Battle for supremacy: nucleic acid interactions between viruses and cells

Elizabeth J. Hennessy, Garret A. FitzGerald

Review

Since the COVID-19 pandemic swept across the globe, researchers have been trying to understand its origin, life cycle, and pathogenesis. There is a striking variability in the phenotypic response to infection with SARS-CoV-2 that may reflect differences in host genetics and/or immune response. It is known that the human epigenome is influenced by ethnicity, age, lifestyle, and environmental factors, including previous viral infections. This Review examines the influence of viruses on the host epigenome. We describe general lessons and methodologies that can be used to understand how the virus evades the host immune response. We consider how variation in the epigenome may contribute to heterogeneity in the response to SARS-CoV-2 and may identify a precision medicine approach to treatment.

Find the latest version:

https://jci.me/144227/pdf

Battle for supremacy: nucleic acid interactions between viruses and cells

Elizabeth J. Hennessy and Garret A. FitzGerald

Institute for Translational Medicine and Therapeutics (ITMAT), University of Pennsylvania, Perelman School of Medicine, Philadelphia, Pennsylvania, USA.

Since the COVID-19 pandemic swept across the globe, researchers have been trying to understand its origin, life cycle, and pathogenesis. There is a striking variability in the phenotypic response to infection with SARS-CoV-2 that may reflect differences in host genetics and/or immune response. It is known that the human epigenome is influenced by ethnicity, age, lifestyle, and environmental factors, including previous viral infections. This Review examines the influence of viruses on the host epigenome. We describe general lessons and methodologies that can be used to understand how the virus evades the host immune response. We consider how variation in the epigenome may contribute to heterogeneity in the response to SARS-CoV-2 and may identify a precision medicine approach to treatment.

Introduction

There are millions of nucleotide interactions within a cell, and their alteration by single-nucleotide polymorphisms (SNPs) can result in changes to pathways implicated in disease, such as the response to viral infection. Most SNPs identified through GWAS are non-functional, having no known effect on a phenotype, but functional SNPs tend to be located in areas of the genome that do not translate into proteins, such as 3′- and 5′-untranslated regions (UTRs), introns, and intergenic regions, hotspots for regulatory elements like enhancers and long noncoding RNAs (lncRNAs) (1, 2). SNPs in these sequences can modify promoter methylation and transcription factor binding. The interplay between chromatin, RNA, and transcription factors is tightly regulated, and SNPs altering these interactions can influence host response to virus and how the virus interacts with host components. Insight into regulation of nucleic acid interactions within cells and between those cells and microbes will afford the opportunity to control them using chemical interventions. These include inhibitors of DNA methyltransferase (DNMT), histone-modifying enzymes, and the use of CRISPR/Cas9 to correct SNPs that result in greater susceptibility to viral infections. In this Review, we describe known mechanisms by which viruses interact with the host genome, including the potential involvement of SNPs and host RNA molecules, leading to differential effects on individual hosts. We also discuss techniques available for elucidating novel virus/host genome interactions and other insights that may inform the study of SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) and development of therapeutics.

Variability in the response to COVID-19

SARS-CoV-2 emerged in 2019 and is the causative agent of the COVID-19 pandemic. The virus spreads mainly through respira-tory droplets and causes a wide range of symptoms, from asymptomatic infection to severe disease with respiratory, renal, and cardiac failure (3). Several studies of people in the United States and China infected with SARS-CoV-2 have reported risk factors for severity of disease including increased age, assignment of male gender at birth, and comorbidities like obesity, hypertension, and type 2 diabetes (4–8). Public Health England reported that people from Black, Asian, and minority ethnic groups had a higher risk of death from COVID-19 than people from a White British background. This analysis was adjusted for sex, age, socioeconomic deprivation, and region but not for comorbidities (9). Two recent studies demonstrated that variability in the type I IFN response to SARS-CoV-2 is caused by gene mutations or varying levels of neutralizing autoantibodies against type I IFNs that lead to their inactivation, with decreased levels of type I IFNs cosegregating with severe phenotype (10, 11). Variability has also been found in the lymphocyte immune response among patients with severe COVID-19 with varying B and T cell phenotypic responses to infection (12, 13). Not all recovered patients have detectable neutralizing antibodies, suggesting a complex relationship between the humoral and cellular responses to COVID-19.

Clearly, host genetics may play a role, and in this respect, there has been interest in SNPs identified in the *ACE2* (angiotensin-converting enzyme 2) gene encoding the host receptor for the SARS-CoV-2 spike protein responsible for entry into the cell (14). SNPs have been identified that affect the expression of ACE2 (rs112171234 and rs75979613) and correlate with hypertension (rs2285666) (15, 16). ACE2 is located on the X chromosome, and females are subject to random X chromosome inactivation. Such a mechanism might regulate viral load among females and also contribute to the preponderance of males with severe disease (17–19). Our understanding of the variability in response to SARS-CoV-2 is at an early stage but would afford important predictive and potentially therapeutic information.

Just as there is variability in the host immune response due to SNPs, there is variability in the impact of viruses on the host. Viruses have evolved ways to hijack cellular processes to evade
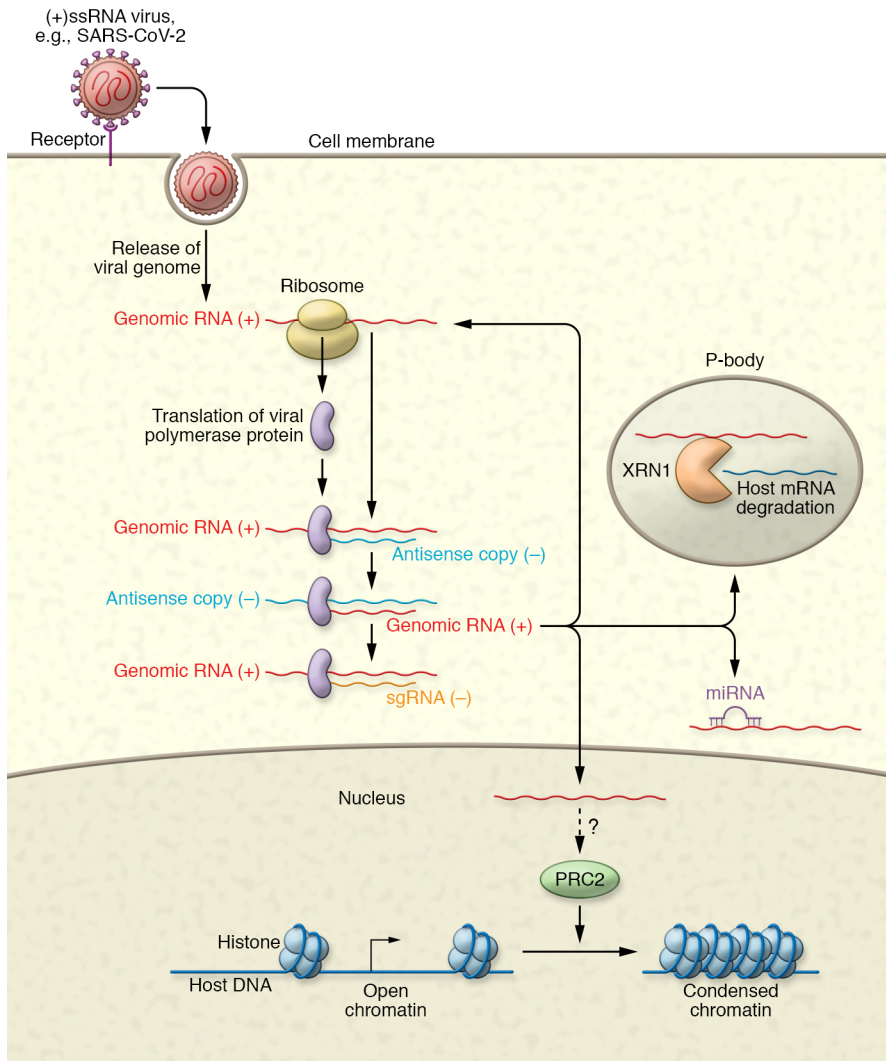
Figure 1. Points of interaction between viral RNA and host RNA factors. When a host cell is infected with a (+)ssRNA virus, both the genomic RNA and subgenomic RNAs produced during RNA replication can interact with endogenous host factors such as miRNAs in the cytoplasm, p-bodies, nuclear factors like PRC2, and tRNAs involved in translation of viral proteins.

