Teaching an Old Virus New Tricks: A Review on New Approaches to Study Age-Old Questions in Influenza Biology

Seema S. Lakdawala [1], Nara Lee [1] and Christopher B. Brooke [2,3]

1 - Department of Microbiology and Molecular Genetics, *University of Pittsburgh, School of Medicine Pittsburgh, PA 15219, USA*
2 - Department of Microbiology, *University of Illinois at Urbana–Champaign, Champaign, IL 61801, USA*
3 - Carl R. Woese Institute for Genomic Biology, *University of Illinois at Urbana–Champaign, Champaign, IL 61801, USA*

Correspondence to Seema S. Lakdawala: *Lakdawala@pitt.edu, Nara.Lee@pitt.edu, cbrooke@illinois.edu*
https://doi.org/10.1016/j.jmb.2019.04.038

Abstract

Influenza viruses have been studied for over 80 years, yet much about the basic viral lifecycle remain unknown. However, new imaging, biochemical, and sequencing techniques have revealed significant insight into many age-old questions of influenza virus biology. In this review, we will cover the role of imaging techniques to describe unique aspects of influenza virus assembly, biochemical techniques to study viral genomic organization, and next-generation sequencing to explore influenza genomic evolution. Our goal is to provide a brief overview of how emerging techniques are being used to answer basic questions about influenza viruses. This is not a comprehensive list of emerging techniques, rather ones that we feel will continue to make significant contributions to field of influenza biology.

Introduction

Influenza viruses are a member of Orthomyxoviridae family of viruses. Seasonal and pandemic influenza viruses pose a large public health risk and cause tens of thousands of death worldwide on a yearly basis [1]. The first recorded influenza pandemic in our history was 100 years ago in 1918 [2]. Since then, influenza viruses have been the target of national vaccination strategies and research to understand the basis of disease.

The influenza viral genome is composed of eight negative sense RNA genome segments. Replication of the viral RNA (vRNA) segments occurs through a virally encoded RNA-dependent RNA polymerase in the nucleus of infected cells. Progeny virions bud from the plasma membrane where all eight segments are selectively packaged into a progeny virion [3]. Pandemic influenza viruses emerge through reassortment and intermingling of viral genomes within coinfected cells. Many aspects of entry, replication, reassortment, packaging, and evolutionary dynamics are still unknown after years of study. In this review, we will cover the use of emerging technologies in microscopy, RNA biology, and next-generation sequencing (NGS) to study important basic aspects of influenza biology from intracellular vRNA assembly to viral evolution.

New Imaging Approaches

Many aspects of the dynamic viral life cycle, particularly within a cell, have been difficult to examine. A lack of resolution both in time and in space have limited our understanding of the intracellular relationships that facilitate assembly of influenza vRNA within a cell. However, many advances within imaging with both new techniques and optical hardware have greatly enhanced our ability to study age-old questions regarding influenza virus assembly.

Recent advances in microscopy techniques have successfully been used to study influenza virus assembly and intracellular dynamics. We will review a few of the emerging techniques and how these tools have advanced our understanding of influenza virus assembly.

Visualization of vRNA segments

The intracellular expression and localization of influenza vRNA segments have be accomplished using fluorescent *in situ* hybridization (FISH) through a variety of probe formats [4–6]. Early studies used long probes targeting a subset of the vRNA segments or conformational antibodies [4,7]. However, long probes require high concentrations of a target to provide a detectable signal given the low labeling density of the probe [8]. An alternative approach was developed to use small oligos of 18–20 nucleotides, each conjugated to a fluorophore, that tile the length of a RNA sequence which allow for single-molecule imaging capability [9]. Single molecule FISH probes have been used to examine the intracellular localization of influenza vRNA segments [5,6,10,11]. Lakdawala and colleagues [6,11] have successfully multiplexed multiple fluorophores to visualize 3–4 vRNA segments within a single cell, providing the ability to compare the localization of multiple segments in relationship to each other as well as cellular features.

One drawback of single-molecule FISH approaches is the inability to distinguish closely related vRNA species, such as gene segments with single nucleotide polymorphisms. An alternative approach using padlock probes was recently used to distinguish vRNA segments between two influenza viruses, which differed in a single point mutation on each segment [12]. In this technique, unique primers are used to amplify a single vRNA segment with sequence specificity and multiple barcode regions for the incubation with fluorescent oligos. Ultimately, padlock probes create a circular DNA structures after replication of an initial probe is bound; thus, detection of a single molecule is amplified significantly for robust detection. The use of multiple barcodes can provide granularity and has been used to differentiate between specific segments from multiple strains [13]. Based on these benefits, padlock probes have been used successfully to examine the temporal regulation of vRNA gene expression during co-infection [12]. This strategy was used effectively to detect vRNA segments from entry through replication. In addition, the ability to distinguish between strains allowed for in-depth analysis of coinfection frequencies. Interestingly, padlock probes were used to detect vRNA segments from two different strains within co-infected cells, and the authors observed that coinfection frequency was restricted at the stage of vRNA import into the nucleus, rather than at entry or genome replication [12].

Multicolor imaging and mathematical modeling

Multicolor imaging is critical to assess the relationships between diverse viral components. It can be achieved through spectral separation of fluorophores with microscopy techniques such as an acoustic optical beam splitter[14] or spectral unmixing [15]. The acoustic optical beam splitter in a white light laser system was used to detect multiple influenza vRNA segments within a single infected cells [6]. These studies allow for the analysis of the intracellular localization of vRNA segments and a glimpse into the relationship between vRNA segments. Multicolor FISH of up to four segments has been used to determine that influenza vRNA segments are exported from the nucleus as subcomplexes containing more than two but less than eight segments, not individually or as a full complex of eight vRNA segments [6]. In addition, this technique was used to examine the relationship of vRNA and host components, such as Rab11A, a small GTPase, and the nucleus or plasma membrane [11].

Segmental interactions between influenza vRNA are thought to mediate selective packaging (reviewed further in the next section) [16]. Therefore, visualization of multiple vRNA segments within a single cell provides an opportunity to examine the intracellular relationship between vRNA segments. To extrapolate vRNA interaction networks using multicolor FISH, work was initiated to use point process algorithms to model the intracellular relationship between vRNA segments sampling only pairs of vRNA segments, triples, or quadruples [17]. Dynamic minimal distance networks using learned point process models were created to examine the relationship between vRNA segments, and found a clear order of assembly between segments [17]. Figure 1 highlights the pipeline that has been developed to use multicolor imaging data to define vRNA assembly networks and provides the most likely network developed from this process.

Other approaches for detection of multiple viral components have also been used to examine aspects of the viral lifecycle. Multicolor imaging of multiple viral proteins on a single virion was recently been developed to assess the size of virions produced during single cell infections [18]. This approach found heterogeneity within virion shape and variation in surface glycoprotein expression suggesting that a phenotypic distribution of virions is produced from a single infection to survive the complex respiratory environment [18]. Overall, multicolor imaging approaches are helping define many aspects of the viral lifecycle and these approaches can be modulated to answer a wide variety of questions.

