by BASU (several hours) (Rhee et al., 2013; Paek et al., 2017; Samavarchi-Tehrani et al., 2020). Note that APEX labeling also requires

FIGURE 5
Schematic of *in vitro* methods to detect RNA-protein interactions. **(A)** *In-vitro* pull down assay. **(B)** Microarray-based binding assay.

treatment of cells with H_2O_2. Hence, choice of the proximity labeling enzyme is dependent on the experimental design.

RaPID assay identifies direct and indirect interaction partners of small RNA fragments (~132 nucleotides), which are expressed as chimeric RNA in fusion with an aptamer sequence such as Box B stem loop, which is recognized by the ΛN peptide (Figure 6B). A biotin ligase [BASU] fused to the ΛN peptide is recruited to the Box B, which biotinylates all proteins in its close proximity (~10-20nm range), including those associated with the RNA of interest. Biotinylated proteins are enriched and identified by LC-MS. RaPID assay has the advantage of detecting weak and transient RNA-protein interactions, however, the assay depends on overexpression of the test RNA. To overcome the limitation of overexpression of the test RNA and improve the efficiency of proximality labeling, Mukherjee et al., developed the RNA-BioID assay using genetically modified mouse embryonic fibroblasts (MEF). They used MEFs in which endogenous β-actin gene copies were replaced by β-actin with 24 MS2 binding sites (MBS) in their distal 3′-UTR and there was stable expression of a fusion of the nuclear localized signal (NLS), MS2 coat protein (MCP), GFP, and BirA* (MCP-GFP-BirA*) (Mukherjee et al., 2019). This approach identified a much higher number of interaction partners of the β-actin RNA, compared to other affinity-based methods.

Another crosslinking independent, RNA-tagging based *in vivo* technique is in-cell protein-RNA interaction (incPRINT). Here, the test protein is tagged with Flag epitope and the test RNA is tagged with MS2 stem loop sequence, which is expressed in cells along with MS2-coat protein fused to Luciferase. Test protein is captured from the cell lysate by Flag affinity beads, followed by Luciferase assay to detect its interaction with the test RNA (Graindorge et al., 2019) (Figure 6C). The assay may be scaled up to screen a library of Flag-tagged proteins against a test RNA.

Recent studies also demonstrated the utility of CRISPR-Cas targeting system in detecting RNA-protein interactions without crosslinking of the samples (Han et al., 2020; Lin et al., 2020; Yi et al., 2020; Zhang et al., 2020; Li et al., 2021). Using guide RNA (gRNA) specific to the test RNA along with a catalytically dead Cas13 (dCas13, dCasRx) fused to a biotin ligase (BASU, PUP-IT), it is possible to biotinylate proteins interacting with any endogenous RNA, which can be subsequently captured by streptavidin beads and identified by LC-MS (Figure 6D). Several modifications in the initial technique have been reported, which further improved the efficacy of the technique (Labun et al., 2019; Wessels et al., 2020). At the same time, more research is required to rule out the possibility of background noise due to off target binding by the gRNA. Note that, towards reducing the background noise in biotin ligase-based

FIGURE 6

Non-crosslinked *in vivo* methodologies to study RNA-protein interactions. **(A)** Yeast three hybrid (Y3H) assay, **(B)** RaPID assay and RNA BioID, **(C)** incPRINT, **(D)** dCas13-based technique.

RNA-protein interaction detection techniques, recent studies have developed split biotin ligases, which gains enzymatic activity only when associated with the target RNA (Shekhawat and Ghosh, 2011; De Munter et al., 2017; Schopp et al., 2017; Cho et al., 2020).

4.1.2.2 *In vivo* methods to detect RNA-protein interactions in crosslinked samples

Crosslinking of the RNA-protein complex *in vivo* arrests the interactions, which helps in capturing of weak and transient interactions. While crosslinking enhances the stability of the complex and increases the number of RBPs in the data set, there

is a possibility of capturing nonspecific proteins due to over crosslinking and loss of bona fide RBPs due to inefficient crosslinking of weak interactions in a multiprotein complex. Both UV and formaldehyde are widely used in different techniques to crosslink the RNA-protein complexes *in vivo*. The choice of technique should be based on approximate information of the abundance of target proteins and test RNA, the strength of the RNA-protein interaction and size of the RNA-protein complex. It is noteworthy that although UV rays irreversibly crosslink nucleotide-protein interactions at zero distance via a covalent bond, it works less efficiently and weak interactions might be missed. On the other

hand, overexposure to UV may have undesired consequences on the cellular processes. Hence, optimization of UV cross linking duration is important for success of the experiment. Formaldehyde reversibly crosslinks protein-protein, protein-DNA and protein-RNA interactions within 2A°, via covalent bond. However, formaldehyde crosslinking is less specific in capturing only RNA-protein interactions. Both RNA-tag and RNA hybridization based approaches have been used to detect RNA-protein interactions in crosslinked samples.

RNA-tag based techniques depend on an aptamer sequence such as the MS2 stem loop element to capture the target RNA. RNA-protein interaction is stabilized by UV crosslinking and interacting proteins are revealed by pull down assay-LC-MS/MS. Techniques such as MS2-BioTrap, MS2-TRAP and MTRAP-MS are based on the above principle (Figure 7) (Tsai et al., 2011; Yoon et al., 2012; Liu et al., 2015).

RNA hybridization based approach have been employed in multiple techniques. Capture Hybridization Analysis of RNA Targets (CHART) and Comprehensive identification of RNA-binding proteins by mass spectrometry (ChIRP-MS) are two popular RNA hybridization-based techniques, in which formaldehyde is used to crosslink the samples (Figure 7) (West et al., 2014; Chu et al., 2015). Cells are treated with formaldehyde, followed by hybridization with biotinylated oligonucleotides (c-oligos). RNA-bound RBPs are purified using streptavidin beads, followed by protein identification by western blot or LC-MS/MS, for CHART and ChIRP, respectively. ChIRP was used to compare the RNA-protein interactome of SARS-CoV-2, Zika, and Ebola viruses (Flynn et al., 2021; Zhang et al., 2022). However, this technique is limited by the inefficiency of c-oligos to bind the different target loci with equal efficiency.

Compared to DNA based oligonucleotide probes used in CHART and ChIRP, antisense RNA based probes are more

FIGURE 7
Schematic of major steps in the *in vivo* experimental methods involving mass spectrometry analysis to study viral RNA-protein interactions. White circle indicates formaldehyde crosslinking, white and Blue star indicate UV crosslinking, yellow circle indicates biotin tag.

TABLE 1 Advantages and limitations of notable methods to study RNA-protein interactions.

Technique	Advantages	Limitations
In-vitro methods	• Ability to use purified and/or well-defined RNA and protein components. • Both direct and indirect interaction can be detected. • Molecular details and biomolecular characteristics of the interactions can be easily analyzed. • Ability to detect the interaction between unstable RNA and proteins.	• Interactions obtained in such artificial condition may not be physiologically relevant. • It may not be possible to generate the *in vivo* structure of the test RNA and/or protein *in vitro*.
Yeast three hybrid	• Ability to unbiasedly detect the RNA-protein interactions between a test in a physiological environment. • Handy technique to characterize molecular details by mutagenesis.	• High false positive rate • Yeast may not be the ideal host for the test RNA-protein interactions.
RaPID, RNA BioID	• *In vivo* technique, independent of cross-linking of the test RNA-protein interactions. • Real time labeling of interacting proteins • Transient and stable interaction can be detected.	• Test RNA is not in its native state. • Length of RNA is restricted to around 200 nucleotides. • Endogenous biotinylated proteins need to be filtered out. • May not be possible to detect all RBPs present in a multiprotein complex.
lncPRINT	• Mass spectrometry-based analysis is not required to identify the proteins.	• A limited number of proteins can be identified.
dCas13-based technique	• Specificity of CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) technology extended to detect RNA-protein interactions.	• The design of effective gRNA is challenging. • Proximity-labeling enzyme biotinylates proteins even when it does not bind to RNA thus increasing non-specificity.
MS2-Bio Trap	• Test done in physiological condition. • High specificity due to UV-crosslinking.	• Test RNA is not in its native state. • Fusion of test RNA with MS2 RNA might alter its property.
RAP-MS iDRiP	• Detection of direct RNA-protein interactions. • Endogenous RNA-protein complex can be purified intact. • High specificity. • Test done in physiological condition.	• A high number of cells is required. • Technically demanding.
PAIR-MS	• Peptide-nucleic-acid (PNA) probe used. • High specificity. • Photoactivated cross-linking and UV cross linking.	• Synthesis of probe is costly.
vRIC-MS VIR-CLASP	• Entire viral RNA interactome can be captured. • High specificity. • Photoactivated cross-linking and UV cross linking.	• Selecting the right nucleoside analogue may be challenging due to its toxicity on the target cell and efficiency of its incorporation into the target RNA.
CHART-MS ChIRP-MS	• RNA-protein interaction are detected in a cellular environment. • Handling is easy due to the use of DNA based probe.	• Additional RNase H step is required to identify free sites for probes. • Inefficiency of C-oligos to bind the different target loci with an equal efficiency. • Formaldehyde cross linking is less specific. • DNA probes show high nonspecific signal.

specific and expected to show less background noise. However, the use of RNA probes require more stringent experimental conditions due to inherently fragile characteristic of the RNA. RNA antisense purification (RAP), identification of Direct RNA-interacting Proteins (iDRiP) and tandem RNA Isolation Procedure (TRIP) are some notable techniques based on antisense RNA hybridization (Figures 4, 7) (McHugh et al., 2015; Minajigi et al., 2015; Matia-González et al., 2017).

Further, Peptide-nucleic-acid (PNA) based probe has been used to detect RNA-protein interactions. In Peptide-nucleic-acid Assisted Identification (PAIR) assay, PNA is used to hybridize with the target RNA. PNA contains a photoactivable amino acid adduct, p-benzoyl phenylalanine (Bpa), which captures the nearby RBPs by photoactivated cross-linking. UV is used to covalently cross-link the PNA-Bpa with adjacent RBPs. Finally, PNA-RBP complexes are isolated using sense oligonucleotide magnetic beads,

followed by LC-MS mediated identification of bound proteins (Zeng et al., 2006).

To address the low efficiency of UV cross linking, photoactivable ribonucleoside-enahnced (PAR) crosslinking techniques such as Viral RNA Interactome Capture (vRIC) and Viral cross-linking and solid-phase purification (VIR-CLASP) have been developed. Here, cellular RNA is metabolically labeled with 4-thiouridine (4SU), followed by UV cross-linking at 365nm (longer wavelength compared to conventional UV crosslinking at 254nm) and capture of the RNA-protein complex. After RNase digestion, quantitative proteomics is employed to reveal the captured RBPs (Figure 7). vRIC identified the RNA-protein interactome of SARS-CoV-2 and Sindbis virus (Kamel et al., 2021a; Kamel et al., 2021b). VIR-CLASP was used to identify the RNA-protein interactome of pre-replicated genome of the Chikungunya virus (Kim et al., 2020). Choice of the nucleoside analogue is decided based on its toxicity on the target

cell and efficiency of its incorporation into the target RNA. A comparison of the advantages and limitations of different methods is summarized in Table 1.