(dsRNA), positive-sense single-stranded RNA [(+)ssRNA], or negative-sense single-stranded RNA [(–)ssRNA]. When a (+)ssRNA virus, like SARS-CoV-2, infects a cell, its RNA is released into the cytoplasm, where it can be translated into viral proteins using the host ribosomal machinery without the need to be reverse-transcribed in cDNA. In contrast, (–)ssRNA viruses must first have their genome copied to form positive-sense RNA by the RNA-dependent RNA polymerase (RdRp) protein. Some viral elements can transit through the nuclear envelope via the nuclear pore complex. Several (+)ssRNA viruses have been found in the nucleus and nucleolus of host cells, where they use proteins, such as the capsid protein, to interfere with RNA-binding proteins and disrupt nuclear architecture and the cell cycle or inhibit transcription (24). Viruses have evolved a multitude of mechanisms for evading and exploiting the host immune response for their benefit. It is becoming evident that variability of both the host response due to genomic elements like SNPs and the mechanisms of viral action is contributing to the range of responses to SARS-CoV-2.

Connecting SNPs to the epigenome

Variability in drug response is influenced by variation in both the genome and the environment through its influence on the epigenome, which consists of chemical modifications to histone proteins and DNA that regulate gene expression (25). The use of biobanks, such as the UK Biobank (26, 27), can aid in understanding the functional importance of variability in the human genome and epigenome. Biobanks differ in diversity, the right of recall, depth of phenotyping, follow-up, and integration with experimental medicine (28). Expression quantitative trait loci (eQTLs) are genomic loci containing SNPs that explain some of the variation in the expression of genes associated with a particular phenotype (29). *Cis* eQTLs act on local gene expression, while *trans* eQTLs act on distal genes and tend to be tissue-specific (30). Analyses of eQTL and SNP databases, such as FANTOM and ENCODE, can determine the statistical association between SNPs located at specific regions of the genome and measure the expression level of a particular gene in a pathway of interest. For example, the database GTEx has gene expression data from 48 different human tissues from 620 donors revealing the tissue specificity of eQTLs and SNPs. Pan-tissue eQTLs tend to be more significant in GWAS compared with those reported as eQTLs in only one tissue type (31). SNPs located in regions of high eQTL density are more likely to occur in regulatory elements, such as in enhancers, where they act as modulators of gene expression. eQTLs and SNPs can

the immune response and exploit the host metabolic and translational machinery to facilitate the completion of their life cycle (Figure 1). This can entail interacting with host RNA or proteins and altering their function. RNA viruses typically have high mutation rates due to a lack of proofreading activity. This can lead to a high frequency of recombination events and the ability of the virus rapidly to adapt to new environments and undergo intrahost evolution to avoid cellular immune responses or antiviral therapies (20, 21). This complicates the development of antiviral therapies targeting RNA viruses, as they can develop drug resistance while maintaining their fitness. The coronavirus family, which includes SARS-CoV-2, encodes an exonuclease proofreading function in the nsp14 open reading frame, so it is thought that it might not mutate as rapidly as other RNA viruses, and this could be a region of the virus with potential for targeting with antivirals (22, 23).

RNA viruses include influenza, hepatitis C virus (HCV), Ebola, rabies, HIV, and SARS-CoV-2. They can be double-stranded RNA
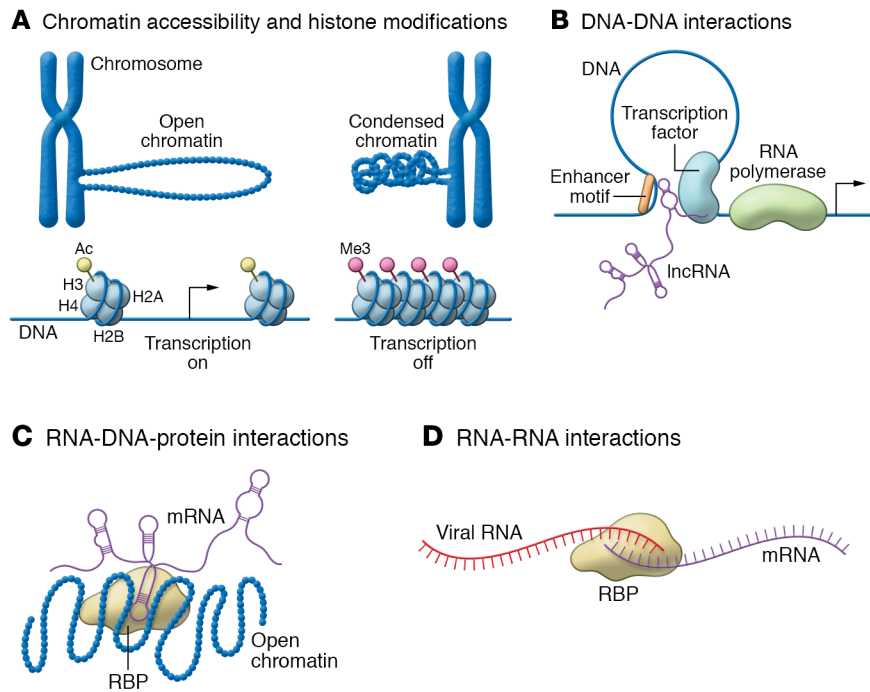
Figure 2. Cellular nucleic acid interactions. Annotating the human epigenome has uncovered nucleic acid interactions that determine the expression of genes involved in the response to viral infections. (**A**) Chromatin accessibility and histone modifications determine whether or not a gene will be transcribed. (**B**) Chromatin capture techniques identify locations in the genome that are interacting. (**C**) Various pull-down approaches use a protein of interest to determine whether specific RNA or chromatin regions are interacting. (**D**) Recent studies have shown that RNA transcripts interact with other RNA transcripts and that this can be independent of RNA-binding proteins.

coding sequences or codons, they do not need to remain in frame. They are transcripts greater than 200 bp long and are often polyadenylated and spliced but do not contain Kozak sequences that act as translation initiation sites. LncRNAs exert their effects on the cell using four general modes of action (signals, scaffolds, guides, or decoys), interacting with chromatin, other RNA molecules, or proteins. LncRNAs can function to control gene expression either in *cis*, where they influence the expression and/or chromatin state of neighboring genes, or in *trans*, where the lncRNA leaves the site of transcription and regulates genes on different chromosomes. An enhancer sequence can contain splicing signals that RNA polymerase II will recognize and transcribe, leading it to be mistaken for an lncRNA. However, the enhancer RNA (eRNA) has no biochemical activity beyond its contribution to "enhancing" the activity of the enhancer. eRNAs are structurally like lncRNAs but are transcribed from the active enhancer site rather than a promoter (36). Enhancer regions and lncRNA genes are hotspots for SNPs, and viruses could use this to their advantage to manipulate how key immune transcription factors are activated. Viruses interact with various host elements to their advantage, and quickly evolving elements like enhancer motifs, lncRNAs, and eRNAs may be active players in the variable responses seen between people infected with SARS-CoV-2.

Regulatory mechanisms by which SNPs can affect gene expression

Mutations within annotated protein-coding genes have traditionally been considered the major genetic causes of human disease, but it is now evident from GWAS that the majority of SNPs are found within noncoding regions of the genome and are likely to be involved in gene regulation (32). Only 12% of 465 SNPs identified in 151 GWAS were located in protein-coding regions, while 40% were found in introns and another 40% in intergenic regions (1). Regions of the genome that were once considered gene deserts harbor sequences encoding enhancer regulatory elements and lncRNA genes. Enhancer sites are recognized by transcription factors and enriched in histone modifications. Because they bind to transcription factors, most enhancer elements are found within open chromatin regions. Enhancers can bypass neighboring genes through chromosome looping to regulate genes located distally along a chromosome. Studies have shown that a substantial fraction of enhancers display weak conservation or no conservation across species (33, 34). Similar to enhancer motifs, lncRNAs are also poorly conserved, and this is thought to be due to the speed at which they evolve (35). Because they do not contain protein-

cause variation in gene expression through a variety of mechanisms, including altered transcription factor binding, histone modifications, and DNA methylation. SNPs can also change splicing sites to affect how mRNAs are degraded and polyadenylated as well as alter microRNA (miRNA) binding sites in 3′-UTRs.