Live cell imaging

Conclusions based on the FISH-based static imaging approaches described above must be confirmed using other methods. Detection and tracking of vRNA in live cells is an emerging area

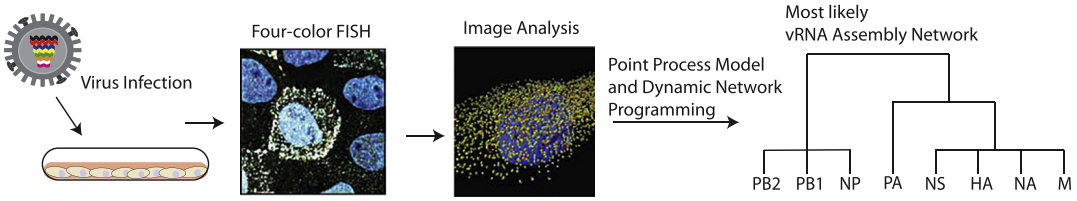

Fig. 1. Generation of vRNA interaction networks using multi-color imaging and machine learning. The intracellular localization of four vRNA segments within a single cell was previously determined using confocal microscopy [6]. Recently, point-process models and dynamic network programing algorithms were used to decipher a potential vRNA–vRNA interaction network [17].

of research and classically requires the use of modified RNA molecules that include conjugated fluorophores or introduction of aptamer repeats such as the MS2 hairpin loops, to the RNA (reviewed in Ref. [19]). Aptamers have successfully been incorporated into the influenza viral genome for purification purposes [20], and this strategy could be adapted for imaging studies. However, the need to alter the vRNA genome may have unwanted impacts on the viral lifecycle including replication fitness and packaging efficiency. An alternative strategy is multiply-labeled tetravalent RNA imaging probes, which uses a streptavidin bead to cluster fluorescent oligos together and amplifies the fluorescent signal; this approach can be used to detect endogenous RNA within live cells [21]. While this technique has been used effectively for detection and tracking of the RSV genome, it is has not yet been used to detect influenza vRNA. Emerging techniques for live cell RNA FISH are also being developed and could be applied to influenza vRNA assembly research.

Many groups have used fluorescently tagged proteins in the viral polymerase as a surrogate for tracking vRNA directly in live cells. The viral genome is bound at the 3′ and 5′ end with three components of the viral polymerase (PB2, PB1, and PA) [22]. Conjugation of fluorescent proteins to either PB2 or PA has been developed to visualize vRNA transport dynamics in live cells [23,24]. These studies revealed that vRNA segments undergo fusion and colocalization events en route to the plasma membrane for packaging [24]. Conjugation of virions with fluorescent dyes has also been used to examine influenza virus entry mechanisms [25]. Recently, a group used quantum dots to examine viral uncoating by labeling both the viral polymerase PA protein and viral membrane [26]. This study found that nuclear import of incoming vRNA segments occurs individually rather than as a complex of eight segments [26].

Live cell detection of RNA is only one piece of the puzzle; image acquisition at high spatio-temporal resolution is necessary to accurately examine the transport dynamics of vRNA transport. Light sheet microscopy is providing an important advancement to the temporal acquisition of live cells [27]. Light sheet microscopy is a technique where only a thin slice of the sample is illuminated perpendicular to the detection objective, thus reducing background. These systems impart low-phototoxicity providing a high signal-to-noise ratio compared to traditional confocal imaging platforms, making them ideal to imaging dynamic movement in live cells. Multiple modifications to classical light sheet systems have been made to address specific biological needs, such as inverted selective plane illumination (SPIM [28],) and lattice light sheet systems [29]). SPIM has been used previously to describe the intracellular movement of influenza vRNA, using a fluorescent virus encoding a GFP-tagged polymerase component, within infected MDCK and A549 cells [24]. The high temporal resolution provided information on the dynamics of vRNA transport and observed frequent fusion and fission events during the transport of vRNA segments form the nucleus to the plasma membrane. In addition, advancements to the SPIM system include a dual-view component (diSPIM) that uses two equivalent objectives on each arm of the diSPIM to provide isotropic resolution in x, y and z planes [30]. Analysis of influenza vRNA dynamics is underway using the diSPIM [31].

Multicolor microscopy has many benefits, and its use has identified many important aspects of influenza virus assembly. However, it is important to understand the limitations of microcopy tools. All of the tools described here, excluding super resolution imaging, are constrained by the limit of resolution (LOR). For conventional confocal microscopy, the resolution can be approximated by the following equation:

$$\text{Resolution } (r) = \lambda/2^{*}\text{NA}$$

where λ is the wavelength used for excitation, that is, 488, 561, and so on, and NA is the numerical aperture of the objective used for image acquisition. A number of factors can impact LOR, including optic alignments. Typically, in the real world, a maximum resolution is 0.25 μm, or 250 nm. Thus, the LOR on a confocal is larger than the size of influenza particles (80–100 nm) and isare unable to resolve direct interactions between vRNA segments without the use of FRET or other approaches. Light sheet microscopy can overcome the temporal obstacle to

live cell imaging, but it is still constrained by LOR. Current light sheet systems, like the diSPIM, have isotropic resolution of ~330 nm [30].

In summary, emerging imaging techniques with advances in resolution and speed will enhance influenza virus research by providing important insight into the movement of viral proteins and vRNA within cells. These types of studies will contribute to the development of novel antiviral strategies that target host proteins to limit the intracellular replication and spread of influenza viruses in cells.

New Biochemical Approaches to Dissect vRNA Biology

While light microscopy has and continues to be central in expanding our understanding of vRNA behaviors within the cell, our perception of influenza virus segment architecture within the virion has mainly been shaped by electron microscopy (EM) studies [32–34], with subsequent investigations further refining our insight [35,36]. Each of the eight vRNA segment is proposed to form an anti-parallel helical structure that associates with the trimeric viral polymerase at the 5′ and 3′ termini, which hybridize with each other to form the so-called "pan-handle" structure, and with a viral nucleoprotein (NP) scaffold along the entire body of the segment. This arrangement is colloquially referred to as the "beads-on-a-string" conformation. Further EM studies have indicated that segments within packaged virions are connected to each other [16,37]. However, given the resolution limit of EM, the identity and precise regions of the segments that form these connections remain unknown, necessitating more sensitive techniques to address this question. Recent advances in RNA methodology have allowed us to study RNA interactions with proteins as well as other RNA molecules with unprecedented sensitivity and resolution. These techniques are now being applied to influenza virology and are expected to complement previous structural studies and provide novel insight into the segment architecture.

vRNA–protein interaction

In recent years, "old school" RNA methods have been coupled with NGS technology to study various aspects of RNA interactions with other macromolecules at nucleotide resolution. Given the fact that the genome of Influenza virus consists of negative-sense single-stranded RNA, these novel RNA-centric techniques are perfectly suited to be applied to influenza virus and further our insight into its genome organization. One such method called CLIP (UV-*c*ross*l*inking and *i*mmuno*p*recipitation) investigates RNA–protein interactions by utilizing UV light irradiation to covalently link RNA to protein [38]. This crosslinking event occurs only within a range of ~1 Å and thus exclusively crosslinks direct interactors, unlike the commonly used crosslinking agent formaldehyde, which has a lower stringency and a wider crosslinking range. Two steps in the CLIP protocol achieve high-confidence identification of *bona fide* RNA–protein interactions: (1) Following UV irradiation prior to immunoprecipitation, partial RNase digest is performed to degrade abundant contaminating RNAs and produce a specific footprint of the associating RNAs, which when coupled to NGS allows for identification of all RNA interactions on a transcriptome-wide scale. (2) Following immunoprecipitation, the RNA–protein adducts are resolved by SDS-gel electrophoresis, which survive the harsh buffer conditions given their covalent linkage, and are subsequently purified by size-selection. The latter step addresses the issue of RNA-reassociation after cell (or virion) lysis, which is a major concern in RNA biology, as RNAs have been shown to artificially bind to normally non-associating proteins following lysis preparation [39].