4.2 Generation of RNA-protein interactome using predictive modeling

Predictive modeling relies on two essential components: algorithms and the data used to train them. In the case of RNA-protein prediction models, large-scale interaction data are required, typically generated through experimental means. Over the years, various databases containing such data have been established through wet lab experiments, literature mining, or computationally predicted interactions (Table 2). For instance, the ENCODE database contains eCLIP datasets from 223 experiments conducted in HepG2 and K562 cell lines, capturing interactions with 150 RBPs. This repository of extensive experimental evidence contributes to the

robustness of the models, particularly as they rely on algorithms that require substantial data inputs. Using the above-mentioned source data, advanced computational tools have been developed to understand the intricacies of RNA-protein interactions and generate the RNA-protein interactome. Several studies have been dedicated to developing sophisticated algorithms that utilize sequence information to identify critical features indicative of RNA-protein binding affinity (Horlacher et al., 2023). The primitive approaches relied on sequence similarity, identifying patterns in sequences and their known interactions. While these models have low computational costs, they lack robustness. Understanding the fact that proteins adopt 3D structures before performing functions, alternative approaches leveraged upon structural information to provide a better understanding of potential affinities. Both methods have advanced significantly over time. However, fully deciphering structural information is complex due to the dynamic nature of proteins, making it challenging to comprehensively capture their true behavior. To address this, hybrid

TABLE 2 Web-based resources for extracting RNA binding proteins, arranged in reverse chronology.

Method	Year	Description	Evidence	Reference
HydRA	2023	It is an ensemble RBP classifier, combines intermolecular protein interactions and internal protein sequence patterns, employing SVMs, CNNs, and Transformer-based models to predict RNA-binding capacity with high specificity and sensitivity.	Prediction	(Jin et al., 2023)
Pprint2	2023	It utilizes machine learning and deep learning for predicting RNA-interacting residues in proteins.	Prediction	(Patiyal et al., 2023)
RBPbind	2022	It predicts RNA-protein interaction probabilities, incorporating sequence specificity from RNAcompete experiments into Vienna RNA package recursions, enhancing accuracy.	Prediction	(Gaither et al., 2022)
RBPmap	2021	It uses sub-sequence that notably aligns with the RBP motif and additionally weaker matches surrounding the motif.	Prediction	(Paz et al., 2022)
RBPsuite	2020	It is a deep learning based method for predicting RBP binding sites on both linear and circular RNAs.	Prediction	(Pan et al., 2020)
RBinds	2020	It integrates RNA structure into networks for binding site prediction, with visualization and simulation tools.	Prediction	(Wang and Zhao, 2020)
ENCODE	2020	Experimental evidence in K562 and HepG2 cells through mapping and analysis of 1223 replicated datasets encompassing 356 RBPs.	Experimental	(Van Nostrand et al., 2020)
NPInter v4.0	2020	It hosts more than 600000 interaction of non-coding RNA with proteins.	Experimental	(Teng et al., 2020)
SMARTIV	2018	It predicts motifs from enriched k-mers, integrating information from ranked RNA sequences and their predicted secondary structure.	Prediction	(Polishchuk et al., 2018)
RNAct	2018	It provides an easy-to-use view of protein-RNA interactions in model organisms.	Experimental	(Lang et al., 2019)
omiXore	2017	It identifies RNA binding sites by discriminating interacting protein-RNA pairs using UV cross-linking data.	Prediction	(Armaos et al., 2017)
BindUP	2016	It predicts nucleic acid binding function using protein electrostatic features and structural properties, providing visualizations of electrostatic surface patches.	Prediction	(Paz et al., 2016)
CLIPdb	2015	It hosts 395 publicly available CLIP-seq data sets for 111 RBPs from four organisms: human, mouse, worm and Yeast.	Experimental	(Yang et al., 2015)
RAID	2014	It hosts more than 6100 RNA-associated interactions obtained by manually reviewing more than 2100 published papers.	Experimental	(Zhang et al., 2014)
PRIDB	2011	The Protein–RNA Interface Database (PRIDB) is a comprehensive database of protein–RNA interfaces extracted from complexes in the Protein Data Bank (PDB).	Experimental	(Lewis et al., 2010)
RBPDB	2010	It is a collection of experimental observations of RNA-binding sites, both *in vitro* and *in vivo*, manually curated from primary literature of four metazoan species (human, mouse, fly and worm).	Experimental	(Cook et al., 2010)

methods combining sequence and structure-based approaches have emerged to balance complexity and computational costs. A chronological list of such methods, along with their methodological descriptions, is listed (Table 3).

Although such algorithms are not exclusively designed to predict viral RNA and protein interactions, researchers have advanced these tools to predict inter-species interactions with promising results (Kazachenka et al., 2018). This endeavor involves a pipeline approach, employing established RBP interaction prediction algorithms (Figure 8). For algorithm training, datasets such as those from the ENCODE project has been used in this pipeline (Van Nostrand et al., 2020). First, the dataset undergoes pre-processing, wherein a defined window is established around each peak, facilitating the identification of potential binding sites for subsequent analysis. Later, this data is compared with peak information from RBPs used as controls, with a two-fold change along with statistical significance is considered. Sampling and randomization strategies are employed to mitigate false positives. Diverse algorithms, including recurrent neural networks (RNNs), extended short-term memory networks (LSTMs), and convolutional neural networks (CNNs) are utilized, employing supervised training methodologies with hyper-parameters such as sequence window size, algorithm layers, and learning rate. A negative sampling strategy is adopted for extracting sequence-level binding information, where the center of the window serves as the nucleotide of interest. Predictions are generated for randomized sequences, and a score is computed like a p-value. A high score indicates similarity in binding affinity between randomized and actual sequences,

suggesting that binding may not be solely sequence or site-driven. This holistic approach yields a nuanced understanding of the RNA-protein interactome within the context of the entire genome.

4.3 Validation of the RNA-protein interactome data generated *in silico* or through Omics based technologies

Validation of the interactions between RBPs and their targets is crucial for distinguishing biologically meaningful associations from non-specific or background signals inherent in techniques like RNA-protein crosslinking or affinity purification. To minimize experimental noise, stringent controls such as mock immunoprecipitations (IPs), untagged viral RNA controls, or genetically modified cell lines with targeted RBP knockouts—are essential for establishing a reliable baseline. Mass spectrometry-based peptide identification platforms, including Mascot, MaxQuant, and FragPipe, enhance specificity by resolving ambiguous spectral data, while the CRAPome database helps systematically filter out common contaminants. Further, computational tools like Differential Enrichment analysis of Proteomics data (DEP), SAINTq, MIST score, and CompPASS, may be used, which apply stringent statistical criteria to enrich high-confidence interactors. A comprehensive computational pipeline processes raw affinity purification-mass spectrometry (AP-MS) data, performs quality control, and ranks biologically relevant bait-prey pairs across replicated experiments using these scoring methods (Verschueren

TABLE 3   Recent evolution of methods (from sequence-based to structural approaches) for predicting RNA-binding protein affinity.

Method	Year	Evidence	Architecture	Reference
BERT-RBP	2022	Sequence	Language model	(Yamada and Hamada, 2022)
DeepPN	2022	Sequence	CNN, GCN	(Zhang et al., 2022)
MultiRBP	2021	Sequence, structure	CNN	(Karin et al., 2021)
PRISMNet	2021	Sequence, structure	CNN	(Xu et al., 2023)
RNAprot	2021	Sequence, genomic annotation conservation	LSTM	(Uhl et al., 2021)
Multi-resBind	2021	Sequence, genomic annotation	CNN	(Zhao and Hamada, 2021)
RBP-ADDA	2021	Sequence	Adversarial domain adaptation	(Liu et al., 2022)
ResidualBind	2021	Sequence	CNN	(Koo et al., 2023)
kDeepBind	2021	Sequence	k-mer embedding CNN	(Tahir et al., 2021)
RBPSpot	2021	Sequence, structure	Embedding DNN	(Sharma et al., 2021)
DeepCLIP	2020	Sequence	CNN, BLSTM	(Grønning et al., 2020)
DeepRiPe	2020	Sequence, genomic annotation	CNN	(Ghanbari and Ohler, 2020)
RPI-NET/RNAonGraph	2020	Sequence, structure	GNN	(Yan et al., 2020)
DeepRKE	2020	Sequence, structure	Embedding, CNN, BLSTM	(Deng et al., 2020)
iDeepMV	2020	Sequence	Multi-view CNNs ensemble	(Yang et al., 2021)
DeepA-RBPBS	2020	Sequence, structure	CNN, biGRU	(Du et al., 2022)
MSC-GRU	2020	Sequence	CNN, biLSTM	(Shen et al., 2019)

FIGURE 8

Schematic of computational methods to study RNA-protein interactions.

et al., 2015). Post-processing filters such as false discovery rate (FDR) thresholds, fold-change cutoffs relative to controls, and consistency across replicates further help to eliminate spurious interactions.

In addition, an unbiased and independent method should be employed to reproduce the RNA-protein interactions identified in one method. For example, a combination of biochemical and imaging techniques are ideal for validating a subset of the interactome data. Super-resolution microscopy and advanced imaging techniques now allow real-time visualization of RBP-viral RNA dynamics within organelles, uncovering transient interactions previously missed. For example, the localization of HEV in recycling endosomes was studied using these imaging technologies (Bentaleb et al., 2022) and DENV and HCV replication and assembly were visualized using transmission electron microscopy (Chatel-Chaix and Bartenschlager, 2014). These methods are robust and broadly applicable across +ss RNA viruses, providing valuable insights into how different viruses manage viral RNA within cells, and identifying conserved or distinct mechanisms for potential antiviral targets.

RNA-SELEX (Systematic Evolution of Ligands by Exponential Enrichment) has emerged as a powerful tool to identify the specific RNA targets through iterative rounds of selection and amplification (Ellington and Szostak, 1990; Tuerk and Gold, 1990). SELEX has been used for determining the binding site of a protein on RNA (Manley, 2013). RNA-based Capture-SELEX has been used for selecting small molecule-binding aptamers (Ye and Jankowsky, 2020). Analogous to SELEX, another notable method, named massively parallel RNA assay combined with immunoprecipitation

(MPRNA-IP) has been developed for high-throughput analysis of RNA–protein interactions *in vivo* (Lee et al., 2024). These methods are useful for improving the accuracy of molecular characterization and validation of the RNA-protein interactions.

Importantly, vast datasets generated in SELEX have expanded the scope of machine learning models by incorporating information about the intermediate interaction steps, in contrast to traditional machine learning (ML) models, which often rely on the final binding information, overlooking the iterative modification in interaction. ML models can learn patterns from SELEX data to predict which RNA sequences will likely bind a given protein with high affinity. It can also help identify sequence motifs, secondary structures, or physicochemical properties important for binding. Databases like HTPSELEX has been developed for training models and tools like DeepPBS (a geometric deep-learning model), GraphProt, BindSpace have been developed to predict RNA-protein binding (Maticzka et al., 2014; Yuan et al., 2019; Mitra et al., 2024). These advanced methods have leveraged the richness of SELEX datasets.

5 Importance of RNA-protein interactome of viruses in decoding the viral life cycle and antiviral discovery

It is important to evaluate the functional significance of the RNA-protein interactions to understand the molecular details of the

viral life cycle and identify new targets for antiviral development. Suitable experimentally amenable tools such as non-infectious replicon of the virus or infectious/attenuated virus strains are useful resources for such studies.

