Journal of Virological Methods 266 (2019) 95–102

Contents lists available at ScienceDirect

Journal of Virological Methods

journal homepage: www.elsevier.com/locate/jviromet

Application of deep sequencing methods for inferring viral population diversity

Sheng-Wen Huang[a], Su-Jhen Hung[b], Jen-Ren Wang[b,c,d,e,*]

[a] National Mosquito-Borne Diseases Control Research Center, National Health Research Institutes, Tainan, Taiwan
[b] Department of Medical Laboratory Science and Biotechnology, National Cheng Kung University, Tainan, Taiwan
[c] Center of Infectious Disease and Signaling Research, National Cheng Kung University, Tainan, Taiwan
[d] Department of Pathology, National Cheng Kung University Hospital, Tainan, Taiwan
[e] National Institute of Infectious Diseases and Vaccinology, National Health Research Institutes, Tainan, Taiwan

ARTICLE INFO

Keywords:
Deep sequencing
Quasispecies
Viral evolution

ABSTRACT

The first deep sequencing method was announced in 2005. Due to an increasing number of sequencing data and a reduction in the costs of each sequencing dataset, this innovative technique was soon applied to genetic investigations of viral genome diversity in various viruses, particularly RNA viruses. These deep sequencing findings documented viral epidemiology and evolution and provided high-resolution data on the genetic changes in viral populations. Here, we review deep sequencing platforms that have been applied in viral quasispecies studies. Further, we discuss recent deep sequencing studies on viral inter- and intrahost evolution, drug resistance, and humoral immune selection, especially in emerging and re-emerging viruses. Deep sequencing methods are becoming the standard for providing comprehensive results of viral population diversity, and their applications are discussed.

1. Introduction

Because error-prone RNA viruses exhibit rapid generation times and high mutation rates, viruses can generate large and diverse populations (Drake, 1993). Genetic drift, genetic shift, and gene recombination contribute to shaping the evolution of virus populations (reviewed in (Duffy et al., 2008)). Among virus populations, a collection of diverse but closely related genetic strains of viruses may coexist within hosts, and they are referred to as viral quasispecies (Domingo et al., 2005; Lauring and Andino, 2010). The quasispecies theory was first described in the investigation of molecular evolution by Eigen and Schuster in 1977 (Eigen and Schuster, 1977). This model has been applied in virology to explain the evolutionary dynamics in population diversity of error-prone viruses, especially RNA viruses (Domingo and Holland, 1997; Nowak, 1992).

Due to utilizing low-fidelity RNA polymerase (RNA-dependent RNA polymerase) without proofreading activity, RNA viruses can rapidly generate many mutations in viral genomes in each replication cycle (Domingo and Holland, 1997), and the genetic variability can be easily observed because RNA viruses exhibit rapid generation times. The average rate of nucleotide substitution using serially sampled gene sequence data is estimated to be approximately 10^{-2} to 10^{-5} nucleotide substitutions per site/year (sub/site/year) (Hanada et al., 2004; Jenkins et al., 2002), with most of them exhibiting a rate of 10^{-3} sub/site/year (reviewed in (Duffy et al., 2008)). Therefore, mutations are introduced into genomes, and diverse viral traits containing mutations, insertions, or deletions are generated in each replication cycle. Genome recombination is an additional source of diverse viral traits in viral populations. As two viruses simultaneously infect the same cell, a recombination or reassortment event may occur and generate mosaic viral genomes with new genetic traits that differ from the progenitor genomes. These evolutionary events aid viruses in the development of a broad spectrum of mutants and lead to population diversity that increases genetic plasticity in the face of various selection pressures. Selection pressures act on viral quasispecies and shape the variation within the population by reducing variability and increasing the frequency distribution of those variants with high viral fitness in response to various stresses. Through the shaping of mutant spectra by selection pressures, certain variants become dominant but are still surrounded by various minor variants with low frequency (Bonhoeffer and Nowak, 1997). This cloud of minor variants will be continually retained in viral quasispecies, and it provides extended population diversity. As virus