Several variations of high-throughput CLIP protocols have been developed, such as HITS-CLIP (*Hi*gh-*t*hroughput *s*equencing of RNA isolated by *c*ross*l*inking *i*mmuno*p*recipitation, also known as CLIP-Seq), iCLIP (*i*ndividual-nucleotide resolution CLIP) or eCLIP (*e*nhanced CLIP) [40], which mainly differ in how the sequencing library is prepared at the final step of the experimental protocol. As UV crosslinking of RNA to protein is notoriously inefficient, as merely 1%–5% of proteins are estimated to form crosslinks [41], photoreactive nucleoside analogues, such as 4-thiouridine, have been incorporated in the CLIP protocol to markedly boost crosslinking efficiency; this approach is referred to as PAR-CLIP (photoactivatable ribonucleoside-enhanced CLIP) [42]. A common caveat to all CLIP variations is that the UV-crosslinking process depends on the right geometry between RNA and interacting protein, as the RNA base to be crosslinked needs to be spatially arranged in a suitable angle to the amino acid residue for crosslinking to occur, and thus, some true RNA binding sites may escape detection. Given this caveat, the most comprehensive study would entail a comparative analysis of conventional CLIP, PAR-CLIP, and RNA-IP assays, as all of these approaches have their own strengths and weaknesses [40,41].

Recently, our groups performed HITS-CLIP to examine the sites of vRNA associated with NP in the vRNP scaffold in virions [43,44]. Of the available CLIP protocols, we chose HITS-CLIP in our experiments due to the fact that this approach reveals the entire binding site covered by NP, as opposed to other CLIP variants, such as iCLIP or eCLIP, which mainly uncover the crosslinked site of RNA to the bound protein. Applying HITS-CLIP of NP to isolated

virions, we have observed that vRNA is not uniformly bound by NP as predicted by the "beads-on-a-string" model. Instead, several regions throughout the viral genome appear to be free of NP association and others show a non-regular binding of NP across the genome. A report using PAR-CLIP for NP on influenza virus-infected host cell lysate reached a similar conclusion that NP is not ubiquitously present on the vRNA segments [45]. Comparative analysis between both data sets is complicated by the fact that different compartments were examined (cell lysate *versus* virions) and that different CLIP approaches were applied (HITS-CLIP *versus* PAR-CLIP). However, both studies arrive at the same conclusion that the "beads-on-a string" conformation may be oversimplified and that the classical architecture of the influenza vRNP should be revised (Fig. 2). This observation is surprising in light of previous studies that demonstrated that NP binds in a sequence-independent manner [46,47].

All CLIP-based approaches provide a snap-shot of RNA–protein interactions by averaging millions of RNA footprints and do not deliver any information regarding the dynamic nature of these interactions. Thus, it remains unclear how the newly synthesized vRNA strands in the nucleus associate with NP molecules to form the distinct NP binding profile observed in virions that contains NP-free as well as NP-enriched regions. It can be assumed that newly synthesized vRNA segments bind not only to NP molecules but also to other host RNA-binding proteins that abound in the nucleus at high concentrations, many of which bind RNA in a non-discriminatory manner without a cognate binding motif [48,49]. Therefore, vRNA regions identified not to be associated with NP in our HITS-CLIP data could be (i) free single-stranded regions, (ii) double-stranded regions, or (iii) regions bound by proteins other than NP. It is unlikely that the NP-free regions are indeed single-stranded, as many putative partners, most notably a plethora of cellular RNA-binding proteins, would be awaiting to undergo non-specific interactions. Certain genomic regions may act as structural elements, such as conserved stem-loops, pertinent to vRNP assembly and therefore not be associated with NP [50,51], and the possibility of NP-free regions forming sites of double-stranded helices is further discussed below. Moreover, some of the NP-free regions may indeed associate with other cellular proteins. For example, it was recently shown that the host RNA-binding protein DAI (DNA-dependent activator of IFN-regulatory factors) associates with vRNA inside infected cells [52], and it remains to be tested whether this interaction persists upon virion assembly. Furthermore, proteomics studies indicated that a number of host proteins are detected in virions [53], insinuating that some of these host factors may be an integral part of vRNPs and their detection not be a result of accidental packaging into virions. Taken together, future studies examining vRNA–protein interactions in greater detail will provide further insight into how many nuances of the textbook "beads-on-a-string" model of vRNP architecture need to be modified.

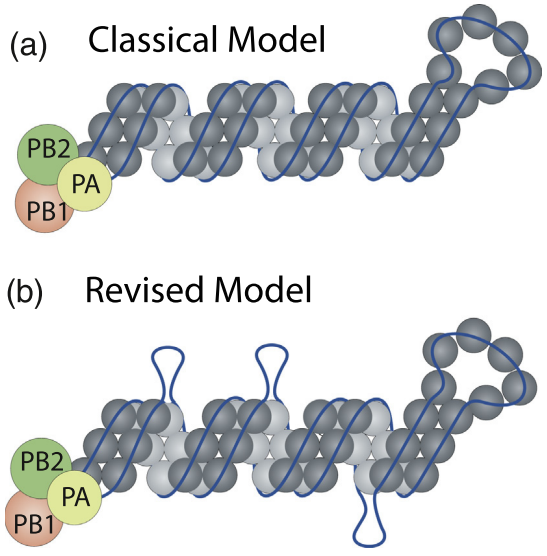

Fig. 2. Redefining the architecture of influenza vRNA. (a) The classical model of influenza vRNA is shown where NP (gray sphere) binds uniformly to the entire length of the vRNA segment. (b) A revised model based on RNA biochemistry technique HITS-CLIP. NP was found not to bind to vRNA in a uniform, random manner [43]. Certain vRNA areas are enriched with NP binding, and others are free of NP association.

vRNA–vRNA interaction

EM studies and *in vitro* RNA–RNA interaction studies suggest that some of the eight vRNA segments interact with each other to form RNA–RNA bridges [16,54,55]. Such crosstalk between segments is further supported by a recent report that examined the genome packaging mechanism by modulating packaging sequences on different segments [56]. Indeed, the formation of an intersegmental RNA network would greatly facilitate the challenging task of coordinating selective packaging of all eight segments into a single infectious virion.