5.1 Approaches to decode the life cycle of +ssRNA viruses using the viral RNA-protein interactome dataset

The RNA-protein interactome of a +ssRNA virus constitutes a set of proteins that directly or indirectly associate with the viral genome. These proteins need to be prudently analyzed to interpret and extrapolate their biological functions in the infected cells. This information forms the basis to hypothesize a mechanism of viral life cycle and pathogenesis, which is subsequently evaluated by suitable experimental models. Enrichment analysis has emerged as a standard approach to analyze large gene lists to produce a data-driven information that is easier to interpret. This analysis involves statistical testing of pathways and processes for over-representation in the experimental gene list compared to what would be expected by chance. Several common statistical tests are utilized, considering factors such as the number of genes detected in the experiment, their relative rankings, and the number of annotated genes. Some well-known web-based applications for such analysis include the Kegg pathway, Reactome pathway, GSEA (gene set enrichment analysis), Panther, and Gene Ontology (Subramanian et al., 2005; Thomas et al., 2022; Aleksander et al., 2023; Kanehisa et al., 2023; Milacic et al., 2024). These tools facilitate the identification of key pathways, functions and processes that are highly influenced by the identified set of genes. Moreover, RBP2GO and the RBP Image Database play crucial roles in elucidating the role of RBPs in viral infections (Caudron-Herger et al., 2021; Benoit Bouvrette et al., 2023). RBP2GO and the RBP Image Databases provide ontological information about the functions, processes, and cellular locations of RBPs, shedding light on their involvement in viral replication, RNA processing, and host immune responses. Once a hypothesis is formulated based on the acquired knowledge, appropriate experimental models are designed to validate the predictions.

5.2 Methods to unlock the therapeutic potential of RNA-protein interactome of +ssRNA viruses

As mentioned above, functional analysis of RNA-protein interactome data provides significant insight into the life cycle and pathogenesis of the corresponding virus. Such intricate understanding of viral lifecycle helps to identify and experimentally validate potential antiviral targets. Both the interactome data and the antiviral targets may be considered for screening antiviral drugs. Computational or experimental model-based screening methods may be followed to identify antiviral drugs either by *de novo* [identification of antiviral

potential of a new chemical entity (NCE)] or drug repurposing [identification of a new therapeutic application of an existing drug] approach. *De novo* drug discovery is extremely expensive and time consuming whereas the drug repurposing strategy holds the potential of immediate therapeutic impact at a much lower cost. The main advantage of repurposed drugs is attributed to their prequalification through safety and toxicity tests in preclinical and human trials. Multiple computational methods may be pursued to discover antiviral drugs. Once a drug candidate is identified, it should be validated using wet lab experiments before proceeding with preclinical studies.

5.2.1 Computational methods for capturing drug targets

The *in silico* drug discovery pipeline begins with target identification, which is very challenging, as proteins may require one or more interaction partners to execute essential functions. To address these limitations, different algorithms are employed to prioritize network nodes/proteins based on sensitivity or their potential to induce phenotypic changes. Some notable tools include CaNDis, CytoHUBBA, and NetEPD (Table 4). Another important aspect is identifying proteins that can control the information flow in the network, which can be obtained using tools like konnect2prot and NetControl4BioMed (Table 4). Sometimes, we are also interested in exploring proteins with similar functions to known therapeutic candidates for various reasons, such as being non-targetable or crucial for the system. In such cases, we can use guilt-by-association-based methods like Netpredictor to identify proteins with similar functions.

Later, the identified target needs to be modulated (activated or inhibited), which often requires small molecules, due to their various pharmacokinetic properties. These molecules could be newly synthesized or already available drugs. Such choices are made based on factors like time, cost, availability etc.

5.2.2 Data driven screening methods for small molecule identification

The data-driven drug discovery process depends on online resources, including clinically oriented drug databases (for example, PharmGKB and RxList) and chemically oriented drug databases (for example, Zinc, TTD and PubChem). While clinically oriented drug databases provide in-depth clinical information, chemically oriented drug databases generally provide nomenclature and structural properties of the compounds (Table 5).

Both these databases have been used by several laboratories as source data for developing drug discovery tools. For example, Drug Bank was developed by combining the attributes of clinically and chemically oriented drug databases to serve as a handy yet comprehensive tool to search drug molecules and get details of their sequence, structure, mode of action, targets as well as biological or physiological consequences of drug action (Knox et al., 2024). Drug central is a platform on similar line (Avram et al., 2023). ChEMBL is another notable online tool that consolidates the bioactivity information of drugs (Zdrazil et al., 2024).

TABLE 4 Useful web-based resources for antiviral discovery (chronologically sorted).

Name	Year	Description	Reference
cytoHUBBA	2014	It is a Cytoscape plugin offering diverse topological analysis methods for ranking network nodes, with the newly proposed "maximum neighborhood component" (MCC), demonstrating superior precision in predicting essential proteins from the Yeast PPI network.	(Chin et al., 2014)
Netpredictor	2018	R package for prediction of missing links in any given unipartite or bipartite network using random walk restart algorithm.	(Seal and Wild, 2018)
NetEPD	2020	It is a web service integrating diverse databases and computational methods for accurate prediction of essential proteins.	(Zhang et al., 2020)
CaNDis	2021	It is a server for exploring a human causal interaction network, expanded with disease and FDA-approved drug data to construct a disease-disease network.	(Škrlj et al., 2021)
NetControl4BioMed	2021	It generates directed protein-protein interaction networks and used in controllability analysis of target structural network.	(Popescu et al., 2021)
Konnect2prot	2022	Build a functional network with context specificity and identifies important proteins that spread information in the network.	(Kumar et al., 2023a)

TABLE 5 Online drug databases and tools for antiviral discovery.

Name	Year	Description	Reference
(A) Clinically oriented drug databases			
PharmGKB	2000	It serves as an openly accessible online repository dedicated to consolidating, curating, integrating, and disseminating information pertaining to the influence of human genetic variability on drug response.	(Whirl-Carrillo et al., 2021)
RxList	1995	It is an online medical resource dedicated to offering detailed and current pharmaceutical information on brand and generic drugs.	(RxList)
(B) Chemically oriented drug database			
TTD	2014	The Therapeutic Target Database (TTD) systematically collects nine categories of established druggability characteristics for a wide range of targets, aiding in the identification and validation of innovative drug targets. Top of Form	(Zhou et al., 2024)
ChemSpider	2007	This database houses data on a vast array of molecules, exceeding 100 million, aggregated from over 270 distinct sources. Each molecule is assigned a unique identifier, known as the ChemSpider Identifier.	(Pence and Williams, 2010)
PubChem	2004	It has an extensive collection of over 293 million substance descriptions, featuring 111 million distinct chemical structures and encompassing 271 million bioactivity data points derived from approximately 1.2 million biological assay experiments.	(Kim et al., 2023)
ZINC	2004	It hosts a library of more than 230 million readily available compounds in 3D formats for docking purposes. Additionally, it offers access to over 750 million purchasable compounds, facilitating rapid analog searches within a minute.	(Sterling and Irwin, 2015)
(C) Online drug discovery tools			
Drug Central	2016	It is a publicly accessible drug information resource that includes 4950 drugs and related additional data sources such as FDA Adverse Event Reporting System (FAERS) and L1000 gene perturbation profile distance matrices.	(Avram et al., 2023)
ChEBI	2009	This database compiles data on compound bioactivity against drug targets, with bioactivity measurements reported in Ki, Kd, IC50, and EC50 values.	(Degtyarenko et al., 2007)
STITCH	2008	Search Tool for Interacting Chemicals (STITCH) consolidates dispersed data on protein-small molecule interactions into a single, user-friendly resource, now expanded to cover 430,000 chemicals. The latest release introduces a network view showcasing binding affinities, facilitating quick insights into chemical effects on interaction partners.	(Szklarczyk et al., 2016)
DrugBank	2006	The latest version of the database contains 4563 FDA approved drugs and 6231 investigational drugs, along with information on drug-drug and drug-food interactions.	(Wishart et al., 2008)

5.2.3 *In silico* methods for discovery of small molecules: structural and machine learning approach

It is a proven fact that the proteins are dynamic in nature and so targeting a static snapshot may not be a very comprehensive approach. Therefore, multiple computational tools have been developed to virtually decipher the three-dimensional (3D) structure of biological molecules in their functionally active state and analyze the interaction among different biological molecules such as RNA-protein interactions, protein-drug interaction and RNA-drug interaction. The hallmark of computational structural study is its ability to generate and analyze multiple interconverting states by studying its thermodynamic properties. Molecular docking and molecular dynamics (MD) simulation techniques are used to characterize the complexities of RBP-drug interactions. Additional methods such as advanced quantum mechanics/molecular mechanics computation, Martini coarse-grained force field molecular modeling and Elastic network models may be adapted to analyze RNA-protein-drug interactions (Monticelli et al., 2008; Pokorna et al., 2018). Further, quantitative structure-activity relationship (QSAR) modeling have played a significant role in computer-aided designing of drug molecules. Recent advances in high-performance computing (HPC) and artificial intelligence (AI) technologies have propelled the transition of QSAR to deep QSAR (a combination of QSAR, more complex statistics and machine learning techniques), which is more robust in structure-based virtual screening (Gini et al., 2019; Selvaraj et al., 2023). Both clinical and chemically oriented drug databases may be used in this approach.

In silico analysis offers the advantage of simultaneous analysis and optimization of 3D structure of the interaction partners as well as drug molecules, which significantly expedites the optimization process. However, such an approach has inherent limitations such as the requirement of high computing power, lack of knowledge regarding 3D structural details of many biological molecules, inability to integrate biological information, and high false positive rate of molecular dynamic simulation analysis. Some of the above-mentioned limitations have been resolved by developing the computational analysis of novel drug opportunities (CANDO) platform (Minie et al., 2014). CANDO is a model independent approach to drug discovery. It leverages the evolutionary basis of protein and small molecule interactions and also considers the known biological data about interaction partners. Importantly, all the analyses do not require high computational power. CANDO platform has been used to identify several drug candidates against COVID-19 (Mangione et al., 2022).

5.2.4 Experimental validation of drug candidates

After identifying a suitable antiviral target in the RNA-protein interactome data and *in silico* screening of drug libraries shortlists potential antiviral candidates, which may be evaluated through suitable wet lab-based assays. If a known drug molecule is identified, its antiviral potential against the corresponding virus may be directly evaluated. Alternatively, small molecule libraries of NCEs (new chemical entities) or FDA (Federal Drug Agency, USA) approved drugs may be screened against a specific target using appropriately validated assays. Potential drug molecules may be characterized by NMR (nuclear magnetic resonance), X-ray-crystallography and structural mass spectrometry (Britt et al., 2021). Once the antiviral potential of the drug molecule is evaluated in cell-based models and small animal models (if available), its efficacy may be evaluated through subsequent pre-clinical studies and clinical trials.

6 Conclusion

A thorough understanding of the RNA-protein interactions prevalent in the life of +ssRNA virus is fundamental to decoding the life cycle and mechanism of viral pathogenesis, knowledge of which is essential for developing specific antivirals. Recent studies have revealed the RNA-protein interactome of a few +ssRNA viruses such as SARS-CoV-2, Dengue Virus and Zika virus. These studies have demonstrated the power and functional utility of omics-based technologies in interrogating the RNA-protein interactome of +ssRNA viruses and provided convincing proof regarding the value of such technologies in gaining deeper insight into the life cycle and pathogenesis mechanism of +ssRNA viruses. Future research should aim at developing more sophisticated and advanced methods, including kinetic models. Further, considering the dynamic nature of the interaction between the RNA, protein and drug, differential equation-based mathematical models should be useful in characterizing them. Finally, the development of more efficient X-ray-crystallography and structural mass spectrometry methods should help in the antiviral discovery process. By coupling the data identified from RNA-protein interactome analysis with the drug discovery pipeline, it should be possible to develop potent antivirals against pathogenic +ssRNA viruses of medical importance.