Measuring chromatin accessibility

Recent efforts to annotate the human epigenome have identified millions of putative regulatory elements like enhancers and lncRNAs using correlative features such as chromatin accessibility and histone modifications (Figure 2A). "Open" and "closed" structural states determine the ability of chromatin to interact with gene regulatory elements. Cells use DNA methylation to lock genes in the "off" position and remain closed, but when chromatin is unmethylated it is open to interacting with elements like transcription factors. Methylation plays a vital role in numerous cellular processes, and abnormal patterns of methylation have been linked to human disease (37). Some viral infections can alter methylation patterns of gene promoters. For example, hepatitis B virus (HBV) upregulates insulin-like growth factor 2 (IGF-2) by hypomethylating the *IGF2* promoter (38). HCV causes hypermethylation at the *SOCS1* (suppressor of cytokine signaling 1) promoter, decreasing its expression and increasing viral infection (39). The extent of repression by DNA methylation can be determined using bisulfite sequencing to measure DNA methylation on cytosines (40). Treatment of DNA with bisulfite converts unmethylated cytosine residues to uracils but leaves 5′ methylated cytosines unaffected. Thus, bisulfite-treated DNA retains only methylated cytosines, which can be quantified by next-generation sequencing.

Deoxyribonuclease I sequencing (DNase-Seq) and assay for transposase-accessible chromatin using sequencing (ATAC-Seq) are alternative techniques used to measure chromatin accessibility. In DNase-Seq, chromatin is treated with DNase I, and the liberated DNA is sequenced to measure accessibility. Next-generation sequencing reveals genomic regions that are bound by regulatory proteins protected from the DNase I digestion (41). ATAC-Seq uses unfixed cells and approximately a thousand-fold fewer cells than DNase-Seq. ATAC-Seq recognizes open chromatin using a highly active transposase that fragments DNA and inserts into open chromatin sites, which are then identified by sequencing (42).

Methods for detecting chromatin interactions

Techniques are rapidly evolving to study interactions within the genome. Chromosome conformation capture methods including 3C (one contact vs. one contact), 4C (one vs. all), 5C (many vs. many), HiC (high-throughput chromosome conformation capture), ChIA-PET (chromatin interaction analysis by paired-end tag sequencing), and HiChIP (high-throughput chromosome conformation capture with chromatin immunoprecipitation) are used to quantify long-range interactions within the genome, creating a map of chromosomal architecture (Figure 2B and ref. 43). Earlier methods — 3C, 4C, and 5C — do not map interacting regions with high resolution, whereas HiC quantifies all possible pairwise interactions between DNA fragments (44). ChIA-PET combines HiC with chromatin immunoprecipitation (ChIP) to identify two distantly located segments of the genome from one fragment whose interaction is mediated by a particular DNA-binding protein (45). HiChIP creates long-range DNA contacts in the nucleus before cells are lysed, helping to minimize false-positive interactions and improve the efficiency of capturing DNA contacts (46).

ChIP-Seq is used to probe the genome for protein associations by immunoprecipitating chromatin with an antibody for a specific DNA-binding protein of interest, such as modified histones (47). Histone proteins comprise the nucleosome, which organizes DNA. Modifications to histones include methylation and acetylation, and these can affect the exposure of enhancer motifs and accessibility of DNA. Active enhancers are commonly marked by monomethylation of histone H3 lysine 4 (H3K4me1) and acetylation of histone H3 lysine 27 (H3K27ac) in a cell type–specific manner. Trimethylation of histone H3 at lysine 4 (H3K4me3) is typically enriched around transcription start sites and regulates gene activation through chromatin remodeling, making the DNA more accessible to transcription factors. Several studies have shown that viruses can manipulate the activity of histone proteins (48). Because some viral nucleic acids and proteins can enter the nucleus, it is unsurprising that some viruses can impact chromatin architecture. Epstein-Barr virus (EBV) modifies trimethylation at lysine 27 on histone H3 (H3K27me3), constituting a repressive mark, at the *BIM* gene promoter. BIM is an inducer of apoptosis and regulator of lymphocyte survival. EBV is thought to alter H3K27me3 by regulating the polycomb repressor complex 2 (PRC2) and inhibiting BIM transcription (49).

Methods for detecting RNA interactions

RNA-binding proteins (RBPs) are typically required to mediate RNA interactions, but RNA can also directly bind to chromatin (Figure 2C). MARGI (mapping RNA-genome interactions) allows for the identification of all chromatin-associated RNAs and their respective genomic target (50). RNA is crosslinked and ligated to target chromatin, resulting in the formation of chimeric sequences. Like MARGI, PIRCh-Seq (profiling interacting RNAs on chromatin followed by deep sequencing) examines RNA interactions with chromatin but uses an immunoprecipitation step to enrich for modified histone proteins (51). This increases specificity, reduces the influence of nascent transcripts, and results in a significantly lower number of intronic reads.

ChIRP-Seq (chromatin isolation by RNA purification) and RAP (RNA antisense purification) use 20-nucleotide or 120-nucleotide biotin-labeled oligonucleotides to label the entire length of a specific RNA. Cells are crosslinked and the nuclei are isolated. Chromatin is fragmented and the labeled oligonucleotides are hybridized to the fragments. Complexes are captured using streptavidin beads, and DNA is isolated and sequenced, identifying regions bound to the RNA of interest (52, 53). HITS-CLIP (high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation), also known as CLIP-Seq, is a method used to detect RNA and protein interactions using a specific protein as bait (54). HITS-CLIP was used to identify mRNAs and miRNAs associated with the RNA-binding protein argonaute-2 (AGO2), a key component of the RNAi silencing complex (55). This study generated a genome-wide interaction map illustrating miRNA binding sites within both 3′-UTR and coding sequences of target mRNAs.

Uncovering RNA-RNA interactions can provide information about dynamic post-transcriptional processes (Figure 2D). Many of the techniques used to detect RNA-RNA interactions require the identification of a specific RBP to pull out bound RNAs from the cell. RBPs bind, guide, and modify RNA transcripts by post-transcriptionally regulating splicing, polyadenylation, and the stabilization, localization, and translation of mRNAs. RBPs contain structural motifs for RNA recognition, dsRNA binding domains, and CCCH zinc finger domains (56). The structure and target of an mRNA, including lncRNAs, can be uncovered by profiling of RBP-mediated RNA interactions. Some methods used to uncover RNA-RNA interactions require the overexpression of a specific RBP, which may perturb the native RNA-RNA interaction network. CLASH (crosslinking, ligation, and sequencing of hybrids) detects RNA-protein complexes by first ectopically expressing a particular RBP and then crosslinking and affinity-purifying the crosslinked complexes (57). RNA-RNA hybrids of base-paired RNA are ligated, isolated, and reverse-transcribed into cDNA that is deep-sequenced, providing high-resolution chimeric reads of RNA-RNA interactions. hiCLIP (RNA hybrid and individual-nucleotide resolution ultraviolet crosslinking and immunoprecipitation) sequences RNA duplexes bound to RBPs in vivo (58, 59). Its unique linker-adapter system identifies whether the RBP-bound RNA duplex originates from the same RNA or two different RNAs. Like HiC for DNA, CLASH and hiCLIP can only assay one overexpressed protein at a time and cannot identify non–RBP-bound RNA duplexes.