Many cellular processes that involve RNA molecules entail RNA–RNA interactions as their basis of molecular mechanism [57,58]. In practice, these RNA–RNA interactions are challenging to uncover experimentally, in part due to the intrinsic intricacy of manipulating RNA, as opposed to DNA. Moreover, computational algorithms deliver only sub-optimal predictions due to the fact that RNA–RNA interactions can occur in several ways, such as canonical Watson–Crick interactions, non-Watson–Crick

interactions between bases, as well as base stacking interactions between single-stranded regions [59]. Furthermore, intersegmental interactions consisting of interspersed multiple short duplexes, partial matches containing non-canonical base pairs, in addition to base stacking, all of which could be concentrated within a given vRNA region to stabilize this particular RNA–RNA bridge, would likely evade detection by a standard prediction algorithm due to its complexity.

To experimentally elucidate RNA–RNA interactions at a transcriptome-wide scale, several methods have recently been described that leverage the sensitivity of NGS technology. These utilize the reversible crosslinking agent psoralen [60], which can generate covalent crosslinks between adjacent pyrimidine bases of opposing double-stranded RNA duplexes when irradiated with long-wave UV light. The covalent linkage allows for stringent washing conditions to lower the false-positive rates and can be subsequently reversed by irradiation with short-wave UV light, which is normally conducted prior to the reverse transcription step during library preparation. The first-generation protocols employed an antisense oligonucleotide targeting a specific RNA of interest in their experimental outline to capture the RNA–RNA interactions formed by a particular RNA only to reduce the complexity of the sequencing library [61–63]. To infer global RNA interaction maps, "shotgun" approaches have then been described [64–67], which require more extensive computational analyses of the resulting NGS data sets. All shotgun approaches, referred to as PARIS (*p*soralen *a*nalysis of *R*NA *i*nteractions and *s*tructures), LIGR-seq (*lig*ation of interacting *R*NA followed by high-throughput *seq*uencing), or SPLASH (*s*equencing of *p*soralen crosslinked, *l*igated, *a*nd *s*elected *h*ybrids), underlie a comparable concept by entailing psoralen crosslinking in combination with intermolecular ligation based on spatial proximity followed by deep-sequencing library preparation, yet differ in how the RNA hybrids are enriched. The functionality of all three experimental outlines has been demonstrated by studying highly abundant cellular noncoding RNAs, which are known to adopt known secondary structures [64–67]. These shotgun approaches are now being applied to influenza virus to examine the putative intersegmental RNA–RNA interactions at a genome-wide scale (bioRxiv doi.org/10.1101/236620).

It should be noted that a major drawback of psoralen-dependent techniques is that RNA–RNA crosslinking is contingent upon the right geometry of RNA duplexes, as not all double-stranded regions are amenable to psoralen-crosslinking, and thus, some interactions may be missed. For this reason, when examining the formation of a particular RNA duplex, the efficacy of psoralen crosslinking on this duplex is typically verified beforehand using *in vitro*-transcribed RNAs [68], which would be impracticable for studying Influenza virus. Furthermore, psoralen crosslinking of RNA duplexes, similar to UV light-mediated crosslinking of RNA–protein complexes, is highly inefficient requiring extended UV light exposure of up to 30 min, and reversal of crosslinking by irradiation with short-wave UV light increases the risk of damaging RNA strands and renders them unusable for ensuing reverse transcription. Moreover, as the exact nature of RNA duplex formation between vRNA segments remains unknown, these psoralen-based approaches may not be appropriate for generating a detailed interactome. For example, vRNA–vRNA bridges that encompass a series of interspersed short stretches of duplexes within an interacting region or that are formed only transiently, as they are required only during the initiation step of segment interaction and are then stabilized by other types of interaction, may escape capture. Another concern is that the proximity ligation step relies on the T4 RNA ligase enzyme, which is notoriously inefficient and normally requires a splint oligonucleotide to boost the ligation efficiency. Ideally, the RNA-equivalent of chromosome conformation capture (3C) [69], a DNA methodology designed to probe for DNA–DNA interactions within chromatin, would be best suited to uncover the three-dimensional viral genome organization. 3C utilizes formaldehyde crosslinking and restriction enzyme digest followed by proximity ligation. However, as the RNA-equivalent tool kits, such as restriction enzymes, are not available, several technical hurdles need yet to be resolved before this concept can be applied to the RNA genome of Influenza virus.

In summary, these promising new techniques to study RNA–RNA interactions in combination with CLIP data outlining the genome-wide NP binding profile will provide intriguing insight into the architecture and interconnections of vRNPs and how the intersegmental interplay is integral to viral genome assembly in virions.

New Sequencing Approaches

In addition to providing new insight into structural and biochemical details of influenza virus biology, new technologies are also allowing us to tackle long standing questions in viral genetics. Like other RNA viruses, influenza viruses exist as genetically diverse populations due to the high mutation rate of the viral polymerase. Recent studies suggest that most, if not nearly all, virions carry at least one random mutation, meaning that viral populations consist of swarms of low-frequency minor sequence variants [70]. This genetic diversity provides the raw material for adaptation, as pre-existing variants can facilitate escape from new selective pressures or expansion into new host environments. These mutant swarms are highly dynamic, as individual

variants fluctuate in frequency due to the combined effects of selection and genetic drift. Tracking these dynamics, which can be observed *in vitro*, *in vivo*, and at the epidemiological scale, can provide enormous insight into virus biology and evolutionary dynamics.

Over the last several years, there has been a revolution in our ability to observe and quantify changes in the viral mutational landscape through the rapid development of NGS methods. These advances in viral population sequencing capabilities have fueled a wave of discoveries surrounding influenza virus evolution and transmission. In addition, new NGS-enabled experimental approaches have opened up new frontiers in our understanding of the basic biology of these viruses.

Advances in NGS technology

The vast majority of viral NGS is now carried out using the Illumina platform, which is both widely available and highly robust and versatile. The Illumina platform can be used to detect minor sequence variants that exist at frequencies as low as ~1% within a viral population, and can easily be multiplexed to examine many individual samples in a single sequencing run [71,72]. The ability to generate whole-genome viral sequence data in a highly multiplexed fashion, combined with advances in the detection and amplification of diverse viral genotypes, has greatly enabled influenza surveillance efforts [73,74]. The primary downside of this platform is that the read lengths are very short, from 50 to 600 bp (shorter than any of the viral gene segments) depending on the specific approach and sequencing chemistry used, meaning that information about full-length segment haplotypes is typically lost.

Multiple approaches have been developed to enhance the sensitivity and accuracy of Illumina-based sequencing. Recently, McCrone and Lauring [75] described the development of experimentally validated best practices for library generation and data filtration that maximize the sensitivity and specificity of minor variant detection. Other groups have enhanced the sensitivity of Illumina sequencing further by decreasing the base calling error rate through the incorporation of unique molecular barcodes during cDNA synthesis, a method commonly known as primerID [76,77]. Pauly *et al*. [78] used this approach in an effort to more precisely quantify the viral mutation rate. Primer ID has the added benefit of permitting the reconstruction of viral haplotypes, information that is generally lost during the fragmentation and PCR steps of standard Illumina sequencing. Kosik *et al*. [79] used this capability to examine the interplay between emerging variants in HA, revealing a strong role for clonal interference in governing the emergence of new antigenic variants.