Author contributions

SG: Writing – original draft, Writing – review & editing. SK: Writing – original draft, Writing – review & editing. RV: Writing – original draft, Writing – review & editing. SA: Writing – original draft, Writing – review & editing. SC: Writing – original draft, Writing – review & editing, Conceptualization, Resources, Supervision. MS: Conceptualization, Resources, Writing – original draft, Writing – review & editing, Funding acquisition, Supervision.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Agrawal, S., Gupta, D., and Panda, S. K. (2001). The 3′ end of hepatitis E virus (HEV) genome binds specifically to the viral RNA-dependent RNA polymerase (RdRp). *Virology* 282, 87–101. doi: 10.1006/viro.2000.0819

Aleksander, S. A., Balhoff, J., Carbon, S., Cherry, J. M., Drabkin, H. J., Ebert, D., et al. (2023). The gene ontology knowledgebase in 2023. *Genetics* 224, iyad031. doi: 10.1093/genetics/iyad031

Ali, I. K., McKendrick, L., Morley, S. J., and Jackson, R. J. (2001). Activity of the hepatitis A virus IRES requires association between the cap-binding translation initiation factor (eIF4E) and eIF4G. *J. Virol.* 75, 7854–7863. doi: 10.1128/JVI.75.17.7854-7863.2001

Al-Sunaidi, M., Williams, Ç.H., Hughes, P. J., Schnurr, D. P., and Stanway, G. (2007). Analysis of a new human parechovirus allows the definition of parechovirus types and the identification of RNA structural domains. *J. Virol.* 81, 1013–1021. doi: 10.1128/JVI.00584-06

Anwar, A., Leong, K. M., Ng, M. L., Chu, J. J., and Garcia-Blanco, M. A. (2009). The polypyrimidine tract-binding protein is required for efficient dengue virus propagation and associates with the viral replication machinery. *J. Biol. Chem.* 284, 17021–17029. doi: 10.1074/jbc.M109.006239

Armaos, A., Cirillo, D., and Gaetano Tartaglia, G. (2017). omiXcore: a web server for prediction of protein interactions with large RNA. *Bioinformatics* 33, 3104–3106. doi: 10.1093/bioinformatics/btx361

. "RxList - The Internet Drug Index for prescription drug information, interactions, and side effects," in *RxList*. Available at: http://www.rxlist.com/.

Avram, S., Wilson, T. B., Curpan, R., Halip, L., Borota, A., Bora, A., et al. (2023). DrugCentral 2023 extends human clinical data and integrates veterinary drugs. *Nucleic Acids Res.* 51, D1276–D1287. doi: 10.1093/nar/gkac1085

Barton, D. J., O'Donnell, B. J., and Flanegan, J. B. (2001). 5′ cloverleaf in poliovirus RNA is a cis-acting replication element required for negative-strand synthesis. *EMBO J* 20, 1439–1448. doi: 10.1093/emboj/20.6.1439

Benoit Bouvrette, L. P., Wang, X., Boulais, J., Kong, J., Syed, E. U., Blue, S. M., et al. (2023). RBP Image Database: A resource for the systematic characterization of the subcellular distribution properties of human RNA binding proteins. *Nucleic Acids Res.* 51, D1549–D1557. doi: 10.1093/nar/gkac971

Bentaleb, C., Hervouet, K., Montpellier, C., Camuzet, C., Ferrié, M., Burlaud-Gaillard, J., et al. (2022). The endocytic recycling compartment serves as a viral factory for hepatitis E virus. *Cell. Mol. Life Sci.* 79, 615. doi: 10.1007/s00018-022-04646-y

Bidet, K., Dadlani, D., and Garcia-Blanco, M. A. (2017). Correction: G3BP1, G3BP2 and CAPRIN1 are required for translation of interferon stimulated mRNAs and are targeted by a dengue virus non-coding RNA. *PloS Pathog.* 13, e1006295. doi: 10.1371/journal.ppat.1006295

Blyn, L. B., Swiderek, K. M., Richards, O., Stahl, D. C., Semler, B. L., and Ehrenfeld, E. (1996). Poly (rC) binding protein 2 binds to stem-loop IV of the poliovirus RNA 5′noncoding region: identification by automated liquid chromatography-tandem mass spectrometry. *Proc. Natl. Acad. Sci.* 93, 11115–11120. doi: 10.1073/pnas.93.20.11115

Britt, H. M., Cragnolini, T., and Thalassinos, K. (2021). Integration of mass spectrometry data for structural biology. *Chem. Rev.* 122, 7952–7986. doi: 10.1021/acs.chemrev.1c00356

Brown, E. A., Zajac, A. J., and Lemon, S. M. (1994). *In vitro* characterization of an internal ribosomal entry site (IRES) present within the 5′nontranslated region of hepatitis A virus RNA: comparison with the IRES of encephalomyocarditis virus. *J. Virol.* 68, 1066–1074. doi: 10.1128/jvi.68.2.1066-1074.1994

Brunner, J. E., Nguyen, J. H., Roehl, H. H., Ho, T. V., Swiderek, K. M., and Semler, B. L. (2005). Functional interaction of heterogeneous nuclear ribonucleoprotein C with poliovirus RNA synthesis initiation complexes. *J. Virol.* 79, 3254–3266. doi: 10.1128/JVI.79.6.3254-3266.2005

Caudron-Herger, M., Jansen, R. E., Wassmer, E., and Diederichs, S. (2021). RBP2GO: a comprehensive pan-species database on RNA-binding proteins, their interactions and functions. *Nucleic Acids Res.* 49, D425–D436. doi: 10.1093/nar/gkaa1040

Chahar, H. S., Chen, S., and Manjunath, N. (2013). P-body components LSM1, GW182, DDX3, DDX6 and XRN1 are recruited to WNV replication sites and positively regulate viral replication. *Virology* 436, 1–7. doi: 10.1016/j.virol.2012.09.041

Chan, Y. K., and Gack, M. U. (2016). Viral evasion of intracellular DNA and RNA sensing. *Nat. Rev. Microbiol.* 14, 360–373. doi: 10.1038/nrmicro.2016.45

Chatel-Chaix, L., and Bartenschlager, R. (2014). Dengue virus-and hepatitis C virus-induced replication and assembly compartments: the enemy inside—caught in the web. *J. Virol.* 88, 5907–5911. doi: 10.1128/JVI.03404-13

Chen, X., Yang, X., Zheng, Y., Yang, Y., Xing, Y., and Chen, Z. (2014). SARS coronavirus papain-like protease inhibits the type I interferon signaling pathway through interaction with the STING-TRAF3-TBK1 complex. *Protein Cell* 5, 369–381. doi: 10.1007/s13238-014-0026-3

Cheung, P., Lim, T., Yuan, J., Zhang, M., Chau, D., McManus, B., et al. (2007). Specific interaction of HeLa cell proteins with coxsackievirus B3 3′ UTR: La autoantigen binds the 3′ and 5′ UTR independently of the poly (A) tail. *Cell. Microbiol.* 9, 1705–1715. doi: 10.1111/j.1462-5822.2007.00904.x

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8, 1–7. doi: 10.1186/1752-0509-8-S4-S11

Cho, K. F., Branon, T. C., Rajeev, S., Svinkina, T., Udeshi, N. D., Thoudam, T., et al. (2020). Split-TurboID enables contact-dependent proximity labeling in cells. *Proc. Natl. Acad. Sci.* 117, 12143–12154. doi: 10.1073/pnas.1919528117

Choudhury, S. M., Ma, X., Abdullah, S. W., and Zheng, H. (2021). Activation and inhibition of the NLRP3 inflammasome by RNA viruses. *J. Inflammation Res.* 14, 1145–1163. doi: 10.2147/JIR.S295706

Chu, C., Zhang, Q. C., Da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., et al. (2015). Systematic discovery of Xist RNA binding proteins. *Cell* 161, 404–416. doi: 10.1016/j.cell.2015.03.025

Cook, K. B., Kazan, H., Zuberi, K., Morris, Q., and Hughes, T. R. (2010). RBPDB: a database of RNA-binding specificities. *Nucleic Acids Res.* 39, D301–D308. doi: 10.1093/nar/gkq1069

Cordey, S., Gerlach, D., Junier, T., Zdobnov, E. M., Kaiser, L., and Tapparel, C. (2008). The cis-acting replication elements define human enterovirus and rhinovirus species. *Rna* 14, 1568–1578. doi: 10.1261/rna.1031408

Corley, M., Burns, M. C., and Yeo, G. W. (2020). How RNA-binding proteins interact with RNA: molecules and mechanisms. *Mol. Cell* 78, 9–29. doi: 10.1016/j.molcel.2020.03.011

Dave, P., George, B., Sharma, D. K., and Das, S. (2017). Polypyrimidine tract-binding protein (PTB) and PTB-associated splicing factor in CVB3 infection: an ITAF for an ITAF. *Nucleic Acids Res.* 45, 9068–9084. doi: 10.1093/nar/gkx519

Degtyarenko, K., De Matos, P., Ennis, M., Hastings, J., Zbinden, M., McNaught, A., et al. (2007). ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Res.* 36, D344–D350. doi: 10.1093/nar/gkm791

de Los Santos, T., Diaz-San Segundo, F., and Grubman, M. J. (2007). Degradation of nuclear factor kappa B during foot-and-mouth disease virus infection. *J. Virol.* 81, 12803–12815. doi: 10.1128/JVI.01467-07

De Munter, S., Görnemann, J., Derua, R., Lesage, B., Qian, J., Heroes, E., et al. (2017). Split-BioID: a proximity biotinylation assay for dimerization-dependent protein interactions. *FEBS Lett.* 591, 415–424. doi: 10.1002/feb2.2017.591.issue-2

Deng, L., Liu, Y., Shi, Y., Zhang, W., Yang, C., and Liu, H. (2020). Deep neural networks for inferring binding sites of RNA-binding proteins by using distributed representations of RNA primary sequence and secondary structure. *BMC Genomics* 21, 1–10. doi: 10.1186/s12864-020-07239-w

Diosa-Toro, M., Prasanth, K. R., Bradrick, S. S., and Garcia Blanco, M. A. (2020). Role of RNA-binding proteins during the late stages of Flavivirus replication cycle. *Virol. J.* 17, 1–14. doi: 10.1186/s12985-020-01329-7

Dreyfuss, G., Kim, V. N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.* 3, 195–205. doi: 10.1038/nrm760

Du, Y., Bi, J., Liu, J., Liu, X., Wu, X., Jiang, P., et al. (2014). 3Cpro of foot-and-mouth disease virus antagonizes the interferon signaling pathway by blocking STAT1/STAT2 nuclear translocation. *J. Virol.* 88, 4908–4920. doi: 10.1128/JVI.03668-13

Du, Z., Xiao, X., and Uversky, V. N. (2022). DeepA-RBPBS: A hybrid convolution and recurrent neural network combined with attention mechanism for predicting RBP binding site. *J. Biomolecular Structure Dynamics* 40, 4250–4258. doi: 10.1080/07391102.2020.1854861

Ellington, A. D., and Szostak, J. W. (1990). *In vitro* selection of RNA molecules that bind specific ligands. *nature* 346, 818–822. doi: 10.1038/346818a0

Ertel, K. J., Brunner, J. E., and Semler, B. L. (2010). Mechanistic consequences of hnRNP C binding to both RNA termini of poliovirus negative-strand RNA intermediates. *J. Virol.* 84, 4229–4242. doi: 10.1128/JVI.02198-09

Flynn, R. A., Belk, J. A., Qi, Y., Yasumoto, Y., Wei, J., Alfajaro, M. M., et al. (2021). Discovery and functional interrogation of SARS-CoV-2 RNA-host protein interactions. *Cell* 184, 2394–2411. doi: 10.1016/j.cell.2021.03.012