MARIO (mapping RNA interactome in vivo) identifies RNA duplexes by double crosslinking to fix all protein-RNA interactions (60). The double crosslinking of the RNA-RNA complexes can lead to the formation of large protein aggregates that may bring non-physiologically relevant RNAs close to each other, resulting

in false RNA interactions. An RNA linker is added to the 5′ end of one strand of the RNA duplex, and proteins are biotinylated on cysteine residues, creating a level of bias. Biotin-tagged proteins are then pulled out from the cell lysates, and the RNA ends are ligated together, followed by sequencing. PARIS (psoralen analysis of RNA interactions and structures) was developed to determine transcriptome-wide interactions between RNAs without the limitation or bias of a bait protein (61). The method uses psoralen to reversibly crosslink RNA-RNA complexes, which has sequence bias, preferring to bind to UpA dinucleotides occurring once every 16 bp in a random duplex. RIC-Seq (RNA in situ conformation sequencing) profiles RNA-RNA interactions at even greater single-nucleotide resolution than PARIS (62). Cells are crosslinked with formaldehyde, RNA transcripts are randomly cut, and 3′ overhangs are dephosphorylated and biotin-labeled. RNA fragments in close proximity are ligated, and total RNA is extracted, fragmented, and converted into strand-specific cDNA libraries for sequencing. (Table 1 summarizes the methods described.) By using novel RNA-RNA interaction techniques like RIC-Seq in the context of the RNA virosphere (Figure 3), new host-pathogen interactions can be revealed. LncRNAs have been implicated in the response to viral infections or to synthetic viral molecules like the dsRNA analog polyinosinic-polycytidylic acid (poly I:C), including lnc-IL7R, AS-IL-1α, NRAV, and NEAT1 (63–66), and it will be interesting to see whether there are immune defense mechanisms occurring where lncRNAs are interacting with viral RNAs. The techniques described here can be adapted to uncover these interactions that could be integral to the viral strategy of immune evasion.

Interactions between RNA viruses and the host

Examples have been described of viruses altering host mRNA stability, manipulating DNA methylation and histone modifications, and hijacking host factors like RBPs to stabilize their transcripts and potentially interfere with host translation (38, 39, 49). The RNA genome of (+)ssRNA viruses has an inherent capacity to form base pair interactions with host RNAs. (+)ssRNA viruses like SARS-CoV-2 are capable of exploiting the host because of the simplicity of their genomes and mechanisms of infiltration. For (+)ssRNA viruses to expand, they must trick the host mRNA decay machinery into turning off. Cytoplasmic mRNA decay occurs via two major pathways: deadenylation-dependent 5′ to 3′ decay and exonucleolytic 3′ to 5′ decay pathways. Viruses can disrupt the cellular decapping machinery to promote translation and replication of their own viral RNA genomes (67, 68). A short noncoding RNA, sfRNA, is transcribed from the 3′-UTR of Dengue virus to inhibit the exonuclease XRN1 during infection to stabilize its own RNA while making host mRNA less stable (69, 70).

Processing bodies (p-bodies) are cytoplasmic granules that contain translationally repressed mRNAs and proteins involved in 5′ to 3′ mRNA decay, such as the DEAD box helicase DDX6, XRN1, and AGO2 (71). Once in p-bodies, mRNAs can be either degraded or stored for future translation. (+)ssRNA viral infections like HCV reduce the cytosolic concentration of proteins essential for p-body formation, leading to disruption of p-bodies and the storage of essential mRNAs, particularly those needed for maintenance of mRNA degradation and decay (72). When HCV infects hepatocytes, DDX6 is recruited to lipid droplets, where it promotes HCV

assembly, allowing viral replication to continue. Sequestration of the cellular decay machinery by (+)ssRNA viruses has the potential to alter the whole transcriptional/translational landscape of the host (73). This is a highly conserved strategy for viruses across species and pinpoints a weak spot that can be exploited for the development of broad-spectrum antiviral drugs.

HIV uses several mechanisms to exploit the host for its benefit, including one where it interacts with the host transfer RNA, tRNALys3, to facilitate its reverse transcription into DNA and integration into the host genome. An HIV strain with mutations in tRNALys3 binding sites in its RNA genome was generated to prevent binding to tRNALys3, resulting in spontaneous reversion back to the wild-type sequence, which could bind to tRNALys3 (74). When the HIV binding site mutants were bound to alternative tRNAs, they were unable to revert to the wild-type sequence and displayed attenuated replication, suggesting that selection of specific tRNAs may affect viral fitness. Additional sequences encoded in the HIV genome that are complementary to tRNALys3 have been described; these can also promote reverse transcription (75). Several studies have reported an antisense transcript transcribed from the nef region of HIV (76, 77). Nef downregulates CD4 and MHC class I on host cells. When the HIV antisense RNA is expressed, there is a loss of the epigenetic modifier DNMT3a, and methyl groups are retained at the viral promoter, resulting in silenced gene transcription. This suggests that the viral antisense RNA suppresses viral gene expression and that it may be involved in epigenetic regulation of HIV. HIV and the dsDNA virus herpes simplex virus (HSV) remain incurable because they exist in latently infected cells where they are dormant and do not produce active virus. The latent silencing of these viruses occurs through epigenetic alterations (78–80). ChIP-Seq experiments demonstrated that during viral latency histones are substantially modified, with histone deacetylase (HDAC) recruited to viral promoters, resulting in transcriptional repression. HDAC inhibitors (HDACis) block the removal of acetyl groups from histones, making chromatin accessible, and gene expression is increased. This strategy is being explored for preventing HIV and HSV from transitioning into their latent phase, and the active virus can be targeted with antiviral therapies (81). Further ChIP-Seq experiments could help elucidate what specific host and viral factors are interacting and lead to better therapies. African green monkeys (AGMs) are natural hosts for simian immunodeficiency virus (SIV) in which the virus is found at high viral loads but remains nonpathogenic (SIVagm). When non-natural hosts like rhesus macaques are infected, the virus is pathogenic and the infection progresses to AIDS. Similar to HIV, SIV uses the CD4 and CCR5 receptors to enter host cells, and natural hosts of SIV can regulate the expression of the receptors (82). CD4+ T cells isolated from AGMs and stimulated with SIV exhibit decreased CD4 expression via hypermethylation at the CD4 locus (83). AGMs have evolved their own mechanisms to compensate for the decrease in CD4 expression and defeat SIV. Because HIV uses the same host entry receptors, the SIVagm model can potentially provide an understanding of the pathology of HIV.

A study examining cells infected with Zika virus (ZIKV) used a modified RAP protocol called COMRADES (crosslinking of matched RNAs and deep sequencing) to identify direct interactions between ZIKV RNA and host RNAs. ZIKV interacted with several

Table 1. Methods to detect chromatin accessibility, DNA-DNA interactions, RNA-DNA-protein interactions, and RNA-RNA interactions