* Corresponding author at: Department of Medical Laboratory Science and Biotechnology, College of Medicine, National Cheng Kung University, One University Road, Tainan, Taiwan.
E-mail address: jrwang@mail.ncku.edu.tw (J.-R. Wang).

https://doi.org/10.1016/j.jviromet.2019.01.013
Received 7 September 2018; Received in revised form 11 January 2019; Accepted 24 January 2019
Available online 25 January 2019
0166-0934/

populations experience other selection pressures, this extended population diversity can serve as a broad reservoir of variants that allows the virus population to rapidly adapt to environmental changes.

Since the divergence of variants in populations confers more advantages for overcoming various selection pressures, including immune response escape, drug resistance, antigenic drift or shift, and pathogenesis within hosts, comparisons and analyses of mutant spectra from viral quasispecies of different origins may provide valuable evolutionary information on virus adaptation when viruses face these selection pressures (Borucki et al., 2013; Gaschen et al., 2002; Johnson et al., 2008; Kuroda et al., 2010; Nowak et al., 1991; Rozera et al., 2014; Tsibris et al., 2009; Vignuzzi et al., 2006; Woo and Reifman, 2012). Sanger sequencing combined with classical molecular cloning or limiting dilutions are the initial methods of choice to determine viral quasispecies in various virus samples (Palmer et al., 2005). To date, Sanger sequencing is still the standard method of defining genetic mutations in viral genomes; however, time- and labor-intensive sequencing processes with limited sensitivity reduce the applicability to large-scale studies of viral quasispecies. With the introduction of the deep sequencing method in 2005, this innovative technique dramatically increased the number of sequencing data and reduced the costs of each sequencing dataset. Thousands of virus sequences can be surveyed by utilizing this method, which improves the analysis sensitivity and precision of viral quasispecies diversity. The method has been applied to determine and compare mutant spectra of viral populations under various conditions, such as immune selection, including antibody or T-cell epitopes (Bimber et al., 2010; Hughes et al., 2010), antiviral drug resistance (Cannon et al., 2008), and different tissue tropisms of viral variants within or between hosts (Wang et al., 2010) (Fig. 1). Studies that have defined viral population diversity in hosts are informative for viral pathogenesis, virulence, persistence, drug therapy, disease progression, and vaccine development. Here, we review the recent findings of viral quasispecies and their impacts on viral evolution, disease prevention or therapy. In addition, we discuss the application of various platforms and data analysis methods for deep sequencing in viral quasispecies investigations.