Frieman, M., Ratia, K., Johnston, R. E., Mesecar, A. D., and Baric, R. S. (2009). Severe acute respiratory syndrome coronavirus papain-like protease ubiquitin-like domain and catalytic domain regulate antagonism of IRF3 and NF-κB signaling. *J. Virol.* 83, 6689–6705. doi: 10.1128/JVI.02220-08

Gaither, J., Lin, Y. H., and Bundschuh, R. (2022). RBPBind: quantitative prediction of protein-RNA interactions. *J. Mol. Biol.* 434, 167515. doi: 10.1016/j.jmb.2022.167515

Gamarnik, A. V., and Andino, R. (2000). Interactions of viral protein 3CD and poly (rC) binding protein with the 5′ untranslated region of the poliovirus genome. *J. Virol.* 74, 2219–2226. doi: 10.1128/JVI.74.5.2219-2226.2000

Gerber, K., Wimmer, E., and Paul, A. V. (2001). Biochemical and genetic studies of the initiation of human rhinovirus 2 RNA replication: identification of a cis-replicating element in the coding sequence of 2Apro. *J. Virol.* 75, 10979–10990. doi: 10.1128/JVI.75.22.10979-10990.2001

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. *Nat. Rev. Genet.* 15, 829–845. doi: 10.1038/nrg3813

Ghanbari, M., and Ohler, U. (2020). Deep neural networks for interpreting RNA-binding protein target preferences. *Genome Res.* 30, 214–226. doi: 10.1101/gr.247494.118

Gini, G., Zanoli, F., Gamba, A., Raitano, G., and Benfenati, E. (2019). Could deep learning in neural networks improve the QSAR models? *SAR QSAR Environ. Res.* 30, 617–642. doi: 10.1080/1062936X.2019.1650827

Goodfellow, I., Chaudhry, Y., Richardson, A., Meredith, J., Almond, J. W., Barclay, W., et al. (2000). Identification of a cis-acting replication element within the poliovirus coding region. *J. Virol.* 74, 4590–4600. doi: 10.1128/JVI.74.10.4590-4600.2000

Graindorge, A., Pinheiro, I., Nawrocka, A., Mallory, A. C., Tsvetkov, P., Gil, N., et al. (2019). In-cell identification and measurement of RNA-protein interactions. *Nat. Commun.* 10, 5317. doi: 10.1038/s41467-019-13235-w

Grønning, A. G. B., Doktor, T. K., Larsen, S. J., Petersen, U. S. S., Holm, L. L., Bruun, G. H., et al. (2020). DeepCLIP: predicting the effect of mutations on protein–RNA binding with deep learning. *Nucleic Acids Res.* 48, 7099–7118. doi: 10.1093/nar/gkaa530

Hackbart, M., Deng, X., and Baker, S. C. (2020). Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors. *Proc. Natl. Acad. Sci.* 117, 8094–8103. doi: 10.1073/pnas.1921485117

Han, J. Q., Townsend, H. L., Jha, B. K., Paranjape, J. M., Silverman, R. H., and Barton, D. J. (2007). A phylogenetically conserved RNA structure in the poliovirus open reading frame inhibits the antiviral endoribonuclease RNase L. *J. Virol.* 81, 5561–5572. doi: 10.1128/JVI.01857-06

Han, S., Zhao, B. S., Myers, S. A., Carr, S. A., He, C., and Ting, A. Y. (2020). RNA–protein interaction mapping via MS2-or Cas13-based APEX targeting. *Proc. Natl. Acad. Sci.* 117, 22068–22079. doi: 10.1073/pnas.2006617117

Hato, S. V., Ricour, C., Schulte, B. M., Lanke, K. H., de Bruijni, M., Zoll, J., et al. (2007). The mengovirus leader protein blocks interferon-α/β gene transcription and inhibits activation of interferon regulatory factor 3. *Cell. Microbiol.* 9, 2921–2930. doi: 10.1111/j.1462-5822.2007.01006.x

Horlacher, M., Oleshko, S., Hu, Y., Ghanbari, M., Cantini, G., Schinke, P., et al. (2023). A computational map of the human-SARS-CoV-2 protein–RNA interactome predicted at single-nucleotide resolution. *NAR Genomics Bioinf.* 5, lqad010. doi: 10.1093/nargab/lqad010

Hou, W., Armstrong, N., Obwolo, L. A., Thomas, M., Pang, X., Jones, K. S., et al. (2017). Determination of the cell permissiveness spectrum, mode of RNA replication, and RNA-protein interaction of Zika virus. *BMC Infect. Dis.* 17, 1–12. doi: 10.1186/s12879-017-2338-4

Huang, P., and Lai, M. M. (2001). Heterogeneous nuclear ribonucleoprotein a1 binds to the 3′-untranslated region and mediates potential 5′-3′-end cross talks of mouse hepatitis virus RNA. *J. Virol.* 75, 5009–5017. doi: 10.1128/JVI.75.11.5009-5017.2001

Huang, L., Xiong, T., Yu, H., Zhang, Q., Zhang, K., Li, C., et al. (2017). Encephalomyocarditis virus 3C protease attenuates type I interferon production through disrupting the TANK–TBK1–IKKε–IRF3 complex. *Biochem. J.* 474, 2051–2065. doi: 10.1042/BCJ20161037

Jin, W., Brannan, K. W., Kapeli, K., Park, S. S., Tan, H. Q., Gosztyla, M. L., et al. (2023). HydRA: Deep-learning models for predicting RNA-binding capacity from protein interaction association context and protein sequence. *Mol. Cell* 83, 2595–2611. doi: 10.1016/j.molcel.2023.06.019

Kamel, W., Noerenberg, M., Cerikan, B., Chen, H., Järvelin, A. I., Kammoun, M., et al. (2021a). Global analysis of protein-RNA interactions in SARS-CoV-2-infected cells reveals key regulators of infection. *Mol. Cell* 81, 2851–2867. doi: 10.1016/j.molcel.2021.05.023

Kamel, W., Ruscica, V., Garcia-Moreno, M., Palmalux, N., Iselin, L., Hannan, M., et al. (2021b). Compositional analysis of Sindbis virus ribonucleoproteins reveals an extensive co-opting of key nuclear RNA-binding proteins. *BioRxiv*, 2021–2010. doi: 10.1101/2021.10.06.463336

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., and Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 51, D587–D592. doi: 10.1093/nar/gkac963

Karin, J., Michel, H., and Orenstein, Y. (2021). "BCB: Bioinformatics, Computational Biology and Biomedicine," in *BCB 2021 Proceeding, Association for Computing Machinery*. 1–9 (New York, United States).

Kazachenka, A., Bertozzi, T. M., Sjoberg-Herrera, M. K., Walker, N., Gardner, J., Gunning, R., et al. (2018). Identification, characterization, and heritability of murine metastable epialleles: implications for non-genetic inheritance. *Cell* 175, 1259–1271. doi: 10.1016/j.cell.2018.09.043

Kelly, J. A., Woodside, M. T., and Dinman, J. D. (2021). Programmed– 1 ribosomal frameshifting in coronaviruses: a therapeutic target. *Virology* 554, 75–82. doi: 10.1016/j.virol.2020.12.010

Kerr, C. H., and Jan, E. (2016). Commandeering the ribosome: Lessons learned from dicistroviruses about translation. *J. Virol.* 90, 5538–5540. doi: 10.1128/JVI.00737-15

Khawaja, A., Vopalensky, V., and Pospisek, M. (2015). Understanding the potential of hepatitis C virus internal ribosome entry site domains to modulate translation initiation via their structure and function. *Wiley Interdiscip. Reviews: RNA* 6, 211–224. doi: 10.1002/wrna.2015.6.issue-2

Khromykh, A. A., Meka, H., Guyatt, K. J., and Westaway, E. G. (2001). Essential role of cyclization sequences in flavivirus RNA replication. *J. Virol.* 75, 6719–6728. doi: 10.1128/JVI.75.14.6719-6728.2001

Kim, B., Arcos, S., Rothamel, K., Jian, J., Rose, K. L., McDonald, W. H., et al. (2020). Discovery of widespread host protein interactions with the pre-replicated genome of CHIKV using VIR-CLASP. *Mol. Cell* 78, 624–640. doi: 10.1016/j.molcel.2020.04.013

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2023). PubChem 2023 update. *Nucleic Acids Res.* 51, D1373–D1380. doi: 10.1093/nar/gkac956

Kim, G. W., Imam, H., Khan, M., and 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

Knox, C., Wilson, M., Klinger, C. M., Franklin, M., Oler, E., Wilson, A., et al. (2024). DrugBank 6.0: the drugBank knowledgebase for 2024. *Nucleic Acids Res.* 52, D1265–D1275. doi: 10.1093/nar/gkad976

Koo, P. K., Ploenzke, M., Anand, P., Paul, S., and Majdandzic, A. (2023). "ResidualBind: Uncovering Sequence-Structure Preferences of RNA-Binding Proteins with Deep Neural Networks," in *RNA Structure Prediction* (Springer US, New York, NY), 197–215.

Kretz, M., Siprashvili, Z., Chu, C., Webster, D. E., Zehnder, A., Qu, K., et al. (2013). Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* 493, 231–235. doi: 10.1038/nature11661

Kumar, S., Sarmah, D. T., Asthana, S., and Chatterjee, S. (2023a). konnect2prot: a web application to explore the protein properties in a functional protein–protein interaction network. *Bioinformatics* 39, btac815. doi: 10.1093/bioinformatics/btac815

Kumar, S., Verma, R., Saha, S., Agrahari, A. K., Shukla, S., Singh, O. N., et al. (2023b). RNA-protein interactome at the Hepatitis E virus internal ribosome entry site. *Microbiol. Spectr.* 11, e02827–e02822. doi: 10.1128/spectrum.02827-22

Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., and Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* 47, W171–W174. doi: 10.1093/nar/gkz365

Lang, B., Armaos, A., and Tartaglia, G. G. (2019). RNAct: Protein–RNA interaction predictions for model organisms with supporting experimental data. *Nucleic Acids Res.* 47, D601–D606. doi: 10.1093/nar/gky967

Lawrence, P., and Rieder, E. (2009). Identification of RNA helicase A as a new host factor in the replication cycle of foot-and-mouth disease virus. *J. Virol.* 83, 11356–11366. doi: 10.1128/JVI.02677-08

Lee, K. M., Chen, C. J., and Shih, S. R. (2017). Regulation mechanisms of viral IRES-driven translation. *Trends Microbiol.* 25, 546–561. doi: 10.1016/j.tim.2017.01.010

Lee, Y. H., Hass, E. P., Campodonico, W., Lee, Y. K., Lasda, E., Shah, J. S., et al. (2024). Massively parallel dissection of RNA in RNA–protein interactions in *vivo*. *Nucleic Acids Res.* 52, e48–e48. doi: 10.1093/nar/gkae334

Lee, H. Y., Haurwitz, R. E., Apffel, A., Zhou, K., Smart, B., Wenger, C. D., et al. (2013). RNA–protein analysis using a conditional CRISPR nuclease. *Proc. Natl. Acad. Sci.* 110, 5416–5421. doi: 10.1073/pnas.1302807110

Lei, J., and Hilgenfeld, R. (2017). RNA-virus proteases counteracting host innate immunity. *FEBS Lett.* 591, 3190–3210. doi: 10.1002/feb2.2017.591.issue-20

Lei, X., Liu, X., Ma, Y., Sun, Z., Yang, Y., Jin, Q., et al. (2010). The 3C protein of enterovirus 71 inhibits retinoid acid-inducible gene I-mediated interferon regulatory factor 3 activation and type I interferon responses. *J. Virol.* 84, 8051–8061. doi: 10.1128/JVI.02491-09