Method	Abbr.	Method details	Adaptation to study viral interactions	Refs.
Chromatin accessibility				
Bisulfite sequencing	Bisulfite-Seq	Converts unmethylated cytosine residues to uracils, leaving 5′ methylated cytosines unaffected; DNA that has been treated retains only methylated cytosines	Examine effects on methylation at specific host gene promoters following viral infection	40
Deoxyribonuclease I sequencing	DNase-Seq	Chromatin is treated with DNase I, and DNA is sequenced to measure accessibility; sequences bound by proteins are protected from DNase I	Examine effects on methylation at specific host gene promoters following viral infection	41
Assay for transposase-accessible chromatin with high-throughput sequencing	ATAC-Seq	Recognizes open chromatin using a highly active transposase that fragments DNA and inserts sequencing primers into open chromatin sites	Examine effects on methylation at specific host gene promoters following viral infection	42
DNA-DNA interactions				
Chromosome conformation capture	3C/4C/5C	Detects long-range chromatin interactions: 3C (one vs. one), 4C (one vs. all), 5C (many vs. many), but all low throughput and not genome-wide	Detect specific regions of chromatin interacting with DNA viruses	43
High-throughput chromosome conformation capture	HiC	Detects all possible pairwise interactions between DNA fragments	Detect genome-wide regions of chromatin interacting with DNA viruses	44
Chromatin interaction analysis by paired-end tag sequencing	ChIA-PET	Combines HiC with ChIP to detect long-range chromatin interactions that are mediated by a protein of interest	Determine whether virus is interacting with a specific protein of interest and chromatin genome-wide	45
High-throughput chromosome conformation capture with chromatin immunoprecipitation	HiChIP	Creates long-range DNA contacts in the nucleus before cells are lysed to minimize false-positive interactions	Examine direct interactions between viral DNA and host chromatin across the genome in actively infected cells	46
RNA-DNA-protein interactions				
Chromatin immunoprecipitation sequencing	ChIP-Seq	Probes the genome for chromatin/protein associations with an antibody for a protein of interest	Determine whether virus is interacting with a specific protein of interest	47
Mapping RNA-genome interactions	MARGI	Identifies all chromatin-associated RNAs and their genomic target	Identify regions of an RNA virus interacting with the whole host genome	50
Profiling interacting RNAs on chromatin followed by deep sequencing	PIRCh-Seq	Detects RNA interactions with chromatin but uses an IP step to enrich for modified histone proteins	Detect regions of an RNA virus that are interacting with genome-wide chromatin and a specific RBP	51
Chromatin isolation by RNA purification with sequencing	ChIRP-Seq	Uses biotin-labeled 20-nt-long oligonucleotides to label the length of an RNA to identify genomic regions that have bound	Use probes to examine a specific viral RNA that is interacting with the genome	52
RNA antisense purification	RAP	Use biotin-labeled 120-nt-long oligonucleotides to label the length of an RNA to identify genomic regions that have bound	Use probes to examine a specific viral RNA that is interacting with the genome	53
High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation	HITS-CLIP	Detects RNA and protein interactions using a specific protein as bait	Determine whether an RNA virus is interacting with a specific protein	54, 55
RNA-RNA interactions				
Crosslinking, ligation, and sequencing of hybrids	CLASH	Detects RNA-protein complexes by overexpressing an RBP and affinity-purifying the crosslinked complexes	Determine whether an RNA virus is interacting with a specific RBP-mediated host RNA	57
RNA hybrid and individual-nucleotide resolution ultraviolet crosslinking and immunoprecipitation	HiCLIP	Identifies RNA duplexes bound to RBPs in vivo; can only assay one overexpressed protein at a time	Determine whether an RNA virus is interacting with a specific RBP-mediated host RNA	58, 59
Mapping RNA interactome in vivo	MARIO	Detects protein-mediated RNA-RNA interactions in cells using any protein	Determine whether an RNA virus is interacting with a host RNA across the transcriptome that is mediated by any protein	60
Psoralen analysis of RNA interactions and structures	PARIS	Identifies transcriptome-wide interactions between RNAs but without the limitation of a bait protein	Examine viral RNAs interacting with host RNAs following crosslinking	61
RNA in situ conformation sequencing	RIC-Seq	Profiles all RNA-RNA interactions	Examine viral RNAs interacting with host RNAs following crosslinking	62

Details of each method are provided, as well as how each method could be adapted to examine interactions between host cells and viruses.

miRNAs, including miR-21. This study demonstrated the ability of a viral RNA genome to engage with multiple host RNAs; however, no experiments were performed to determine whether the interactions were functional (84). NeST is a conserved lncRNA found within the same locus as the *IL22* and *IFNG* genes but transcribed from the opposite strand (85). NeST interacts with histone methyltransferase WDR5 to alter histone methylation of the *IFNG* promoter and stimulate its expression. CD8+ T cells isolated from transgenic
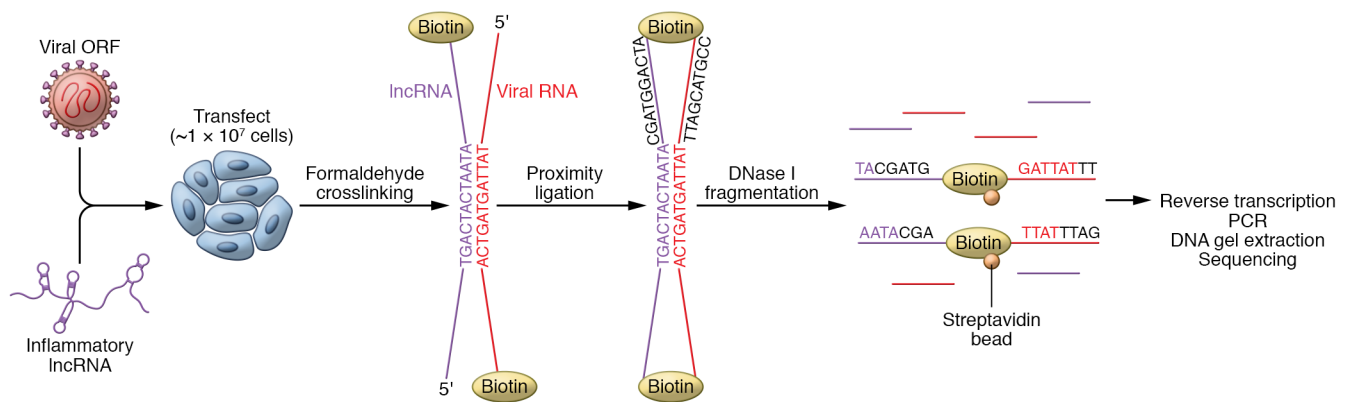
Figure 3. Chromatin and RNA interaction techniques can be used to uncover novel interactions with viral nucleic acids. RNA-RNA interaction methods such as RIC-Seq can be modified to determine whether a viral RNA interacts with host RNAs like lncRNA transcripts. Overexpression of both the viral RNA and the host RNA in a target cell can identify interacting regions through proximity ligation and fragmentation followed by reverse transcription into cDNA and PCR to identify interacting sequences. Adapted with permission from *Nature* (62).

mice overexpressing NeST are more susceptible to infection by the murine (+)ssRNA virus Theiler's murine encephalomyelitis virus (TMEV) while exhibiting decreased IFN-γ expression. NeST alters the magnitude and timing of the inflammatory response, activating basal inflammation to attenuate subsequent inflammatory events. It remains possible that NeST could have other targets in addition to IFN-γ, perhaps interacting directly with viral RNA.

A study using AGO-CLIP examined a panel of 15 RNA viruses to identify host AGO2-bound miRNAs that interact with viral RNA (86). Interactions were identified between miR-17, let-7, and the 3′-UTR of bovine viral diarrhea virus (BVDV; related to HCV) enhancing viral RNA stability. Because the virus was sequestering miRNAs and AGO2, there was reduced miRNA binding to host mRNAs, including IFN-stimulated genes, resulting in the derepression of their expression during infection. Treatment with miRNA antagonists targeting virus-associated miRNAs could be used as an antiviral strategy to reduce miRNA association with the virus. HCV replication is dependent on liver-specific miR-122 expression (87). miR-122 binds two sites in the HCV 5′-UTR. AGO2 is recruited to the viral internal ribosome entry site and binds to the miR-122 sites, leading to promotion of viral protein translation. Mutating the miRNA binding sites in the viral genome or blocking endogenous miR-122 with antisense oligonucleotides (ASOs) decreased HCV growth, demonstrating that this interaction is required to sustain HCV replication. HCV sequesters miR-122 to redirect miRNA repression away from its endogenous host mRNA targets. The 3′-UTR of HIV contains a miR-29 binding site. T cells infected with a virus expressing a wild-type 3′-UTR exhibited decreased viral replication when exposed to miR-29. However, when cells were infected with a virus expressing a miR-29 seed site mutant, there was elevated replication, demonstrating the impact of host miR-29 on viral proliferation (88). Simian foamy virus (SFV) encodes its own miRNA called miR-S4-3p, which mimics the seed sequence of cellular miR-155 (89). Several targets of miR-155 regulate cell proliferation, leading to the hypothesis that viral miRNAs such as miR-S4-3p stimulate proliferative activity of SFV-infected cells. Table 2 summarizes the viruses described and their interactions with the host.

SARS-CoV-2 belongs to the coronavirus family, members of which have the largest ssRNA genomes. Because it is a (+)ssRNA virus, its RNA transcription and translation are controlled by interactions with its own RNA and proteins as well as host elements. Genes located at the 3′ end of the viral genome are transcribed into subgenomic negative-strand mRNAs (subgenomic RNAs) (90). An in silico analysis of the SARS-CoV-2 genome identified a set of virus-encoded miRNAs that potentially regulate host signaling pathways, including miR-33a-3p, which has been implicated in the regulation of cell proliferation and lipid metabolism, two processes relevant to the pathology of viral infections (91–93).