2. Deep sequencing platform comparisons for viral quasispecies analysis

Each deep sequencing platform varies with respect to the throughput, read numbers, error rates, and cost per base. Among the compared platforms (Table 1), the suite of Roche 454 and Illumina instruments are widely applied in viral quasispecies studies. Roche 454 pyrosequencing was the first platform developed for deep sequencing, and researchers utilized this system to comprehensively analyze mutant spectra of various viruses including human immunodeficiency virus (HIV), hepatitis B virus (HBV), hepatitis C virus (HCV), enterovirus, and influenza viruses (Bartolini et al., 2015; Bednar et al., 2016; Bourret et al., 2015, 2013; Bunnik et al., 2011; Chabria et al., 2014; Cortes et al., 2014; Dialdestoro et al., 2016; Fischer et al., 2010; Fordyce et al., 2013; Gall et al., 2013; Hansen et al., 2016; Hedegaard et al., 2017; Jabara et al., 2011; Lequime et al., 2016; Ode et al., 2015; Parameswaran et al., 2012; Poon et al., 2016; Prabhakara et al., 2013; Redd et al., 2012; Rodriguez-Roche et al., 2016; Romano et al., 2013; Sede et al., 2014; Sobel Leonard et al., 2016; Stapleford et al., 2016; Tellez-Sosa et al., 2013; Tsai and Chen, 2011; Wang et al., 2010; Wilker et al., 2013; Zanini et al., 2016). Through single-nucleotide addition, the Roche 454 platform uses enzymatic reactions to detect deoxyribonucleotide triphosphate (dNTP) incorporation into a DNA template, thereby resulting in a single bioluminescence signal. Moreover, this platform offers long read lengths at up to an average of 700 bp, which provides greater advantages for repetitive or complex sequences, such as those in specific viral genome regions, compared with short-read sequencers, such as Illumina platforms. Long read length is also advantageous for haplotype reconstruction of viral quasispecies. The Roche 454 platform has drawbacks for insertion and deletion (indel) errors, although the overall error rate is equal to that of other platforms except in homopolymer regions. For measuring homopolymers longer than 6–8 bp, the Roche 454 platform exhibits low single-base accuracy for sequencing results. Unfortunately, Roche discontinued this platform in 2016 due to the low yield and high cost per sequencing base relative to other platforms. The Illumina suite of instruments is increasingly being utilized within various fields of investigation due to high confidence in data generated by the platform. Recently, Illumina has become the dominant platform for short-read sequencing because of its mature technology, wide application range, and high compatibility between platforms. Illumina provides various platforms ranging from low-throughput MiniSeq to ultrahigh-throughput HiSeq X. Compared with sequencing by single-nucleotide addition in the Roche 454 platform, the Illumina platform relies on the cyclic reversible termination technique, which is less susceptible to the homopolymer errors that

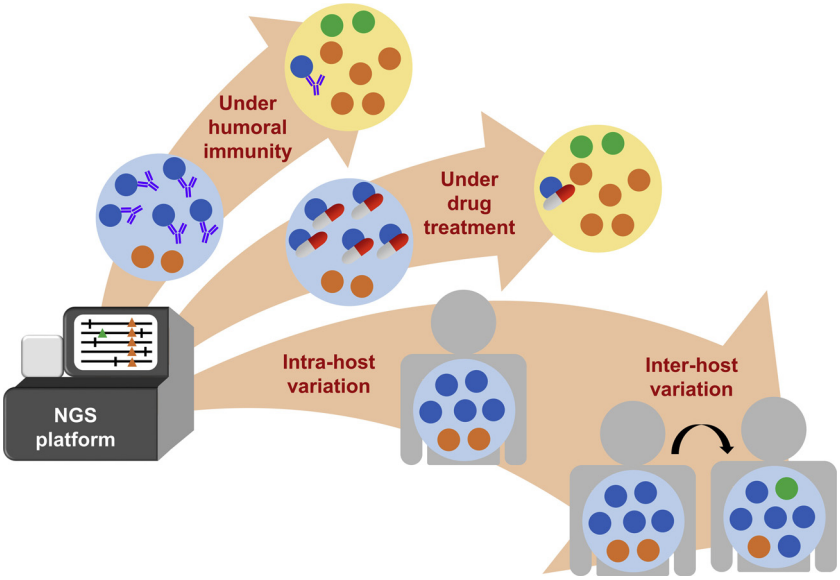

Fig. 1. Applications of deep sequencing in the field of viral population diversity. Deep sequencing methods have been applied in studies of viral evolution under various selection pressures, including humoral immunity, drug treatment, and intra- or interhost adaptations.

Table 1
Platforms commonly applied in viral quasispecies investigations.