Levengood, J. D., Tolbert, M., Li, M. L., and Tolbert, B. S. (2013). High-affinity interaction of hnRNP A1 with conserved RNA structural elements is required for translation and replication of enterovirus 71. *RNA Biol.* 10, 1136–1145. doi: 10.4161/rna.25107

Lewis, B. A., Walia, R. R., Terribilini, M., Ferguson, J., Zheng, C., Honavar, V., et al. (2010). PRIDB: a protein–RNA interface database. *Nucleic Acids Res.* 39, D277–D282. doi: 10.1093/nar/gkq1108

Li, H. P., Huang, P., Park, S., and Lai, M. M. (1999). Polypyrimidine tract-binding protein binds to the leader RNA of mouse hepatitis virus and serves as a regulator of viral transcription. *J. Virol.* 73, 772–777. doi: 10.1128/JVI.73.1.772-777.1999

Li, N., Hui, H., Bray, B., Gonzalez, G. M., Zeller, M., Anderson, K. G., et al. (2021). METTL3 regulates viral m6A RNA modification and host cell innate immune responses during SARS-CoV-2 infection. *Cell Rep.* 35, 1–21. doi: 10.1016/j.celrep.2021.109091

Li, Y., Liu, S., Cao, L., Luo, Y., Du, H., Li, S., et al. (2021). CBRPP: a new RNA-centric method to study RNA–protein interactions. *RNA Biol.* 18, 1608–1621. doi: 10.1080/15476286.2021.1873620

Lin, X., Fonseca, M. A., Corona, R. I., and Lawrenson, K. (2020). *In vivo* discovery of RNA proximal proteins in human cells via proximity-dependent biotinylation. *BioRxiv,* 2020–2002. doi: 10.1101/2020.02.28.970442

Lin, J. Y., Li, M. L., Huang, P. N., Chien, K. Y., Horng, J. T., and Shih, S. R. (2008). Heterogeneous nuclear ribonuclear protein K interacts with the enterovirus 71 5′ untranslated region and participates in virus replication. *J. Gen. Virol.* 89, 2540–2549. doi: 10.1099/vir.0.2008/003673-0

Lind, K., Svedin, E., Domsgen, E., Kapell, S., Laitinen, O. H., Moll, M., et al. (2016). Coxsackievirus counters the host innate immune response by blocking type III interferon expression. *J. Gen. Virol.* 97, 1368–1380. doi: 10.1099/jgv.0.000443

Liu, Y., Li, R., Luo, J., and Zhang, Z. (2022). Inferring RNA-binding protein target preferences using adversarial domain adaptation. *PloS Comput. Biol.* 18, e1009863. doi: 10.1371/journal.pcbi.1009863

Liu, P., Li, L., Millership, J. J., Kang, H., Leibowitz, J. L., and Giedroc, D. P. (2007). A U-turn motif-containing stem–loop in the coronavirus 5′ untranslated region plays a functional role in replication. *Rna* 13, 763–780. doi: 10.1261/rna.261807

Liu, Y., Zhang, Y., Wang, M., Cheng, A., Yang, Q., Wu, Y., et al. (2020). Structures and functions of the 3′ untranslated regions of positive-sense single-stranded RNA viruses infecting humans and animals. *Front. Cell. infection Microbiol.* 10, 453. doi: 10.3389/fcimb.2020.00453

Liu, S., Zhu, J., Jiang, T., Zhong, Y., Tie, Y., Wu, Y., et al. (2015). Identification of lncRNA MEG3 binding protein using MS2-tagged RNA affinity purification and mass spectrometry. *Appl. Biochem. Biotechnol.* 176, 1834–1845. doi: 10.1007/s12010-015-1680-5

Lo, C. Y., Tsai, T. L., Lin, C. N., Lin, C. H., and Wu, H. Y. (2019). Interaction of coronavirus nucleocapsid protein with the 5′-and 3′-ends of the coronavirus genome is involved in genome circularization and negative-strand RNA synthesis. *FEBS J.* 286, 3222–3239. doi: 10.1111/febs.v286.16

Lobert, P. E., Escriou, N., Ruelle, J., and Michiels, T. (1999). A coding RNA sequence acts as a replication signal in cardioviruses. *Proc. Natl. Acad. Sci.* 96, 11560–11565. doi: 10.1073/pnas.96.20.11560

Lunde, B. M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* 8, 479–490. doi: 10.1038/nrm2178

Ma, Z., Moore, R., Xu, X., and Barber, G. N. (2013). DDX24 negatively regulates cytosolic RNA-mediated innate immune signaling. *PloS Pathog.* 9, e1003721. doi: 10.1371/journal.ppat.1003721

Ma, D. Y., and Suthar, M. S. (2015). Mechanisms of innate immune evasion in re-emerging RNA viruses. *Curr. Opin. Virol.* 12, 26–37. doi: 10.1016/j.coviro.2015.02.005

Madhugiri, R., Fricke, M., Marz, M., and Ziebuhr, J. (2016). Coronavirus cis-acting RNA elements. *Adv. Virus Res.* 96, 127–163. doi: 10.1016/bs.aivir.2016.08.007

Mangione, W., Falls, Z., and Samudrala, R. (2022). Optimal COVID-19 therapeutic candidate discovery using the CANDO platform. *Front. Pharmacol.* 13, 970494. doi: 10.3389/fphar.2022.970494

Manley, J. L. (2013). SELEX to identify protein-binding sites on RNA. *Cold Spring Harbor Protoc.* 2013, pdb–prot072934. doi: 10.1101/pdb.prot072934

Manokaran, G., Finol, E., Wang, C., Gunaratne, J., Bahl, J., Ong, E. Z., et al. (2015). Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness. *Science* 350, 217–221. doi: 10.1126/science.aab3369

Mason, P. W., Bezborodova, S. V., and Henry, T. M. (2002). Identification and characterization of a cis-acting replication element (cre) adjacent to the internal ribosome entry site of foot-and-mouth disease virus. *J. Virol.* 76, 9686–9694. doi: 10.1128/JVI.76.19.9686-9694.2002

Matia-González, A. M., Iadevaia, V., and Gerber, A. P. (2017). A versatile tandem RNA isolation procedure to capture *in vivo* formed mRNA-protein complexes. *Methods* 118, 93–100. doi: 10.1016/j.ymeth.2016.10.005

Maticzka, D., Lange, S. J., Costa, F., and Backofen, R. (2014). GraphProt: modeling binding preferences of RNA-binding proteins. *Genome Biol.* 15, 1–18. doi: 10.1186/gb-2014-15-1-r17

McHugh, C. A., Chen, C. K., Chow, A., Surka, C. F., Tran, C., McDonel, P., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521, 232–236. doi: 10.1038/nature14443

Mcknight, K. L., and Lemon, S. M. (1998). The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication. *Rna* 4, 1569–1584. doi: 10.1017/S1355838298981006

Miao, Z., Tidu, A., Eriani, G., and Martin, F. (2021). Secondary structure of the SARS-CoV-2 5′-UTR. *RNA Biol.* 18, 447–456. doi: 10.1080/15476286.2020.1814556

Milacic, M., Beavers, D., Conley, P., Gong, C., Gillespie, M., Griss, J., et al. (2024). The reactome pathway knowledgebase 2024. *Nucleic Acids Res.* 52, D672–D678. doi: 10.1093/nar/gkad1025

Minajigi, A., Froberg, J. E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., et al. (2015). A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* 349, aab2276. doi: 10.1126/science.aab2276

Minie, M., Chopra, G., Sethi, G., Horst, J., White, G., Roy, A., et al. (2014). CANDO and the infinite drug discovery frontier. *Drug Discov. Today* 19, 1353–1363. doi: 10.1016/j.drudis.2014.06.018

Mitra, R., Li, J., Sagendorf, J. M., Jiang, Y., Cohen, A. S., Chiu, T. P., et al. (2024). Geometric deep learning of protein–DNA binding specificity. *Nat. Methods* 21, 1674–1683. doi: 10.1038/s41592-024-02372-w

Monticelli, L., Kandasamy, S. K., Periole, X., Larson, R. G., Tieleman, D. P., and Marrink, S. J. (2008). The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* 4, 819–834. doi: 10.1021/ct700324x

Mukherjee, J., Hermesh, O., Eliscovich, C., Nalpas, N., Franz-Wachtel, M., Maček, B., et al. (2019). β-Actin mRNA interactome mapping by proximity biotinylation. *Proc. Natl. Acad. Sci.* 116, 12863–12872. doi: 10.1073/pnas.1820737116

Nagy, P. D., and Pogany, J. (2012). The dependence of viral RNA replication on co-opted host factors. *Nat. Rev. Microbiol.* 10, 137–149. doi: 10.1038/nrmicro2692

Nicholson, B. L., and White, K. A. (2014). Functional long-range RNA–RNA interactions in positive-strand RNA viruses. *Nat. Rev. Microbiol.* 12, 493–504. doi: 10.1038/nrmicro3288

Paek, J., Kalocsay, M., Staus, D. P., Wingler, L., Pascolutti, R., Paulo, J. A., et al. (2017). Multidimensional tracking of GPCR signaling via peroxidase-catalyzed proximity labeling. *Cell* 169, 338–349. doi: 10.1016/j.cell.2017.03.028

Pan, X., Fang, Y., Li, X., Yang, Y., and Shen, H. B. (2020). RBPsuite: RNA-protein binding sites prediction suite based on deep learning. *BMC Genomics* 21, 1–8. doi: 10.1186/s12864-020-07291-6

Papon, L., Oteiza, A., Imaizumi, T., Kato, H., Brocchi, E., Lawson, T. G., et al. (2009). The viral RNA recognition sensor RIG-I is degraded during encephalomyocarditis virus (EMCV) infection. *Virology* 393, 311–318. doi: 10.1016/j.virol.2009.08.009

Patiyal, S., Dhall, A., Bajaj, K., Sahu, H., and Raghava, G. P. (2023). Prediction of RNA-interacting residues in a protein using CNN and evolutionary profile. *Briefings Bioinf.* 24, bbac538. doi: 10.1093/bib/bbac538

Paul, A. V., Rieder, E., Kim, D. W., van Boom, J. H., and Wimmer, E. (2000). Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the *in vitro* uridylylation of VPg. *J. Virol.* 74, 10359–10370. doi: 10.1128/JVI.74.22.10359-10370.2000

Paz, I., Argoetti, A., Cohen, N., Even, N., and Mandel-Gutfreund, Y. (2022). RBPmap: a tool for mapping and predicting the binding sites of RNA-binding proteins considering the motif environment. *Post-Transcriptional Gene Regul.* 2404, 53–65. doi: 10.1007/978-1-0716-1851-6_3. Available at: https://www.springer.com/series/7651

Paz, I., Kligun, E., Bengad, B., and Mandel-Gutfreund, Y. (2016). BindUP: a web server for non-homology-based prediction of DNA and RNA binding proteins. *Nucleic Acids Res.* 44, W568–W574. doi: 10.1093/nar/gkw454

Pence, H. E., and Williams, A. (2010). ChemSpider: an online chemical information resource. *J. Chemical Education* 87(11). doi: 10.1021/ed100697w

Pestova, T. V., and Hellen, C. U. (2003). Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes Dev.* 17, 181–186. doi: 10.1101/gad.1040803

Pokorna, P., Kruse, H., Krepl, M., and Sponer, J. (2018). QM/MM calculations on protein–RNA complexes: Understanding limitations of classical MD simulations and search for reliable cost-effective QM methods. *J. Chem. Theory Comput.* 14, 5419–5433. doi: 10.1021/acs.jctc.8b00670