A recent study mapped the RNA-RNA interactome for SARS-CoV-2 using COMRADES (84, 94). There is a high prevalence of long-distance RNA base-pairing along the SARS-CoV-2 genome, with ORF1a having the most long-range connectivity. There are site-specific interactions between viral RNA and small nuclear RNAs, which function in splicing. The RNA subunit of the RNase mitochondrial RNA processing (MRP) enzyme complex base-pairs with a SARS-CoV-2 subgenomic RNA and is implicated in preribosomal processing and viral RNA degradation. Because there are overlapping sequences among coronaviruses, these can potentially be targeted for identifying new antiviral drug targets. Understanding how viral genomic RNA, subgenomic RNAs, and host RNAs are brought together to manipulate the virus is key to defeating it.

A controversy surrounding nucleic acid interactions

Because a majority of the methods described here use in situ techniques in which cells are fixed at a certain point in time and only nucleic acid interactions occurring at that precise moment are examined, several considerations are necessary to decide whether the nucleotide interactions detected are physiologically possible, including determining the number of transcripts per cell of each interacting partner using quantitative PCR and a standard curve of known quantities of the DNA or RNA to compare with the unknown samples. Another factor to consider is the

Table 2. Viruses, their type of genome, targets in the host cell, and effect of the interaction on the virus

Virus	Viral genome	Target	Effect on virus	Refs.
Hepatitis B virus (HBV)	DNA	Upregulates IGF-2 by hypomethylating the *IGF2* promoter	Promotes viral replication	38
Hepatitis C virus (HCV)	(+)ssRNA	Causes hypermethylation at the *SOCS1* promoter, decreasing its expression	Increases viral infection	39, 72, 73, 87
		Reduces the cytosolic concentration of p-body proteins essential for p-body formation; binds to the DEAD box helicase DDX6 and recruits it to lipid droplets	Promotes HCV RNA translation and replication	
		miR-122 binds two sites in the HCV 5′-UTR	Promotes viral replication	
Epstein-Barr virus (EBV)	dsDNA	Regulates PRC2 to alter H3K27me2 at the *BIM* gene promoter, inhibiting its expression	Promotes viral replication	48
Dengue virus	(+)ssRNA	Noncoding RNA; sfRNA is transcribed from the 3′-UTR of virus, inhibiting the exonuclease activity of XRN1	Stabilizes its own RNA while disrupting host mRNA stability	69, 70
Human immunodeficiency virus (HIV)	(+)ssRNA/ retrovirus	Interacts with host transfer RNA, tRNALys3	Facilitates HIV reverse transcription into DNA and integration into the host genome	74–77, 88
		Antisense transcript from the nef region of HIV downregulates CD4 and MHC class I on host cells	Viral antisense RNA suppresses viral gene expression	
		3′-UTR of HIV contains a miR-29 binding site	Decreases viral replication	
HIV/Herpes simplex virus (HSV)	(+)ssRNA/ retrovirus dsDNA	During viral latency histones are substantially modified with HDAC recruited to viral promoters	Represses viral transcription; virus goes into latent phase	78–81
Simian immunodeficiency virus (SIV)	(+)ssRNA/ retrovirus	CD4+ T cells isolated from AGMs and stimulated with SIV exhibit decreased CD4 expression via hypermethylation at the CD4 locus	Nonpathogenic symptoms in AGMs but high viral loads	82, 83
Zika virus	(+)ssRNA	Interacts with miR-19, miR-512, miR-5151, miR-1323, and miR-21	Not determined	84
Theiler's murine encephalomyelitis virus (TMEV)	(+)ssRNA	Decreases IFN-γ and decreases susceptibility to *Salmonella* infection by increasing lncRNA NeST, which binds to histone modification complex	Increases Theiler's virus persistence	85
Bovine viral diarrhea virus (BVDV) (related to HCV)	(+)ssRNA	Interaction between miR-17 and let-7 and the 3′-UTR of BVDV	Promotes viral replication	86
Simian foamy virus (SFV)	(+)ssRNA/retrovirus	Encodes miRNA miR-S4-3p, which mimics miR-155	miR-S4-3p stimulates proliferative activity of SFV-infected cells	89

thermodynamics of the nucleic acid interactions. Every reaction in the cell costs energy, and nucleic acid interactions require energy-intensive conformational changes to their structures (95). An example in host cells that leads to controversy is the hypothesis that miRNAs and lncRNAs can compete for binding to target mRNAs. The current guidelines put forth for measuring the ability of these competing endogenous RNAs (ceRNAs) to interact come from theoretical and prediction-based models and use non-physiological levels, at which miRNAs are overexpressed (96). These computational analyses suggest that comparable levels of a single target RNA carrying a single binding motif are unlikely to be consequential because the number of other binding sites for a specific RNA in the transcriptome and overall target occupancy need to be considered. It is currently unclear how a single transcript with few binding motifs, no matter how abundant, can compete against the pool of thousands of sites found in the rest of the transcriptome. However, if there is a cooperative mechanism between binding sites and multiple RNAs, the likelihood of a ceRNA effect increases (97). The number of virus and host molecules, thermodynamic potential, and number of binding sites need to be considered to determine whether interactions uncovered using the described methods are physiologically possible.

Methods to alter nucleotide interactions as potential antiviral strategies

Several chemical inhibitors targeting epigenetic modifications have reached clinical application (98). Histone modifications, such as acetylation and methylation, have been pursued as drug targets because they are the most immediate contributors to epigenetic regulation. HDACis block removal of acetyl groups from histones, leaving chromatin accessible and gene expression increased. These compounds might particularly impact DNA viruses. U2OS cells infected with HSV and treated with the HDACi trichostatin A showed a specific increase in expression of the antiviral genes *ATRX* and *PML*, leading to a reduction in parental viral genomes (98). DNMT inhibitors (DNMTis) prevent methylation of DNA. Hypermethylation has been found on tumor suppressor genes in cancers, and DNMTis can remove the methyl groups, making the DNA accessible and reactivating these tumor suppressors. DNMT-deficient mice show upregulation of inflammatory mediators and increased atherosclerosis and inflammation (99, 100). Dysregulated lipid metabolism has been associated with hypermethylation at the promoter of the cholesterol transporter ABCA1 and hypomethylation of the cholesterol sensor INSIG (101). DNMTi removes methyl groups from the hypermethylated promoter of *ABCA1*, increasing its expression and decreasing

total cellular cholesterol. The DNA methyltransferases DNMT1 and DNMT3 are required for HCV propagation, and the DNMTis 5-Aza-C and 5-Aza-dC significantly degrade DNMT1 protein and suppress HCV infection, replication, and protein expression (102).

CRISPR/Cas9 creates breaks in DNA via the endonuclease activity of Cas9, and through endogenous DNA repair mechanisms, gene sequences can be precisely edited. Dead Cas9 (dCas9) is a modified Cas9 enzyme that lacks the endonuclease activity but can still be used to guide oligonucleotides to a specific region of the genome. dCas9 can employ transcriptional activators or inhibitors to increase or decrease expression of genes of interest. Enhancer CRISPR activator (enCRISPRa) uses dCas9 fused with the core domain of histone acetyltransferase p300, and, together with the MS2-tagged sgRNA sequence, they recruit the activator domains of the MS2 coat protein (MCP) and its fusion partner VP64, acting like artificial transcription factors. Enhancer CRISPR interference (enCRISPRi) uses dCas9 fused with the lysine-specific demethylase LSD1. Together with MS2-sgRNA they recruit the MCP-KRAB repressor domain to block enhancer activity. These methods can be used to alter SNPs in enhancers, lncRNAs, or eRNAs that are interacting with viral nucleic acids and further understand their impact.

ASOs, small interfering RNAs (siRNAs), short hairpin RNAs (shRNAs), and locked nucleic acid (LNA) antagonist oligonucleotides are designed to target mRNAs and could potentially be used to target viral RNAs in infected cells. Lipid nanoparticles and viral vectors are being studied to ensure intracellular delivery. There has been some success in vitro using ASOs targeting HIV; researchers targeted both viral and host factors, delivering the ASOs using a self-inactivating lentiviral vector system to HIV-infected human primary cells, and produced strong suppression of HIV replication (103). HITS-CLIP found miR-122 binding directly to the 5′-UTR of HCV, protecting HCV RNA from degradation and promoting viral replication. An LNA antagomiR (miRNA antagonist) targeting miR-122, miravirsen (also called SPC3649), is in clinical trials and has shown potential to suppress HCV (104).