Two other sequencing platforms have recently emerged that overcome one of the main practical limitations of Illumina by generating much longer reads that can span the entirety of the viral genome segments [80]. The Pacific Biosciences (PacBio) platform is based on the direct imaging of base additions within single molecules and can achieve read lengths of tens of thousands of base pairs while retaining individual haplotypes. Rogers *et al*. [81] took advantage of these features to describe the dynamics of drug resistance during chronic influenza virus infection in single hosts. Combining viral template circularization with the long reads generated during PacBio sequencing allows for the repeated sequencing of individual molecules multiple times during a single run, significantly reducing the error rate [82]. This approach, known as circular consensus sequencing, was recently reported for influenza virus for the first time and has also been used to examine hepatitis C virus within-host evolution [83,84].

More recently, Oxford Nanopore Technologies (ONT) has developed a new sequencing platform based on passing nucleic acid molecules through tiny pores ("nanopores," if you will) and determining base identity based on current fluctuations across the pore. The ONT can generate long-read, single-molecule reads, similar to PacBio, but also does so in a highly accessible and relatively affordable package [85]. In particular, the low cost and portability of the ONT MinION may facilitate the expansion of influenza surveillance capabilities in areas where other NGS technologies are not readily available. The ONT is unique among NGS platforms in that it also allows for the direct sequencing of RNA, without the need for PCR amplification [86]. This approach was recently used to perform the first full-genome RNA sequencing of influenza virus [87]. The ability to directly sequence viral genomic RNA may facilitate future studies of RNA splicing, base modification, and defective interfering RNA formation during influenza virus infection.

Genotype–phenotype mapping

One of the most exciting areas in which NGS is advancing the field is in allowing us to explore basic questions in a much more comprehensive and systematic way than ever possible before. One prominent example is the proliferation of studies that use deep mutational scanning (DMS) to examine the effects of all possible amino acid substitutions on influenza virus protein function. In contrast with the historical approach of individually generating mutants and comparing effects of viral growth side by side, DMS allows investigators to generate swarms of all possible substitutions through PCR and then collectively phenotype all mutants using NGS-based screens. This approach was first used to examine

the mutational tolerance of the HA protein, revealing that the head domain (at least for some H1 viruses) may be more tolerant of mutations than the stem region [88,89]. These *in vitro* analyses of mutational tolerance can now be incorporated into phylogenetic models to better estimate the evolutionary dynamics of viral proteins over time at the epidemiological scale [90]. Subsequent studies have used DMS to comprehensively identify residues involved in escape from neutralizing antibodies, interferon sensitivity, and polymerase function [89,91–93].

Single-cell virology

The combination of NGS with recent advances in microfluidics technology has finally allowed us to move beyond bulk studies to examine viral infection at the single-cell level through single-cell RNAseq. By partitioning individual cells within liquid droplets surrounded by oil, it is possible to tag all cDNAs with cell-specific barcodes that identify the specific cell that its parent RNA molecule was isolated from. These barcodes can be used to reconstruct the paired host and viral transcriptomes from thousands of individual cells, revealing patterns of heterogeneity that are lost during bulk analysis. Russell *et al*. [94] recently used this approach to examine the extent of heterogeneity present during influenza infection, revealing a surprising degree of variation in overall levels of viral transcription between individual cells, even when viral input was largely normalized. Interestingly, they observed that only a tiny percentage of infected cells actually upregulated type I interferon expression, highlighting the need to better understand the role of stochasticity in governing host–virus interactions. The field of single-cell RNAseq is still very young, and the bioinformatics tools used to analyze the enormous NGS data sets involved are still in their infancy; however, this powerful technology will likely enable many more exciting discoveries in the future [95].

Analysis approaches

Many of the advances in viral sequencing have come from the development of new analysis tools and statistical methods, rather than advances in sequencing hardware. In particular, the application of novel statistical analyses and modeling approaches have allowed researchers to use clinical NGS data to address critical questions about the transmission and evolution of influenza viruses in humans for the first time.

One of the biggest questions concerning influenza transmission in humans is the size of the virus population transmitted during natural infection. The size of this transmission bottleneck has important implications for understanding viral evolutionary dynamics, but is very hard to study in a clinical setting. Sobel Leonard *et al*. [96] developed a novel approach for estimating bottleneck sizes based on beta-binomial sampling of NGS data from known transmission pairs while accounting for the variation in variant calling thresholds and stochastic replication dynamics that may have skewed previous estimates. McCrone *et al*. [97,98] applied the same beta-binomial sampling method to a more rigorously generated clinical NGS data set and estimated a mean bottleneck size of one to two viral genomes, much smaller than previous estimates but in line with what has been observed during airborne transmission in animal models.

Another big question concerns the extent of reassortment that occurs during natural infection. Studies in laboratory animals have suggested that reassortment is widespread; however, it has been impossible to do the same sorts of experiments in humans [99]. Sobel Leonard *et al*. [100] developed a model that allowed them to estimate the extent of reassortment in humans using clinical NGS data. They measured the degree to which specific allele combinations were maintained across genome segments in longitudinal NGS samples, and observed results consistent with a very low frequency of effective reassortment in humans. NGS-enabled analysis of infection and evolutionary dynamics in humans will continue to be an area of active development as the number and quality of clinical NGS data sets increases, and as more sophisticated statistical analyses are developed and implemented.

Overall Summary

Technology development in microscopy tools, RNA-biology techniques, and sequencing pipelines are dramatically altering the landscape of influenza research. These new approaches will shape upcoming discoveries on fundamental properties of the influenza virus lifecycle and answer age-old questions. The cross-pollination of disciplines ranging from chemistry, virology, and physics to bioinformatics are setting the stage for in-depth exploration of basic viral properties at a resolution previously unimagined.

Acknowledgments

S.S.L. is funded by NIH grant (1R01AI139063-01A1) and the American Lung Association Biomedical Research grant. Both S.S.L. and N.L. are funded through a New Initiative Award from the Charles E. Kaufman Foundation, a supporting organization of The Pittsburgh Foundation. C.B.B. is funded by an NIH grant (1R01AI139246-01A1), a DARPA contract (DARPA-16-35-INTERCEPT-FP-018), and a gift from the Roy J. Carver Charitable Trust (17-4905).

Received 24 January 2019;
Received in revised form 12 April 2019;
Accepted 23 April 2019
Available online 30 April 2019

Keywords:
influenza virus;
microscopy;
RNA−protein interactions;
RNA−RNA interactions;
next-generation sequencing

Abbreviations used:
vRNA, viral RNA; NGS, next-generation sequencing; FISH, fluorescent *in situ* hybridization; SPIM, selective plane illumination microscope; LOR, limit of resolution; EM, electron microscopy; NP, nucleoprotein; PacBio, Pacific Biosciences; ONT, Oxford Nanopore Technologies; DMS, deep mutational scanning.