Polishchuk, M., Paz, I., Yakhini, Z., and Mandel-Gutfreund, Y. (2018). SMARTIV: combined sequence and structure *de-novo* motif discovery for *in-vivo* RNA binding data. *Nucleic Acids Res.* 46, W221–W228. doi: 10.1093/nar/gky453

Popescu, V. B., Sánchez-Martín, J.Á., Schacherer, D., Safadoust, S., Majidi, N., Andronescu, A., et al. (2021). NetControl4BioMed: a web-based platform for controllability analysis of protein–protein interaction networks. *Bioinformatics* 37, 3976–3978. doi: 10.1093/bioinformatics/btab570

Queiroz, R. M., Smith, T., Villanueva, E., Marti-Solano, M., Monti, M., Pizzinga, M., et al. (2019). Comprehensive identification of RNA–protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nat. Biotechnol.* 37, 169–178. doi: 10.1038/s41587-018-0001-2

Ramanathan, M., Majzoub, K., Rao, D. S., Neela, P. H., Zarnegar, B. J., Mondal, S., et al. (2018). RNA–protein interaction detection in living cells. *Nat. Methods* 15, 207–212. doi: 10.1038/nmeth.4601

Rangan, R., Zheludev, I. N., Hagey, R. J., Pham, E. A., Wayment-Steele, H. K., Glenn, J. S., et al. (2020). RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: a first look. *Rna* 26, 937–959. doi: 10.1261/rna.076141.120

Rhee, H. W., Zou, P., Udeshi, N. D., Martell, J. D., Mootha, V. K., Carr, S. A., et al. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* 339, 1328–1331. doi: 10.1126/science.1230593

Robinson, M., Schor, S., Barouch-Bentov, R., and Einav, S. (2018). Viral journeys on the intracellular highways. *Cell. Mol. Life Sci.* 75, 3693–3714. doi: 10.1007/s00018-018-2882-0

Rodríguez Pulido, M., Sánchez-Aparicio, M. T., Martínez-Salas, E., García-Sastre, A., Sobrino, F., and Sáiz, M. (2018). Innate immune sensor LGP2 is cleaved by the Leader protease of foot-and-mouth disease virus. *PloS Pathog.* 14, e1007135. doi: 10.1371/journal.ppat.1007135

Samavarchi-Tehrani, P., Samson, R., and Gingras, A. C. (2020). Proximity dependent biotinylation: key enzymes and adaptation to proteomics approaches. *Mol. Cell. Proteomics* 19, 757–773. doi: 10.1074/mcp.R120.001941

Schmidt, N., Lareau, C. A., Keshishian, H., Ganskih, S., Schneider, C., Hennig, T., et al. (2021). The SARS-CoV-2 RNA–protein interactome in infected human cells. *Nat. Microbiol.* 6, 339–353. doi: 10.1038/s41564-020-00846-z

Schopp, I. M., Amaya Ramirez, C. C., Debeljak, J., Kreibich, E., Skribbe, M., Wild, K., et al. (2017). Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. *Nat. Commun.* 8, 15690. doi: 10.1038/ncomms15690

Seal, A., and Wild, D. J. (2018). Netpredictor: R and Shiny package to perform drug-target network analysis and prediction of missing links. *BMC Bioinf.* 19, 1–10. doi: 10.1186/s12859-018-2254-7

Selvaraj, C., Elakkiya, E., Prabhu, P., Velmurugan, D., and Singh, S. K. (2023). "Advances in QSAR through artificial intelligence and machine learning methods," in *QSAR in Safety Evaluation and Risk Assessment* (USA: Academic Press), 101–116.

SenGupta, D. J., Zhang, B., Kraemer, B., Pochart, P., Fields, S., and Wickens, M. (1996). A three-hybrid system to detect RNA–protein interactions in *vivo. Proc. Natl. Acad. Sci.* 93, 8496–8501. doi: 10.1073/pnas.93.16.8496

Sharma, N. K., Gupta, S., Kumar, A., Kumar, P., Pradhan, U. K., and Shankar, R. (2021). RBPSpot: Learning on appropriate contextual information for RBP binding sites discovery. *Iscience* 24, 1–32. doi: 10.1016/j.isci.2021.103381

Shekhawat, S. S., and Ghosh, I. (2011). Split-protein systems: beyond binary protein–protein interactions. *Curr. Opin. Chem. Biol.* 15, 789–797. doi: 10.1016/j.cbpa.2011.10.014

Shen, Z., Deng, S. P., and Huang, D. S. (2019). RNA-protein binding sites prediction via multi scale convolutional gated recurrent unit networks. *IEEE/ACM Trans. Comput. Biol. Bioinf.* 17, 1741–1750. doi: 10.1109/TCBB.2019.2910513

Škrlj, B., Eržen, N., Lavrač, N., Kunej, T., and Konc, J. (2021). CaNDis: a web server for investigation of causal relationships between diseases, drugs and drug targets. *Bioinformatics* 37, 885–887. doi: 10.1093/bioinformatics/btaa762

Sola, I., Almazán, F., Zúñiga, S., and Enjuanes, L. (2015). Continuous and discontinuous RNA synthesis in coronaviruses. *Annu. Rev. Virol.* 2, 265–288. doi: 10.1146/annurev-virology-100114-055218

Sola, I., Mateos-Gomez, P. A., Almazan, F., Zuniga, S., and Enjuanes, L. (2011). RNA-RNA and RNA-protein interactions in coronavirus replication and transcription. *RNA Biol.* 8, 237–248. doi: 10.4161/rna.8.2.14991

Sooryanarain, H., Heffron, C. L., and Meng, X. J. (2020). The U-rich untranslated region of the hepatitis E virus induces differential type I and type III interferon responses in a host cell-dependent manner. *MBio* 11, 10–1128. doi: 10.1128/mBio.03103-19

Sorokin, I., Vassilenko, K. S., Terenin, I. M., Kalinina, N. O., Agol, V. I., and Dmitriev, S. E. (2021). Non-canonical translation initiation mechanisms employed by eukaryotic viral mRNAs. *Biochem. (Moscow)* 86, 1060–1094. doi: 10.1134/S0006297921090042

Srisawat, C., and Engelke, D. R. (2001). Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. *Rna* 7, 632–641. doi: 10.1017/S135583820100245X

Stavrou, S., Feng, Z., Lemon, S. M., and Roos, R. P. (2010). Different strains of Theiler's murine encephalomyelitis virus antagonize different sites in the type I interferon pathway. *J. Virol.* 84, 9181–9189. doi: 10.1128/JVI.00603-10

Sterling, T., and Irwin, J. J. (2015). ZINC 15–ligand discovery for everyone. *J. Chem. Inf. modeling* 55, 2324–2337. doi: 10.1021/acs.jcim.5b00559

Stern-Ginossar, N., Thompson, S. R., Mathews, M. B., and Mohr, I. (2019). Translational control in virus-infected cells. *Cold Spring Harbor Perspect. Biol.* 11, a033001. doi: 10.1101/cshperspect.a033001

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* 102, 15545–15550. doi: 10.1073/pnas.0506580102

Surjit, M., Jameel, S., and Lal, S. K. (2004). The ORF2 protein of hepatitis E virus binds the 5′ region of viral RNA. *J. Virol.* 78, 320–328. doi: 10.1128/JVI.78.1.320-328.2004

Szklarczyk, D., Santos, A., Von Mering, C., Jensen, L. J., Bork, P., and Kuhn, M. (2016). STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Res.* 44, D380–D384. doi: 10.1093/nar/gkv1277

Tahir, M., Tayara, H., Hayat, M., and Chong, K. T. (2021). kDeepBind: prediction of RNA-Proteins binding sites using convolution neural network and k-gram features. *Chemometrics Intelligent Lab. Syst.* 208, 104217. doi: 10.1016/j.chemolab.2020.104217

Tavares, R. D. C. A., Mahadeshwar, G., Wan, H., Huston, N. C., and Pyle, A. M. (2021). The global and local distribution of RNA structure throughout the SARS-CoV-2 genome. *J. Virol.* 95, 10–1128. doi: 10.1128/JVI.02190-20

Teng, X., Chen, X., Xue, H., Tang, Y., Zhang, P., Kang, Q., et al. (2020). NPInter v4. 0: an integrated database of ncRNA interactions. *Nucleic Acids Res.* 48, D160–D165. doi: 10.1093/nar/gkz969

Thomas, P. D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L. P., and Mi, H. (2022). PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* 31, 8–22. doi: 10.1002/pro.v31.1

Tian, Y., Han, X., and Tian, D. L. (2012). The biological regulation of ABCE1. *IUBMB Life* 64, 795–800. doi: 10.1002/iub.v64.10

Trendel, J., Schwarzl, T., Horos, R., Prakash, A., Bateman, A., Hentze, M. W., et al. (2019). The human RNA-binding proteome and its dynamics during translational arrest. *Cell* 176, 391–403. doi: 10.1016/j.cell.2018.11.004

Tsai, W. C., and Lloyd, R. E. (2014). Cytoplasmic RNA granules and viral infection. *Annu. Rev. Virol.* 1, 147–170. doi: 10.1146/annurev-virology-031413-085505

Tsai, B. P., Wang, X., Huang, L., and Waterman, M. L. (2011). Quantitative profiling of *in vivo*-assembled RNA-protein complexes using a novel integrated proteomic approach. *Mol. Cell. Proteomics* 10, 1–15. doi: 10.1074/mcp.M110.007385

Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *science* 249, 505–510. doi: 10.1126/science.2200121

Uhl, M., Tran, V. D., Heyl, F., and Backofen, R. (2021). RNAProt: an efficient and feature-rich RNA binding protein binding site predictor. *GigaScience* 10, giab054. doi: 10.1093/gigascience/giab054

Urdaneta, E. C., and Beckmann, B. M. (2020). Fast and unbiased purification of RNA-protein complexes after UV cross-linking. *Methods* 178, 72–82. doi: 10.1016/j.ymeth.2019.09.013

Van Nostrand, E. L., Freese, P., Pratt, G. A., Wang, X., Wei, X., Xiao, R., et al. (2020). A large-scale binding and functional map of human RNA-binding proteins. *Nature* 583, 711–719. doi: 10.1038/s41586-020-2077-3

Verma, R., Saha, S., Kumar, S., Mani, S., Maiti, T. K., and Surjit, M. (2021). RNA-protein interaction analysis of SARS-CoV-2 5′ and 3′ untranslated regions reveals a role of lysosome-associated membrane protein-2a during viral infection. *Msystems* 6, 10–1128. doi: 10.1128/msystems.00643-21

Verschueren, E., Von Dollen, J., Cimermancic, P., Gulbahce, N., Sali, A., and Krogan, N. J. (2015). Scoring large-scale affinity purification mass spectrometry datasets with MiST. *Curr. Protoc. Bioinf.* 49, 8–19. doi: 10.1002/0471250953.2015.49.issue-1

Viktorovskaya, O. V., Greco, T. M., Cristea, I. M., and Thompson, S. R. (2016). Identification of RNA binding proteins associated with dengue virus RNA in infected cells reveals temporally distinct host factor requirements. *PloS neglected Trop. Dis.* 10, e0004921. doi: 10.1371/journal.pntd.0004921

Villordo, S. M., and Gamarnik, A. V. (2009). Genome cyclization as strategy for flavivirus RNA replication. *Virus Res.* 139, 230–239. doi: 10.1016/j.virusres.2008.07.016

Walsh, D., and Mohr, I. (2011). Viral subversion of the host protein synthesis machinery. *Nat. Rev. Microbiol.* 9, 860–875. doi: 10.1038/nrmicro2655