Conclusion

The pandemic has emphasized the need to enhance our understanding of virus RNA–host RNA interactions. Such insights promise to elucidate the remarkable diversity apparent in the clinical response to COVID-19. For example, a lingering impact of previous viral infections on the epigenome of immune cells may be relevant to the association of disease severity with age, poverty, and comorbid conditions such as obesity and diabetes. Viruses are skilled at harnessing host cells to enhance their replication. Our challenge is to distinguish their interactions with host RNA that favor this objective from those that are intrinsic to viral clearance. To address this challenge, a range of novel technologies have emerged that permit us to interrogate RNA-RNA interactions; perhaps they will lead us to novel targets for antiviral drugs and a new era of precision medicine.

Acknowledgments

Address correspondence to: Elizabeth J. Hennessy, Institute for Translational Medicine and Therapeutics (ITMAT), University of Pennsylvania, Perelman School of Medicine, 3400 Civic Center Boulevard, Philadelphia, Pennsylvania 19104, USA. Phone: 215.898.1185; Email: ehenn@pennmedicine.upenn.edu.

1. Hindorff LA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*. 2009;106(23):9362–9367.
2. Deplancke B, et al. The genetics of transcription factor DNA binding variation. *Cell*. 2016;166(3):538–554.
3. Huang C, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):497–506.
4. Palaiodimos L, et al. Severe obesity, increasing age and male sex are independently associated with worse in-hospital outcomes, and higher in-hospital mortality, in a cohort of patients with COVID-19 in the Bronx, New York. *Metabolism*. 2020;108:154262.
5. Harrison SL, et al. Comorbidities associated with mortality in 31,461 adults with COVID-19 in the United States: a federated electronic medical record analysis. *PLoS Med*. 2020;17(9):e1003321.
6. Chen J, et al. Distinct clinical characteristics and risk factors for mortality in female inpatients with coronavirus disease 2019 (COVID-19): a sex-stratified, large-scale cohort study in Wuhan, China [published online July 8, 2020]. *Clin Infect Dis*. https://doi.org/10.1093/cid/ciaa920.
7. Ye C, et al. Impact of comorbidities on patients with COVID-19: a large retrospective study in Zhejiang, China. *J Med Virol*. 2020;92(11):2821–2829.
8. Richardson S, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *JAMA*. 2020;323(20):2052–2059.
9. Ravi K. Ethnic disparities in COVID-19 mortality: are comorbidities to blame? *Lancet*. 2020;396(10243):22.
10. Zhang Q, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science*. 2020;370(6515):eabd4570.
11. Bastard P, et al. Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science*. 2020;370(6515):eabd4585.
12. Chen Z, John Wherry E. T cell responses in patients with COVID-19. *Nat Rev Immunol*. 2020;20(9):529–536.
13. Mathew D, et al. Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science*. 2020;369(6508):eabc8511.
14. Ortiz-Fernandez L, Sawalha AH. Genetic variability in the expression of the SARS-CoV-2 host cell entry factors across populations. *Genes Immun*. 2020;21(4):269–272.
15. Lu N, et al. ACE2 gene polymorphism and essential hypertension: an updated meta-analysis involving 11,051 subjects. *Mol Biol Rep*. 2012;39(6):6581–6589.
16. Bosso M, et al. The two faces of ACE2: the role of ACE2 receptor and its polymorphisms in hypertension and COVID-19. *Mol Ther Methods Clin Dev*. 2020;18:321–327.
17. Gebhard C, et al. Impact of sex and gender on COVID-19 outcomes in Europe. *Biol Sex Differ*. 2020;11(1):29.
18. Klein SL, et al. Biological sex impacts COVID-19 outcomes. *PLoS Pathog*. 2020;16(6):e1008570.
19. Scully EP, et al. Considering how biological sex impacts immune responses and COVID-19 outcomes. *Nat Rev Immunol*. 2020;20(7):442–447.
20. Dolan PT, et al. Mapping the evolutionary potential of RNA viruses. *Cell Host Microbe*. 2018;23(4):435–446.
21. Taft RJ, et al. The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays*. 2007;29(3):288–299.
22. Gorbalenya AE, et al. Nidovirales: evolving the largest RNA virus genome. *Virus Res*. 2006;117(1):17–37.
23. Robson F, et al. Coronavirus RNA proofreading: molecular basis and therapeutic targeting. *Mol Cell*. 2020;79(5):710–727.
24. Rawlinson SM, Moseley GW. The nucleolar interface of RNA viruses. *Cell Microbiol*.

2015;17(8):1108–1120.

25. Pai AA, et al. The genetic and mechanistic basis for variation in gene regulation. *PLoS Genet*. 2015;11(1):e1004857.

26. Raisi-Estabragh Z, Petersen SE. Cardiovascular research highlights from the UK Biobank: opportunities and challenges. *Cardiovasc Res*. 2020;116(1):e12–e15.

27. Littlejohns TJ, et al. UK Biobank: opportunities for cardiovascular research. *Eur Heart J*. 2019;40(14):1158–1166.

28. Melamud E, et al. The promise and reality of therapeutic discovery from large cohorts. *J Clin Invest*. 2020;130(2):575–581.

29. Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. *Philos Trans R Soc Lond B Biol Sci*. 2013;368(1620):20120362.

30. Michaelson JJ, et al. Detection and interpretation of expression quantitative trait loci (eQTL). *Methods*. 2009;48(3):265–276.

31. Gorlov I, et al. SNP eQTL status and eQTL density in the adjacent region of the SNP are associated with its statistical significance in GWA studies. *BMC Genet*. 2019;20(1):85.

32. Farh KK, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. 2015;518(7539):337–343.

33. Pennacchio LA, et al. Enhancers: five essential questions. *Nat Rev Genet*. 2013;14(4):288–295.

34. Schmidt D, et al. Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science*. 2010;328(5981):1036–1040.

35. Djebali S, et al. Landscape of transcription in human cells. *Nature*. 2012;489(7414):101–108.

36. Espinosa JM. Revisiting lncRNAs: how do you know yours is not an eRNA? *Mol Cell*. 2016;62(1):1–2.

37. Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*. 2019;20(10):590–607.

38. Tang S, et al. Hepatitis B virus X protein promotes P3 transcript expression of the insulin-like growth factor 2 gene via inducing hypomethylation of P3 promoter in hepatocellular carcinoma. *Liver Int*. 2015;35(2):608–619.

39. Ko E, et al. CpG island hypermethylation of SOCS-1 gene is inversely associated with HBV infection in hepatocellular carcinoma. *Cancer Lett*. 2008;271(2):240–250.

40. Li Y, Tollefsbol TO. DNA methylation detection: bisulfite genomic sequencing analysis. *Methods Mol Biol*. 2011;791:11–21.

41. Song L, Crawford GE. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb Protoc*. 2010;2010(2):pdb prot5384.

42. Buenrostro JD, et al. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol*. 2015;109:21.29.21–21.29.9.

43. Kempfer R, Pombo A. Methods for mapping 3D chromosome architecture. *Nat Rev Genet*. 2020;21(4):207–226.

44. Lieberman-Aiden E, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326(5950):289–293.

45. Li X, et al. Long-read ChIA-PET for base-pair-resolution mapping of haplotype-specific chromatin interactions. *Nat Protoc*. 2017;12(5):899–915.

46. Mumbach MR, et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods*. 2016;13(11):919–922.

47. Park PJ. ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet*. 2009;10(10):669–680.

48. Paschos K, Allday MJ. Epigenetic reprogramming of host genes in viral and microbial pathogenesis. *Trends Microbiol*. 2010;18(10):439–447.

49. Paschos K, et al. Epstein-barr virus latency in B cells leads to epigenetic repression and CpG methylation of the tumour suppressor gene Bim. *PLoS Pathog*. 2009;5(6):e1000492.

50. Sridhar B, et al. Systematic mapping of RNA-chromatin interactions in vivo. *Curr Biol*. 2017;27(4):602–609.

51. Fang J, et al. PIRCh-seq: functional classification of non-coding RNAs associated with distinct histone modifications. *Genome Biol*. 2019;20(1):292.

52. Chu C, et al. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell*. 2011;44(4):667–678.

53. Engreitz JM, et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science*. 2013;341(6147):1237973.

54. Licatalosi DD, et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*. 2008;456(7221):464–469.

55. Chi SW, et al. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*. 2009;460(7254):479–486.

56. Dominguez D, et al. Sequence, structure, and context preferences of human RNA binding proteins. *Mol Cell*. 2018;70(5):854–867.e9.