	Roche	Illumina	Pacific BioSciences	Thermo Fisher Scientific
Platform	**454 GS Junior** **454 GS FLX** **Titanium**	**iSeq** **MiniSeq** **MiSeq** **NextSeq** **HiSeq** **HiSeq X** **NovaSeq**	**RSII** **Sequel**	**Ion PGM** **Ion Proton** **Ion S5**
Read length	**˜700 bp**	**25-300 bp (Pair end)** **36-150 bp (Single end)**	**400 bp - 20 kbp**	**200-400 bp**
Drawback	**Homopolymers > 6-8 bases with high error rates for insertions and deletions**	**Under-representation in AT-rich and GC-rich regions increasing sequencing errors**	**Approximately 13% error rate for insertions and deletions (can be reduced by circular consensus sequences to 1%)**	**Homopolymers with high error rates for insertions and deletions**
Average error rate	**˜1%, indel**	**˜0.1%, substitution**	**˜13%, indel**	**˜1%, indel**
Cost per Gb (USD $)	**9500 to 40000**	**7 to 1000**	**1000**	**25 to 3500**
Recently available	**No**	**Yes**	**Yes**	**Yes**

occurred in the Roche 454 platform. However, Illumina results underrepresent AT-rich and GC-rich regions, which tends to increase sequencing errors. Although Illumina has an average accuracy of > 99.5% and dominates in recent viral quasispecies research, concerns have been raised that a systematic error may occur in these studies when utilizing a single deep sequencing platform. Therefore, more approaches have been developed, and multiple sequencing methods with complementary strengths have been applied.

The Pacific Biosciences (PacBio) and Ion Torrent systems are recent alternative sequencing methods for viral mutant spectra studies; these two platforms generate longer length reads of approximately up to 20 kbp and 400 bp, respectively. Superior length reads compared to other platforms provide more accurate haplotype reconstruction results to resolve viral evolution in mutant spectra investigations. PacBio utilizes a flow cell containing thousands of picoliter wells, and the polymerase is fixed in the bottom of transparent wells (zero-mode waveguides, ZMW). As the DNA template binds, the polymerase can process through the DNA strand and incorporate dNTP at a consistent location. The PacBio system focuses on the dNTP incorporation process of a single molecule and captures real-time images of a single molecule template for each well. When a dNTP-bound fluorophore is cleaved by the polymerase during incorporation, the fluorophore signal will diffuse from the sensor area until the next dNTP is incorporated. The PacBio platform can generate an average read length of 10–15 kb for a long-insert library, which nearly covers the complete genome of certain RNA viruses. This platform provides advantages for viral haplotype analysis; however, the error rate of long single-pass reads is approximately 15% with mainly indel errors. Because these errors randomly appear for each read, sufficiently high coverage can overcome the high error rate. In addition, the use of circular consensus sequences (CCSs) with the PacBio platform can reduce the error rate to a degree. The PacBio platform utilizes unique circular molecules to repeatedly sequence the template multiple times. The consensus sequences of inserts are generated by multiple sequencing passes of CCSs that aid in reducing indel errors.

Ion Torrent systems, which are similar to the Roche 454 system, rely on single-nucleotide addition rather than using enzymatic methods. Ion Torrent detects the H^+ ions that are released as the dNTPs are incorporated and utilizes a pH change sensor to detect the change in pH value. Unfortunately, the reduced sensitivity of the pH sensor for quantifying amounts of nucleotides reduces the accuracy of detecting the length of homopolymers, which is similar to the Roche 454 system. Although the Roche 454 system has been discontinued, the Ion Torrent platform developed several types of chips and instruments to customize the performance and throughput for various investigative needs. These systems shorten run times to 2–7 hours with 50 Mb to 15 Gb throughput, representing some of the fastest detection systems among the deep sequencing platforms that can be applied in clinical applications, such as sequencing drug-resistant viruses.

Recently, several viral diversity studies combined the short and long read platforms to retrieve the full resolution of genomic variations (Bull et al., 2016; Moldovan et al., 2017a, b; Rogers et al., 2015; Zhu et al., 2017). The short reads from the platforms including illumina or Ion Torrent provided deep population of viral quasispecies for identification of polymorphisms, but it is difficult to display quasispecies only based on allele frequencies alone. In contrast, PacBio platform determines the quasispecies structures by directly producing long reads for individual viral genetic variants without requiring a reference sequence. Combining with the advantage of short and long read platforms, these recent studies characterized and revealed complete population diversity and population dynamics in viral evolution.