Wang, L. C., Chen, S. O., Chang, S. P., Lee, Y. P., Yu, C. K., Chen, C. L., et al. (2015). Enterovirus 71 proteins 2A and 3D antagonize the antiviral activity of gamma interferon via signaling attenuation. *J. Virol.* 89, 7028–7037. doi: 10.1128/JVI.00205-15

Wang, D., Fang, L., Li, P., Sun, L., Fan, J., Zhang, Q., et al. (2011). The leader proteinase of foot-and-mouth disease virus negatively regulates the type I interferon pathway by acting as a viral deubiquitinase. *J. Virol.* 85, 3758–3766. doi: 10.1128/JVI.02589-10

Wang, D., Fang, L., Li, K., Zhong, H., Fan, J., Ouyang, C., et al. (2012). Foot-and-mouth disease virus 3C protease cleaves NEMO to impair innate immune signaling. *J. Virol.* 86, 9311–9322. doi: 10.1128/JVI.00722-12

Wang, D., Fang, L., Luo, R., Ye, R., Fang, Y., Xie, L., et al. (2010). Foot-and-mouth disease virus leader proteinase inhibits dsRNA-induced type I interferon transcription by decreasing interferon regulatory factor 3/7 in protein levels. *Biochem. Biophys. Res. Commun.* 399, 72–78. doi: 10.1016/j.bbrc.2010.07.044

Wang, D., Fang, L., Shi, Y., Zhang, H., Gao, L., Peng, G., et al. (2016). Porcine epidemic diarrhea virus 3C-like protease regulates its interferon antagonism by cleaving NEMO. *J. Virol.* 90, 2090–2101. doi: 10.1128/JVI.02514-15

Wang, H., and Zhao, Y. (2020). RBinds: a user-friendly server for RNA binding site prediction. *Comput. Struct. Biotechnol. J.* 18, 3762–3765. doi: 10.1016/j.csbj.2020.10.043

Ward, A. M., Bidet, K., Yinglin, A., Ler, S. G., Hogue, K., Blackstock, W., et al. (2011). Quantitative mass spectrometry of DENV-2 RNA-interacting proteins reveals that the DEAD-box RNA helicase DDX6 binds the DB1 and DB2 3'UTR structures. *RNA Biol.* 8, 1173–1186. doi: 10.4161/rna.8.6.17836

Wessels, H. H., Méndez-Mancilla, A., Guo, X., Legut, M., Daniloski, Z., and Sanjana, N. E. (2020). Massively parallel Cas13 screens reveal principles for guide RNA design. *Nat. Biotechnol.* 38, 722–727. doi: 10.1038/s41587-020-0456-9

West, J. A., Davis, C. P., Sunwoo, H., Simon, M. D., Sadreyev, R. I., Wang, P. I., et al. (2014). The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell* 55, 791–802. doi: 10.1016/j.molcel.2014.07.012

Whirl-Carrillo, M., Huddart, R., Gong, L., Sangkuhl, K., Thorn, C. F., Whaley, R., et al. (2021). An evidence-based framework for evaluating pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* 110, 563–572. doi: 10.1002/cpt.2350

White, J. P., and Lloyd, R. E. (2012). Regulation of stress granules in virus systems. *Trends Microbiol.* 20, 175–183. doi: 10.1016/j.tim.2012.02.001

Wilson, J. E., Pestova, T. V., Hellen, C. U., and Sarnow, P. (2000). Initiation of protein synthesis from the A site of the ribosome. *Cell* 102, 511–520. doi: 10.1016/S0092-8674(00)00055-6

Wishart, D. S., Knox, C., Guo, A. C., Cheng, D., Shrivastava, S., Tzur, D., et al. (2008). DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* 36, D901–D906. doi: 10.1093/nar/gkm958

Wu, S. Y., Chen, Y. L., Lee, Y. R., Lin, C. F., Lan, S. H., Lan, K. Y., et al. (2021). The autophagosomes containing dengue virus proteins and full-length genomic RNA are infectious. *Viruses* 13, 2034. doi: 10.3390/v13102034

Wu, B., Pogany, J., Na, H., Nicholson, B. L., Nagy, P. D., and White, K. A. (2009). A discontinuous RNA platform mediates RNA virus replication: building an integrated model for RNA–based regulation of viral processes. *PloS Pathog.* 5, e1000323. doi: 10.1371/journal.ppat.1000323

Xu, J., Zhu, J., Huang, W., Xu, K., Yang, R., Zhang, Q. C., et al. (2023). PrismNet: predicting protein–RNA interaction using *in vivo* RNA structural information. *Nucleic Acids Res.* 51, W468–W477. doi: 10.1093/nar/gkad353

Yamada, K., and Hamada, M. (2022). Prediction of RNA–protein interactions using a nucleotide language model. *Bioinf. Adv.* 2, vbac023. doi: 10.1093/bioadv/vbac023

Yan, Z., Hamilton, W. L., and Blanchette, M. (2020). Graph neural representational learning of RNA secondary structures for predicting RNA-protein interactions. *Bioinformatics* 36, i276–i284. doi: 10.1093/bioinformatics/btaa456

Yang, H., Deng, Z., Pan, X., Shen, H. B., Choi, K. S., Wang, L., et al. (2021). RNA-binding protein recognition based on multi-view deep feature and multi-label learning. *Briefings Bioinf.* 22, bbaa174. doi: 10.1093/bib/bbaa174

Yang, Y. C. T., Di, C., Hu, B., Zhou, M., Liu, Y., Song, N., et al. (2015). CLIPdb: a CLIP-seq database for protein-RNA interactions. *BMC Genomics* 16, 1–8. doi: 10.1186/s12864-015-1273-2

Yang, D., and Leibowitz, J. L. (2015). The structure and functions of coronavirus genomic 3' and 5' ends. *Virus Res.* 206, 120–133. doi: 10.1016/j.virusres.2015.02.025

Ye, X., and Jankowsky, E. (2020). High throughput approaches to study RNA-protein interactions *in vitro*. *Methods* 178, 3–10. doi: 10.1016/j.ymeth.2019.09.006

Yi, W., Li, J., Zhu, X., Wang, X., Fan, L., Sun, W., et al. (2020). CRISPR-assisted detection of RNA–protein interactions in living cells. *Nat. Methods* 17, 685–688. doi: 10.1038/s41592-020-0866-0

Yoon, J. H., Srikantan, S., and Gorospe, M. (2012). MS2-TRAP (MS2-tagged RNA affinity purification): tagging RNA to identify associated miRNAs. *Methods* 58, 81–87. doi: 10.1016/j.ymeth.2012.07.004

Yuan, H., Kshirsagar, M., Zamparo, L., Lu, Y., and Leslie, C. S. (2019). BindSpace decodes transcription factor binding signals by large-scale sequence embedding. *Nat. Methods* 16, 858–861. doi: 10.1038/s41592-019-0511-y

Zdrazil, B., Felix, E., Hunter, F., Manners, E. J., Blackshaw, J., Corbett, S., et al. (2024). The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res.* 52, D1180–D1192. doi: 10.1093/nar/gkad1004

Zeng, F., Peritz, T., Kannanayakal, T. J., Kilk, K., Eiriksdottir, E., Langel, U., et al. (2006). A protocol for PAIR: PNA-assisted identification of RNA binding proteins in living cells. *Nat. Protoc.* 1, 920–927. doi: 10.1038/nprot.2006.81

Zhang, S., Huang, W., Ren, L., Ju, X., Gong, M., Rao, J., et al. (2022). Comparison of viral RNA–host protein interactomes across pathogenic RNA viruses informs rapid antiviral drug discovery for SARS-CoV-2. *Cell Res.* 32, 9–23. doi: 10.1038/s41422-021-00581-y

Zhang, J., Li, W., Zeng, M., Meng, X., Kurgan, L., Wu, F. X., et al. (2020). NetEPD: a network-based essential protein discovery platform. *Tsinghua Sci. Technol.* 25, 542–552. doi: 10.1109/TST.5971803

Zhang, J., Liu, B., Wang, Z., Lehnert, K., and Gahegan, M. (2022). DeepPN: a deep parallel neural network based on convolutional neural network and graph convolutional network for predicting RNA-protein binding sites. *BMC Bioinf.* 23, 257. doi: 10.1186/s12859-022-04798-5

Zhang, Z., Sun, W., Shi, T., Lu, P., Zhuang, M., and Liu, J. L. (2020). Capturing RNA–protein interaction via CRUIS. *Nucleic Acids Res.* 48, e52–e52. doi: 10.1093/nar/gkaa143

Zhang, X., Wu, D., Chen, L., Li, X., Yang, J., Fan, D., et al. (2014). RAID: a comprehensive resource for human RNA-associated (RNA–RNA/RNA–protein) interaction. *Rna* 20, 989–993. doi: 10.1261/rna.044776.114

Zhang, K., Zheludev, I. N., Hagey, R. J., Haslecker, R., Hou, Y. J., Kretsch, R., et al. (2021). Cryo-EM and antisense targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA genome. *Nat. Struct. Mol. Biol.* 28, 747–754. doi: 10.1038/s41594-021-00653-y

Zhao, S., and Hamada, M. (2021). Multi-resBind: a residual network-based multi-label classifier for *in vivo* RNA binding prediction and preference visualization. *BMC Bioinf.* 22, 1–15. doi: 10.1186/s12859-021-04430-y

Zheng, X., Cho, S., Moon, H., Loh, T. J., Jang, H. N., and Shen, H. (2016). Detecting RNA–protein interaction using end-labeled biotinylated RNA oligonucleotides and immunoblotting. *RNA-Protein Complexes Interactions: Methods Protoc.* 1421, 35–44. doi: 10.1007/978-1-4939-3591-8_4. Available at: https://www.springer.com/series/7651.

Zhou, Y., Zhang, Y., Zhao, D., Yu, X., Shen, X., Zhou, Y., et al. (2024). TTD: Therapeutic Target Database describing target druggability information. *Nucleic Acids Res.* 52, D1465–D1477. doi: 10.1093/nar/gkad751

Zhu, X., Fang, L., Wang, D., Yang, Y., Chen, J., Ye, X., et al. (2017a). Porcine deltacoronavirus nsp5 inhibits interferon-β production through the cleavage of NEMO. *Virology* 502, 33–38. doi: 10.1016/j.virol.2016.12.005

Zhu, Z., Li, C., Du, X., Wang, G., Cao, W., Yang, F., et al. (2017). Foot-and-mouth disease virus infection inhibits LGP2 protein expression to exaggerate inflammatory response and promote viral replication. *Cell Death Dis.* 8, e2747–e2747. doi: 10.1038/cddis.2017.170

Zhu, X., Wang, D., Zhou, J., Pan, T., Chen, J., Yang, Y., et al. (2017b). Porcine deltacoronavirus nsp5 antagonizes type I interferon signaling by cleaving STAT2. *J. Virol.* 91, 10–1128. doi: 10.1128/JVI.00003-17

Ziv, O., Price, J., Shalamova, L., Kamenova, T., Goodfellow, I., Weber, F., et al. (2020). The short-and long-range RNA-RNA interactome of SARS-CoV-2. *Mol. Cell* 80, 1067–1077. doi: 10.1016/j.molcel.2020.11.004

Zoll, J., Melchers, W. J., Galama, J. M., and van Kuppeveld, F. J. (2002). The mengovirus leader protein suppresses alpha/beta interferon production by inhibition of the iron/ferritin-mediated activation of NF-κB. *J. Virol.* 76, 9664–9672. doi: 10.1128/JVI.76.19.9664-9672.2002