References

[1] Centers for Disease Control and Prevention, N. C. f. I. a. R. D. N, Disease burden of influenza. https://www.cdc.gov/flu/about/disease/burden.htm>, 2017.
[2] D.M. Morens, J.K. Taubenberger, Pandemic influenza: certain uncertainties, Rev. Med. Virol. (2011), https://doi.org/10.1002/rmv.689.
[3] B.N. Fields, D.M. Knipe, P.M. Howley, Fields Virology, 6th edn, Wolters Kluwer/Lippincott Williams & Wilkins Health, 2013.
[4] M.C. Arcangeletti, et al., Host-cell-dependent role of actin cytoskeleton during the replication of a human strain of influenza A virus, Arch. Virol. 153 (2008) 1209−1221, https://doi.org/10.1007/s00705-008-0103-0.
[5] Y.Y. Chou, et al., Colocalization of different influenza viral RNA segments in the cytoplasm before viral budding as shown by single-molecule sensitivity FISH analysis, PLoS Pathog. 9 (2013), e1003358, https://doi.org/10.1371/journal.ppat.1003358.
[6] S.S. Lakdawala, et al., Influenza A virus assembly intermediates fuse in the cytoplasm, PLoS Pathog. 10 (2014), e1003971, https://doi.org/10.1371/journal.ppat.1003971.
[7] A.J. Eisfeld, E. Kawakami, T. Watanabe, G. Neumann, Y. Kawaoka, RAB11A is essential for transport of the influenza virus genome to the plasma membrane, J. Virol. 85 (2011) 6117−6126, https://doi.org/10.1128/JVI.00378-11.
[8] J.M. Levsky, R.H. Singer, Fluorescence in situ hybridization: past, present and future, J. Cell Sci. 116 (2003) 2833−2838, https://doi.org/10.1242/jcs.00633.
[9] M. Batish, A. Raj, S. Tyagi, Single molecule imaging of RNA in situ, Methods Mol. Biol. 714 (2011) 3−13, https://doi.org/10.1007/978-1-61779-005-8_1.
[10] Y.Y. Chou, et al., One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 9101−9106, https://doi.org/10.1073/pnas.1206069109.
[11] E. Nturibi, A.R. Bhagwat, S. Coburn, M.M. Myerburg, S.S. Lakdawala, Intracellular colocalization of influenza viral RNA and Rab11A is dependent upon microtubule filaments, J. Virol. (2017), https://doi.org/10.1128/JVI.01179-17.
[12] D. Dou, et al., Analysis of IAV replication and co-infection dynamics by a versatile RNA viral genome labeling method, Cell Rep. 20 (2017) 251−263, https://doi.org/10.1016/j.celrep.2017.06.021.
[13] F. Neumann, et al., Padlock probe assay for detection and subtyping of seasonal influenza, Clin. Chem. 64 (2018) 1704−1712, https://doi.org/10.1373/clinchem.2018.292979.
[14] R.T. Borlinghaus, The white confocal, Eur. Phys. J. Plus 127 (2012), https://doi.org/10.1140/epjp/i2012-12131-x.
[15] M.E. Dickinson, G. Bearman, S. Tille, R. Lansford, S.E. Fraser, Multi-spectral imaging and linear unmixing add a whole new dimension to laser scanning fluorescence microscopy, Biotechniques 31 (2001) 1272, https://doi.org/10.2144/01316bt01 (1274−1276, 1278).
[16] E. Fournier, et al., A supramolecular assembly formed by influenza A virus genomic RNA segments, Nucleic Acids Res. 40 (2012) 2197−2209, https://doi.org/10.1093/nar/gkr985.
[17] T.D. Majarian, R.F. Murphy, S.S. Lakdawala, Learning the sequence of influenza A genome assembly during viral replication using point process models and fluorescence in situ hybridization, PLoS Comput. Biol. 15 (2019), e1006199, https://doi.org/10.1371/journal.pcbi.1006199.
[18] M.D. Vahey, D.A. Fletcher, Low-fidelity assembly of influenza A virus promotes escape from host cells, Cell 176 (2019), https://doi.org/10.1016/j.cell.2018.10.056 (281−294 e219).
[19] E.A. Specht, E. Braselmann, A.E. Palmer, A critical and comparative review of fluorescent tools for live-cell imaging, Annu. Rev. Physiol. 79 (2017) 93−117, https://doi.org/10.1146/annurev-physiol-022516-034055.
[20] A. York, E. Fodor, Biogenesis, assembly, and export of viral messenger ribonucleoproteins in the influenza A virus infected cell, RNA Biol. 10 (2013) 1274−1282, https://doi.org/10.4161/rna.25356.
[21] P.J. Santangelo, et al., Single molecule-sensitive probes for imaging RNA in live cells, Nat. Methods 6 (2009) 347−349, https://doi.org/10.1038/nmeth.1316.
[22] A.J. Te Velthuis, E. Fodor, Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis, Nat. Rev. Microbiol. 14 (2016) 479−493, https://doi.org/10.1038/nrmicro.2016.87.
[23] S.V. Avilov, et al., Replication-competent influenza A virus that encodes a split-green fluorescent protein-tagged PB2 polymerase subunit allows live-cell imaging of the virus life cycle, J. Virol. 86 (2012) 1433−1448, https://doi.org/10.1128/JVI.05820-11.
[24] S.S. Lakdawala, et al., Correction: influenza A virus assembly intermediates fuse in the cytoplasm, PLoS Pathog. 12 (2016), e1006121, https://doi.org/10.1371/journal.ppat.1006121.
[25] M. Lakadamyali, M.J. Rust, H.P. Babcock, X. Zhuang, Visualizing infection of individual influenza viruses, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 9280−9285, https://doi.org/10.1073/pnas.0832269100.
[26] C. Qin, et al., Real-time dissection of dynamic uncoating of individual influenza viruses, Proc. Natl. Acad. Sci. U. S. A. (2019), https://doi.org/10.1073/pnas.1812632116.
[27] R.M. Power, J. Huisken, A guide to light-sheet fluorescence microscopy for multiscale imaging, Nat. Methods 14 (2017) 360−373, https://doi.org/10.1038/nmeth.4224.