Although deep sequencing showed advanced advantages for revealing population diversity, some challenges of implementing the technique remained to be overcome. In contrast to conventional sanger sequencing method, deep sequencing assay posed high degrees of complexity due to large sizes of analyzed genomes and various bioinformatic pipelines for analysis. Therefore, to determine the assay performance parameters and quality control metrics, including analytical sensitivity, analytical specificity, accuracy, precision, limit of detection, and sequencing depth and allelic frequency cutoffs, are essential for validating sequencing results within and between assays or even different studies (Alame et al., 2016; Fisher et al., 2016; Kluk et al., 2016; Lowe et al., 2016; Roy et al., 2016). For example, instead of only use illumina platform to determine genetic variations, the Ion Torrent platform can be applied as the validating platforms (use similar targeting regions) that might facilitate to compare and to confirm the results of viral genetic diversity. In addition to assay performance and quality control, analysis pipelines or programs/softwares need to be validated and compared. Analysis programs/softwares are frequently upgraded and might result in diverse results due to the improvements that require to be validated among different deep sequencing studies.

3. Application of deep sequencing for inter- and intrahost evolution of RNA viruses

Since deep sequencing methods are capable of providing extreme sequence read depths, they can be widely applied to detect various genetic traits in a mixed population, such as viral quasispecies. As mentioned previously, RNA viruses accumulate abundant nucleotide substitutions to generate diverse traits as a reservoir of viruses with

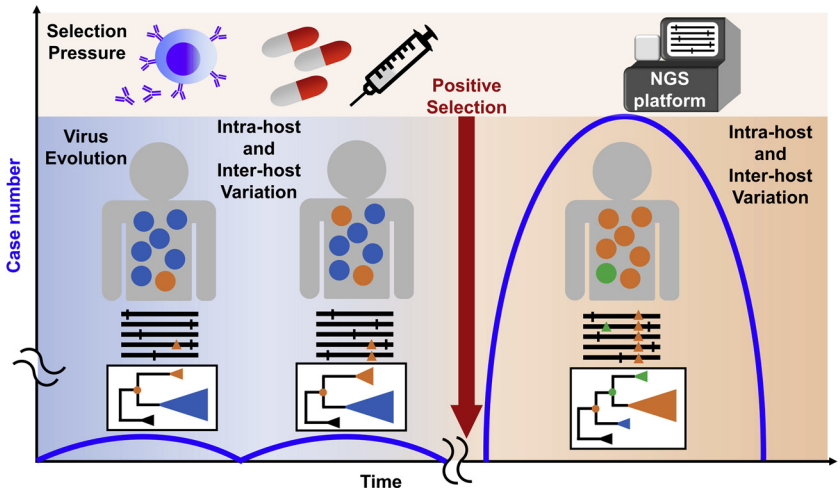

Fig. 2. Insights into viral genome diversity and viral evolution under various selection pressures through deep sequencing. As viruses are transmitted among hosts such as humans, purifying selection pressure is the dominant force of viral evolution. During this period, virus populations accumulate various synonymous mutations (short black lines) with few nonsynonymous mutations (small orange triangles) among genomes (long black lines). Under positive selection pressures, nonsynonymous mutations that have advantages for certain traits that allow viruses to resist selection pressures, such as viral spread and growth traits, will become the dominant trait in a viral population. The emerging virus trait may increase the viral activity and cause a wave of outbreaks or epidemics.

various haplotypes. In addition to nucleotide substitutions, genetic rearrangement, including reassortment and recombination, are additional resources of viral genetic traits. Genome-segmented viruses generate new virus variants by reassortment to exchange genomes between different viruses infecting the same cell. Compared with genome-segmented viruses, single-stranded RNA viruses recombine heterogeneous genomes as two virus variants simultaneously infect a cell. Both of these processes rapidly accumulate several substitutions located in viral variant genomes and obviously increase population diversity, which is advantageous for viral adaptation within hosts. Recently, deep sequencing methods have been applied to examine the occurrence, mechanism, and clinical significance of substitution accumulation in viral evolution, especially for emerging and re-emerging RNA viruses (Eshaghi et al., 2014; Huang et al., 2017; Jonges et al., 2014; Moncla et al., 2016; Rogers et al., 2015).