57. Kudla G, et al. Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc Natl Acad Sci U S A*. 2011;108(24):10010–10015.

58. Sugimoto Y, et al. Using hiCLIP to identify RNA duplexes that interact with a specific RNA-binding protein. *Nat Protoc*. 2017;12(3):611–637.

59. Sugimoto Y, et al. hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. *Nature*. 2015;519(7544):491–494.

60. Nguyen TC, et al. Mapping RNA-RNA interactome and RNA structure in vivo by MARIO. *Nat Commun*. 2016;7:12023.

61. Lu Z, et al. RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell*. 2016;165(5):1267–1279.

62. Cai Z, et al. RIC-seq for global in situ profiling of RNA-RNA spatial interactions. *Nature*. 2020;582(7812):432–437.

63. Ding L, et al. The TLR3 agonist inhibit drug efflux and sequentially consolidates low-dose cisplatin-based chemoimmunotherapy while reducing side effects. *Mol Cancer Ther*. 2017;16(6):1068–1079.

64. Chan J, et al. Cutting edge: a natural antisense transcript, AS-IL1α, controls inducible transcription of the proinflammatory cytokine IL-1α. *J Immunol*. 2015;195(4):1359–1363.

65. Ouyang J, et al. NRAV, a long noncoding RNA, modulates antiviral responses through suppres-sion of interferon-stimulated gene transcription. *Cell Host Microbe*. 2014;16(5):616–626.

66. Morchikh M, et al. HEXIM1 and NEAT1 long non-coding RNA form a multi-subunit complex that regulates DNA-mediated innate immune response. *Mol Cell*. 2017;67(3):387–399.e5.

67. Sokoloski KJ, et al. Sindbis virus usurps the cellular HuR protein to stabilize its transcripts and promote productive infections in mammalian and mosquito cells. *Cell Host Microbe*. 2010;8(2):196–207.

68. Palusa S, et al. The 3′ untranslated region of the rabies virus glycoprotein mRNA specifically interacts with cellular PCBP2 protein and promotes transcript stability. *PLoS One*. 2012;7(3):e33561.

69. Moon SL, et al. A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular exoribonuclease XRN1 and alters host mRNA stability. *RNA*. 2012;18(11):2029–2040.

70. Chapman EG, et al. The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production. *Science*. 2014;344(6181):307–310.

71. Jungfleisch J, et al. A novel translational control mechanism involving RNA structures within coding sequences. *Genome Res*. 2017;27(1):95–106.

72. Beckham CJ, Parker R. P bodies, stress granules, and viral life cycles. *Cell Host Microbe*. 2008;3(4):206–212.

73. Moon SL, Wilusz J. Cytoplasmic viruses: rage against the (cellular RNA decay) machine. *PLoS Pathog*. 2013;9(12):e1003762.

74. Ni N, Morrow CD. Impact of forced selection of tRNAs on HIV-1 replication and genome stability highlight preferences for selection of certain tRNAs. *Virus Res*. 2007;124(1–2):29–37.

75. Piekna-Przybylska D, et al. A sequence similar to tRNA 3 Lys gene is embedded in HIV-1 U3-R and promotes minus-strand transfer. *Nat Struct Mol Biol*. 2010;17(1):83–89.

76. Landry S, et al. Detection, characterization and regulation of antisense transcripts in HIV-1. *Retrovirology*. 2007;4:71.

77. Kobayashi-Ishihara M, et al. HIV-1-encoded antisense RNA suppresses viral replication for a prolonged period. *Retrovirology*. 2012;9:38.

78. Divsalar DN, et al. Novel histone deacetylase inhibitors and HIV-1 latency-reversing agents identified by large-scale virtual screening. *Front Pharmacol*. 2020;11:905.

79. Darcis G, et al. HIV latency: should we shock or lock? *Trends Immunol*. 2017;38(3):217–228.

80. Cliffe AR, et al. Transcription of the herpes simplex virus latency-associated transcript promotes the formation of facultative heterochromatin on lytic promoters. *J Virol*. 2009;83(16):8182–8190.

81. Archin NM, Margolis DM. Emerging strategies to deplete the HIV reservoir. *Curr Opin Infect Dis*. 2014;27(1):29–35.

82. Beaumier CM, et al. CD4 downregulation by memory CD4⁺ T cells in vivo renders African green monkeys resistant to progressive SIVagm infection. *Nat Med*. 2009;15(8):879–885.

83. Mudd JC, et al. Epigenetic silencing of CD4 expression in nonpathogenic SIV infection in African green monkeys. *JCI Insight*. 2020;5(18):e139043.

84. Ziv O, et al. COMRADES determines in vivo RNA structures and interactions. *Nat Methods*. 2018;15(10):785–788.

85. Gomez JA, et al. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-γ locus. *Cell*. 2013;152(4):743–754.

86. Scheel TK, et al. A broad RNA virus survey reveals both miRNA dependence and functional sequestration. *Cell Host Microbe*. 2016;19(3):409–423.

87. Shimakami T, et al. Stabilization of hepatitis C virus RNA by an Ago2-miR-122 complex. *Proc Natl Acad Sci U S A*. 2012;109(3):941–946.

88. Nathans R, et al. Cellular microRNA and P bodies modulate host-HIV-1 interactions. *Mol Cell*. 2009;34(6):696–709.

89. Kincaid RP, et al. 2014. Noncanonical microRNA (miRNA) biogenesis gives rise to retroviral mimics of lymphoproliferative and immunosuppressive host miRNAs. *mBio*. 2014;5(2):e00074.

90. Madhugiri R, et al. Coronavirus cis-acting RNA elements. *Adv Virus Res*. 2016;96:127–163.

91. Goedeke L, et al. A regulatory role for microRNA 33* in controlling lipid metabolism gene expression. *Mol Cell Biol*. 2013;33(11):2339–2352.

92. Han SY, et al. MicroRNA-33a-3p suppresses cell migration and invasion by directly targeting PBX3 in human hepatocellular carcinoma. *Oncotarget*. 2016;7(27):42461–42473.

93. Periwal N, et al. In-silico analysis of SARS-CoV-2 genomes: insights from SARS encoded non-coding RNAs [preprint]. https://doi.org/10.1101/2020.03.31.018499. Posted on bioRxiv April 04, 2020.

94. Ziv O, et al. The short- and long-range RNA-RNA interactome of SARS-CoV-2. *Mol Cell*. 2020;80(6):1067–1077.e5.

95. Reyes CM, Kollman PA. Structure and thermodynamics of RNA-protein binding: using molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change. *J Mol Biol*. 2000;297(5):1145–1158.

96. Jens M, Rajewsky N. Competition between target sites of regulators shapes post-transcriptional gene regulation. *Nat Rev Genet*. 2015;16(2):113–126.

97. Denzler R, et al. Impact of microRNA levels, target-site complementarity, and cooperativity on competing endogenous RNA-regulated gene

expression. *Mol Cell*. 2016;64(3):565–579.

98. Ganesan A, et al. The timeline of epigenetic drug discovery: from reality to dreams. *Clin Epigenetics*. 2019;11(1):174.

99. Li E, et al. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*. 1992;69(6):915–926.

100. Yu J, et al. DNMT1-PPARγ pathway in macrophages regulates chronic inflammation and atherosclerosis development in mice. *Sci Rep*. 2016;6:30053.

101. Kim GH, et al. The role of redox signaling in epigenetics and cardiovascular disease. *Antioxid Redox Signal*. 2013;18(15):1920–1936.

102. Chen C, et al. DNA methyltransferases 1 and 3B are required for hepatitis C virus infection in cell culture. *Virology*. 2013;441(1):57–65.

103. Takahashi M, et al. Dual mechanisms of action of self-delivering, anti-HIV-1 FANA oligonucleotides as a potential new approach to HIV therapy. *Mol Ther Nucleic Acids*. 2019;17:615–625.

104. Kunden RD, et al. The role of the liver-specific microRNA, miRNA-122 in the HCV replication cycle. *Int J Mol Sci*. 2020;21(16):E5677.