[28] Y. Wu, et al., Inverted selective plane illumination microscopy (iSPIM) enables coupled cell identity lineaging and neurodevelopmental imaging in Caenorhabditis elegans, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 17708–17713, https://doi.org/10.1073/pnas.1108494108.
[29] A.M. Valm, et al., Applying systems-level spectral imaging and analysis to reveal the organelle interactome, Nature 546 (2017) 162–167, https://doi.org/10.1038/nature22369.
[30] A. Kumar, et al., Dual-view plane illumination microscopy for rapid and spatially isotropic imaging, Nat. Protoc. 9 (2014) 2555–2573, https://doi.org/10.1038/nprot.2014.172.
[31] A.R. Bhagwat, V. Le Sage, S.S. Lakdawala, Live imaging of influenza viral ribonucleoproteins using light-sheet microscopy, Methods Mol. Biol. 1836 (2018) 303–327, https://doi.org/10.1007/978-1-4939-8678-1_15.
[32] R.W. Compans, J. Content, P.H. Duesberg, Structure of the ribonucleoprotein of influenza virus, J. Virol. 10 (1972) 795–800.
[33] A. Harris, et al., Influenza virus pleiomorphy characterized by cryoelectron tomography, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 19123–19127, https://doi.org/10.1073/pnas.0607614103.
[34] T. Noda, et al., Architecture of ribonucleoprotein complexes in influenza A virus particles, Nature 439 (2006) 490–492, https://doi.org/10.1038/nature04378.
[35] R. Arranz, et al., The structure of native influenza virion ribonucleoproteins, Science 338 (2012) 1634–1637, https://doi.org/10.1126/science.1228172.
[36] A. Moeller, R.N. Kirchdoerfer, C.S. Potter, B. Carragher, I.A. Wilson, Organization of the influenza virus replication machinery, Science 338 (2012) 1631–1634, https://doi.org/10.1126/science.1227270.
[37] T. Noda, Y. Kawaoka, Packaging of influenza virus genome: robustness of selection, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 8797–8798, https://doi.org/10.1073/pnas.1206736109.
[38] J. Ule, K. Jensen, A. Mele, R.B. Darnell, CLIP: a method for identifying protein–RNA interaction sites in living cells, Methods 37 (2005) 376–386, https://doi.org/10.1016/j.ymeth.2005.07.018.
[39] S. Mili, J.A. Steitz, Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses, RNA 10 (2004) 1692–1694, https://doi.org/10.1261/rna.7151404.
[40] F.C.Y. Lee, J. Ule, Advances in CLIP technologies for studies of protein–RNA interactions, Mol. Cell 69 (2018) 354–369, https://doi.org/10.1016/j.molcel.2018.01.005.
[41] R.B. Darnell, HITS-CLIP: panoramic views of protein–RNA regulation in living cells, Wiley Interdiscip. Rev. RNA 1 (2010) 266–286, https://doi.org/10.1002/wrna.31.
[42] M. Hafner, et al., Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP, Cell 141 (2010) 129–141, https://doi.org/10.1016/j.cell.2010.03.009.
[43] N. Lee, et al., Genome-wide analysis of influenza viral RNA and nucleoprotein association, Nucleic Acids Res. 45 (2017) 8968–8977, https://doi.org/10.1093/nar/gkx584.
[44] V. Le Sage, et al., Non-uniform and non-random binding of nucleoprotein to influenza A and B viral RNA, Viruses 10 (2018), https://doi.org/10.3390/v10100522.
[45] G.D. Williams, et al., Nucleotide resolution mapping of influenza A virus nucleoprotein–RNA interactions reveals RNA features required for replication, Nat. Commun. 9 (2018) 465, https://doi.org/10.1038/s41467-018-02886-w.
[46] F. Baudin, C. Bach, S. Cusack, R.W. Ruigrok, Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent, EMBO J. 13 (1994) 3158–3165.
[47] K. Yamanaka, A. Ishihama, K. Nagata, Reconstitution of influenza virus RNA–nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores, J. Biol. Chem. 265 (1990) 11151–11155.
[48] M.W. Hentze, A. Castello, T. Schwarzl, T. Preiss, A brave new world of RNA-binding proteins, Nat. Rev. Mol. Cell Biol. 19 (2018) 327–341, https://doi.org/10.1038/nrm.2017.130.
[49] G. Singh, G. Pratt, G.W. Yeo, M.J. Moore, The clothes make the mRNA: past and present trends in mRNP fashion, Annu. Rev. Biochem. 84 (2015) 325–354, https://doi.org/10.1146/annurev-biochem-080111-092106.
[50] A.P. Gultyaev, et al., RNA structural constraints in the evolution of the influenza A virus genome NP segment, RNA Biol. 11 (2014) 942–952, https://doi.org/10.4161/rna.29730.
[51] Y. Kobayashi, et al., Computational and molecular analysis of conserved influenza A virus RNA secondary structures involved in infectious virion production, RNA Biol. 13 (2016) 883–894, https://doi.org/10.1080/15476286.2016.1208331.
[52] R.J. Thapa, et al., DAI senses influenza A virus genomic RNA and activates RIPK3-dependent cell death, Cell Host Microbe 20 (2016) 674–681, https://doi.org/10.1016/j.chom.2016.09.014.
[53] E.C. Hutchinson, et al., Conserved and host-specific features of influenza virion architecture, Nat. Commun. 5 (2014) 4816, https://doi.org/10.1038/ncomms5816.
[54] B. Gilbertson, et al., Influenza NA and PB1 gene segments interact during the formation of viral progeny: localization of the binding region within the PB1 gene, Viruses 8 (2016), https://doi.org/10.3390/v8080238.
[55] C. Gavazzi, et al., An in vitro network of intermolecular interactions between viral RNA segments of an avian H5N2 influenza A virus: comparison with a human H3N2 virus, Nucleic Acids Res. 41 (2013) 1241–1254, https://doi.org/10.1093/nar/gks1181.
[56] H. Bolte, M.E. Rosu, E. Hagelauer, A. Garcia-Sastre, M. Schwemmle, Packaging of the influenza virus genome is governed by a plastic network of RNA- and nucleoprotein-mediated interactions, J. Virol. 93 (2019), https://doi.org/10.1128/JVI.01861-18.
[57] A. Fatica, I. Bozzoni, Long non-coding RNAs: new players in cell differentiation and development, Nat. Rev. Genet. 15 (2014) 7–21, https://doi.org/10.1038/nrg3606.
[58] J. Steitz, RNA–RNA base-pairing: theme and variations, RNA 21 (2015) 476–477, https://doi.org/10.1261/rna.050427.115.
[59] B. Van Treeck, R. Parker, Emerging roles for intermolecular RNA–RNA interactions in RNP assemblies, Cell 174 (2018) 791–802, https://doi.org/10.1016/j.cell.2018.07.023.
[60] G.D. Cimino, H.B. Gamper, S.T. Isaacs, J.E. Hearst, Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry, Annu. Rev. Biochem. 54 (1985) 1151–1193, https://doi.org/10.1146/annurev.bi.54.070185.005443.
[61] J. Engreitz, E.S. Lander, M. Guttman, RNA antisense purification (RAP) for mapping RNA interactions with chromatin, Methods Mol. Biol. 1262 (2015) 183–197, https://doi.org/10.1007/978-1-4939-2253-6_11.
[62] C. Gorbea, T. Mosbruger, D. Cazalla, A viral Sm-class RNA base-pairs with mRNAs and recruits microRNAs to inhibit