Emerging and re-emerging RNA viruses are selected between susceptible hosts with various selection pressures, which causes virus evolution in divergent directions. Various evolutionary directions lead to increases in the genetic diversity in the virus quasispecies, which leave genetic signature sequences as “footprints” among the viral genomes in the process of viral evolution; however, conventional Sanger sequencing only reveals the consensus sequences in viral quasispecies and loses an abundance of virus genetic information. To determine viral dynamic evolution among various hosts, even those of different species, recent viral evolution studies, especially on emerging and re-emerging RNA viruses, have utilized deep sequencing instead of Sanger sequencing to gain insights into viral evolution with transmission among various hosts and evaluate the selection pressures faced by viruses (Eshaghi et al., 2014; Huang et al., 2017; Jonges et al., 2014; Moncla et al., 2016; Rogers et al., 2015).

Influenza virus is a globally emerging and re-emerging RNA virus, and deep sequencing has been applied to this virus to study its genetic transmission among hosts. Recently, Wilker et al. identified a selective bottleneck of avian influenza viruses during virus transmission among ferrets (Wilker et al., 2013). Genetic diversity in the hemagglutinin gene was analyzed in the experimental ferret model during transmission of reassorted H5N1 influenza viruses. In the initial infection, high virus diversity appeared during viral replication in inoculated ferrets; however, the genetic diversity was obviously reduced in the process of transmission through airway droplets between ferrets, with only 1 or 2 haplotypes becoming established. Some minor haplotypes in the source animals became dominant in the infected ferrets through air droplet transmission, which demonstrates that a selection bottleneck occurred in mammalian transmission. Another evolutionary bottleneck has been observed for the 1918-like avian influenza virus. In the ferret transmission model, hemagglutinin diversity initially increased within hosts. Through airborne virus transmission, two polymerase mutations, PA-V253 M and PB2-A684D, were fixed in the population, although they did not increase virus replication. The additional transmissions resulted in HA-S113 N, HA-D265 V, or HA-I187 T mutations near receptor-binding regions in the variants, which indicated multiple genetic evolutionary pathways in virus transmission. These results suggest that a relaxed bottleneck occurred in the initial infection in hosts, and it became a strong selection pressure between hosts. In addition to investigating viral transmission among animals, transmission among humans has been investigated. Another study utilized deep sequencing results of eight segmented influenza virus genomes to quantify virus evolutionary changes within humans in the first wave of the 2009 H1N1 outbreak (Kuroda et al., 2010). Viruses from household donor-recipient pairs among human hosts in the community were analyzed to examine virus variants that confer transmission to new individuals. The results revealed that the same variant could be identified in multiple members of the community. Although variants showed different frequencies, the frequency patterns within households were more similar than those between households (Poon et al., 2016). Variant swarming was observed within hosts (intrahosts) as well as between hosts (interhosts). To precisely determine the intrahost evolution of the influenza virus in humans, Leonard et al. challenged 17 healthy human subjects with passaged viruses and examined the viral populations in these subjects using deep sequencing (Sobel Leonard et al., 2016). A selection bottleneck was detected with intranasal inoculation with virus decreasing nonsynonymous mutations in viral quasispecies. In contrast, abundant synonymous mutations of hemagglutinin and nucleoprotein occurred in the process of infection within hosts, suggesting that purifying selection was the dominant selection pressure in intrahost evolution. These results indicate that viral quasispecies provide influenza virus with a variant reservoir to dynamically evolve when faced with various selection pressures, e.g., intrahost and interhost conditions (Fig. 2). Compared with purifying selection in influenza viruses, positive selection was the dominant pressure in the evolution of respiratory syncytial virus (RSV), another respiratory tract virus. Do et al. performed deep sequencing on RSV-positive samples collected from the upper airways of hospitalized children (Do et al., 2015) and found that the G gene was the most variable and positively selected gene throughout the viral genome. The selection sites of the G gene were either in close proximity or overlapped with potential glycosylation motifs, suggesting that amino acid glycosylation might play important roles in viral genetic diversity.