apoptosis, Nature 550 (2017) 275–279, https://doi.org/10.1038/nature24034.
[63] A.V. Nanni, N. Lee, Identification of host RNAs that interact with EBV noncoding RNA EBER2, RNA Biol. (2018), https://doi.org/10.1080/15476286.2018.1518854.
[64] Z. Lu, et al., RNA duplex map in living cells reveals higher-order transcriptome structure, Cell 165 (2016) 1267–1279, https://doi.org/10.1016/j.cell.2016.04.028.
[65] E. Sharma, T. Sterne-Weiler, D. O'Hanlon, B.J. Blencowe, Global mapping of human RNA–RNA interactions, Mol. Cell 62 (2016) 618–626, https://doi.org/10.1016/j.molcel.2016.04.030.
[66] J.G. Aw, et al., In vivo mapping of eukaryotic RNA Interactomes reveals principles of higher-order organization and regulation, Mol. Cell 62 (2016) 603–617, https://doi.org/10.1016/j.molcel.2016.04.028.
[67] T.C. Nguyen, et al., Mapping RNA–RNA interactome and RNA structure in vivo by MARIO, Nat. Commun. 7 (2016) 12023, https://doi.org/10.1038/ncomms12023.
[68] N. Lee, W.N. Moss, T.A. Yario, J.A. Steitz, EBV noncoding RNA binds nascent RNA to drive host PAX5 to viral DNA, Cell 160 (2015) 607–618, https://doi.org/10.1016/j.cell.2015.01.015.
[69] J. Dekker, K. Rippe, M. Dekker, N. Kleckner, Capturing chromosome conformation, Science 295 (2002) 1306–1311, https://doi.org/10.1126/science.1067799.
[70] C.B. Brooke, Population diversity and collective interactions during influenza virus infection, J Virol 91 (2017), https://doi.org/10.1128/JVI.01164-17.
[71] N.J. Loman, et al., Performance comparison of benchtop high-throughput sequencing platforms, Nat. Biotechnol. 30 (2012) 434–439, https://doi.org/10.1038/nbt.2198.
[72] M. Schirmer, et al., Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform, Nucleic Acids Res. 43 (2015), e37, https://doi.org/10.1093/nar/gku1341.
[73] B. Zhou, et al., Multiplex reverse transcription-PCR for simultaneous surveillance of influenza A and B viruses, J. Clin. Microbiol. 55 (2017) 3492–3501, https://doi.org/10.1128/JCM.00957-17.
[74] B. Zhou, et al., Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and swine origin human influenza A viruses, J. Virol. 83 (2009) 10309–10313, https://doi.org/10.1128/JVI.01109-09.
[75] J.T. McCrone, A.S. Lauring, Measurements of Intrahost viral diversity are extremely sensitive to systematic errors in variant calling, J. Virol. 90 (2016) 6884–6895, https://doi.org/10.1128/JVI.00667-16.
[76] T.H. Zhang, N.C. Wu, R. Sun, A benchmark study on error-correction by read-pairing and tag-clustering in amplicon-based deep sequencing, BMC Genomics 17 (2016) 108, https://doi.org/10.1186/s12864-016-2388-9.
[77] S. Zhou, C. Jones, P. Mieczkowski, R. Swanstrom, Primer ID validates template sampling depth and greatly reduces the error rate of next-generation sequencing of HIV-1 genomic RNA populations, J. Virol. 89 (2015) 8540–8555, https://doi.org/10.1128/JVI.00522-15.
[78] M.D. Pauly, M.C. Procario, A.S. Lauring, A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses, Elife 6 (2017), https://doi.org/10.7554/eLife.26437.
[79] I. Kosik, et al., Influenza A virus hemagglutinin glycosylation compensates for antibody escape fitness costs, PLoS Pathog. 14 (2018), e1006796, https://doi.org/10.1371/journal.ppat.1006796.
[80] F.J. Sedlazeck, H. Lee, C.A. Darby, M.C. Schatz, Piercing the dark matter: bioinformatics of long-range sequencing and mapping, Nat. Rev. Genet. 19 (2018) 329–346, https://doi.org/10.1038/s41576-018-0003-4.
[81] M.B. Rogers, et al., Intrahost dynamics of antiviral resistance in influenza A virus reflect complex patterns of segment linkage, reassortment, and natural selection, MBio 6 (2015), https://doi.org/10.1128/mBio.02464-14.
[82] J. Eid, et al., Real-time DNA sequencing from single polymerase molecules, Science 323 (2009) 133–138, https://doi.org/10.1126/science.1162986.
[83] C.K. Ho, et al., Characterization of hepatitis C virus (HCV) envelope diversification from acute to chronic infection within a sexually transmitted HCV cluster by using single-molecule, real-time sequencing, J. Virol. 91 (2017), https://doi.org/10.1128/JVI.02262-16.
[84] A.B. Russell, J.R. Kowalsky, J.D. Bloom, Single-cell virus sequencing of influenza infections that trigger innate immunity, bioRxiv (2018), https://doi.org/10.1101/437277.
[85] J. Wang, N.E. Moore, Y.M. Deng, D.A. Eccles, R.J. Hall, MinION nanopore sequencing of an influenza genome, Front. Microbiol. 6 (2015) 766, https://doi.org/10.3389/fmicb.2015.00766.
[86] D.R. Garalde, et al., Highly parallel direct RNA sequencing on an array of nanopores, Nat. Methods 15 (2018) 201–206, https://doi.org/10.1038/nmeth.4577.
[87] M.W. Keller, et al., Author correction: direct RNA sequencing of the coding complete influenza A virus genome, Sci. Rep. 8 (2018) 15746, https://doi.org/10.1038/s41598-018-34067-6.
[88] M.B. Doud, J.D. Bloom, Accurate measurement of the effects of all amino-acid mutations on influenza hemagglutinin, Viruses 8 (2016), https://doi.org/10.3390/v8060155.
[89] N.C. Wu, et al., High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution, Sci. Rep. 4 (2014) 4942, https://doi.org/10.1038/srep04942.
[90] S.K. Hilton, J.D. Bloom, Modeling site-specific amino-acid preferences deepens phylogenetic estimates of viral sequence divergence, Virus Evol. 4 (2018), vey033, https://doi.org/10.1093/ve/vey033.
[91] M.B. Doud, S.E. Hensley, J.D. Bloom, Complete mapping of viral escape from neutralizing antibodies, PLoS Pathog. 13 (2017), e1006271, https://doi.org/10.1371/journal.ppat.1006271.
[92] J.M. Lee, et al., Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) E8276–E8285, https://doi.org/10.1073/pnas.1806133115.
[93] N.C. Wu, et al., Functional constraint profiling of a viral protein reveals discordance of evolutionary conservation and functionality, PLoS Genet. 11 (2015), e1005310, https://doi.org/10.1371/journal.pgen.1005310.
[94] A.B. Russell, C. Trapnell, J.D. Bloom, Extreme heterogeneity of influenza virus infection in single cells, Elife 7 (2018), https://doi.org/10.7554/eLife.32303.
[95] V.Y. Kiselev, T.S. Andrews, M. Hemberg, Challenges in unsupervised clustering of single-cell RNA-seq data, Nat. Rev. Genet. (2019), https://doi.org/10.1038/s41576-018-0088-9.
[96] A. Sobel Leonard, D.B. Weissman, B. Greenbaum, E. Ghedin, K. Koelle, Transmission bottleneck size estimation from pathogen deep-sequencing data, with an

application to human influenza A virus, J. Virol. 91 (2017), https://doi.org/10.1128/JVI.00171-17.

[97] J.T. McCrone, et al., Stochastic processes constrain the within and between host evolution of influenza virus, Elife 7 (2018), https://doi.org/10.7554/eLife.35962.

[98] A. Varble, et al., Influenza A virus transmission bottlenecks are defined by infection route and recipient host, Cell Host Microbe 16 (2014) 691−700, https://doi.org/10.1016/j.chom.2014.09.020.

[99] N. Marshall, L. Priyamvada, Z. Ende, J. Steel, A.C. Lowen, Influenza virus reassortment occurs with high frequency in the absence of segment mismatch, PLoS Pathog. 9 (2013), e1003421, https://doi.org/10.1371/journal.ppat.1003421.

[100] A. Sobel Leonard, et al., The effective rate of influenza reassortment is limited during human infection, PLoS Pathog. 13 (2017), e1006203, https://doi.org/10.1371/journal.ppat.1006203.