Inter- and intrahost evolution was observed not only among mammals but also between vertebrates and invertebrates. Recent genetic diversity investigations of dengue virus also focused on intrahost or interhost populations in various infected patients and genetic evolution

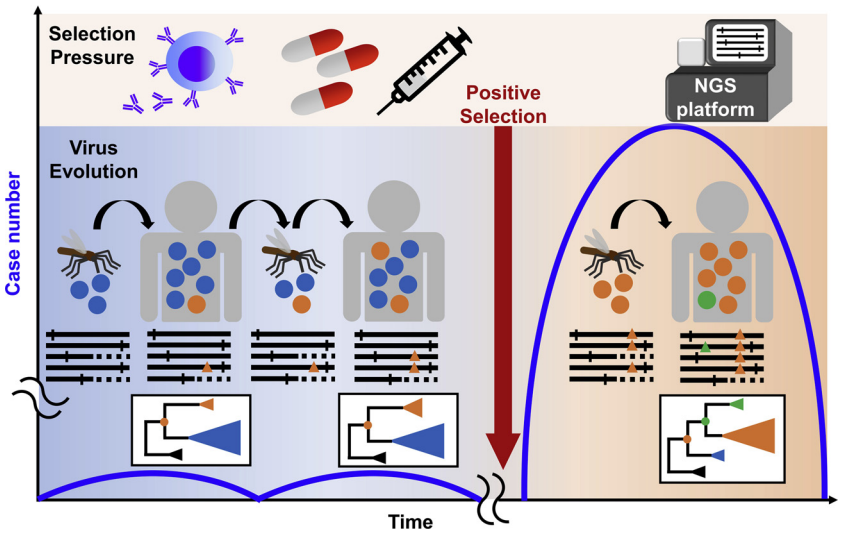

Fig. 3. Virus evolution in transmissions between hosts and vectors by using deep sequencing. Similar to the intra- and interhost evolution among hosts in the same species, purifying selection plays a major role in vector-borne virus evolution. As a virus transmits between hosts and vectors, it continuously accumulates synonymous (small orange triangles)/nonsynonymous mutations (short black lines) and deletions (black dashed line). As positive selection pressure occurs, nonsynonymous mutations will become the dominant strains in the population, including vectors and hosts, which may result in an emerging outbreak.

between humans and mosquitos. As different genotypes or lineages of dengue viruses become predominant in different countries, quasispecies of these genotypes or lineages need to be investigated among various outbreaks to understand dengue virus evolution. Parameswaran et al. (2012) analyzed dengue virus serotype 2 genomes from the sera of Nicaraguan patients with secondary dengue virus infection by deep sequencing. Approximately 75% of genomes for each sample were retrieved among the sequencing reads; accordingly, significantly high genetic diversity within hosts was detected in the structural E gene and the nonstructural NS5 gene; in particular, domain II of the E protein exhibited higher diversity than the other domains. A comparison of genetic diversity between diverse clades showed a significant difference because the NI-I clade exhibited higher diversity than the NI-2B clade in E, NS4A, and NS4B genes, suggesting that different clades of dengue virus might shape various viral evolution pathways. In examining viral evolution across dengue virus transmission events, most mutations accumulated in dengue virus were synonymous instead of nonsynonymous, which is similar to influenza virus evolution among humans and strongly suggests that purifying selection pressure shapes the evolutionary pathway of dengue virus. Similar findings were reported in other dengue virus outbreaks. Rodriguez-Roche et al. (2016) investigated intra- and interhost diversity of dengue virus 3 genomes from Havana in the 2001–2002 epidemic by genome-wide analysis. According to the consensus sequences of the virus genomes, purifying selection predominantly drove dengue virus 3 evolution, while abundant interhost synonymous mutations and few nonsynonymous mutations were identified in the epidemic. Intrahost populations also accumulated synonymous mutations in nonstructural regions during the progression of the epidemic. The minor synonymous intrahost variants continuously increased with time during the Havana epidemics; however, only a minor variant containing the nonsynonymous NS2B mutation in the early epidemic became the major variant at the end of the epidemic among patients. In another 2010 Brazil outbreak, strong purifying selection in dengue virus type 2 evolution was reported within hosts as well (Romano et al., 2013). Romano et al. (2013) showed that two diverse clades of dengue virus type 2 circulated among patients in this outbreak. Although the genomes of these diverse strains contain several nonsynonymous mutations due to the virus originating from various geographic regions in Brazil, more synonymous mutations were identified throughout these genomes. A relatively low intrahost diversity appeared in Brazil dengue virus type 2, and it indicated that most mutations were synonymous and accumulated in nonstructural protein regions. Since dengue virus transmission cycles between human hosts and mosquito vectors, the virus needs to retain transmission ability across two species. Intrahost genetic diversity was suggested as a survival advantage for rapid viral adaptation between these two hosts. To determine the role of genetic diversity in the viral adaptation of dengue virus, Sim et al. directly fed *Aedes aegypti* on dengue virus type 2-infected patients and examined single-nucleotide variations (SNVs) that appeared in the mosquito vectors (Sim et al., 2015). Although the SNV frequency did not change from human to mosquito abdomen and from mosquito abdomen to its salivary gland, less than 10% of the SNVs were preserved in the population through the process of human-to-mosquito and intra-mosquito transmissions. Instead of observing dominant purifying selection pressure, these results indicated that dengue virus faces positive selection pressure targeting the E, prM, and NS1 genes in humans but not in mosquitos, which was suggested to be driven by the host antibody response in humans that does not exist in mosquitos. In addition to the difference in antibody response in humans and mosquitos, the genetic divergence of mosquitos was observed to affect the composition of the variant population. Lequime et al. (2016) recently monitored the intrahost evolution of diverse mosquito lineages by deep sequencing. The evolution of SNVs is primarily determined by strong purifying selection and by genetically diverse viral populations in the midgut of mosquitos of different genetic backgrounds. Overall, dengue virus continues its intrahost and interhost evolution in humans or in cycling transmission between humans and mosquitos. In viral evolution, purifying selection is suggested as the predominant driving force that mainly accumulates synonymous mutations (Fig. 3). This finding might be due to dengue virus having to maintain genetic consistency for cross-infection among invertebrate and vertebrate hosts, although synonymous mutations are also suggested to have viral fitness effects in the hosts (Cuevas et al., 2012). Arboviruses can increase the divergence of variants through not only SNV but also genome duplication. Stapleford et al. performed a whole-genome deep sequencing analysis of Chikungunya virus in a Caribbean outbreak and in the Americas (Stapleford et al., 2016). Combined with identifying several minor variants with SNVs in the population, a novel 3′-UTR element duplication was identified. The 3′-UTR element duplication can aid virus replication in insect cells but not in mammalian cells, indicating that viruses utilize genome rearrangement to adapt to different host/vector species.

Through a deep sequencing analysis, we found that another RNA virus, enterovirus A71, emerged from a selection bottleneck in the dissemination route of humans (Huang et al., 2017). Analysis of the virus quasispecies in various tissues of an autopsy patient infected by enterovirus A71 revealed that the abundance of virus containing a glycine substitution at the VP1-31 residue dynamically increased as the dominant virus from peripheral respiratory tissues to the central nervous system. In addition, the clinical isolates from deceased patients

exhibited a higher proportion of VP1-31 G than those from mild hand-foot-and-mouth disease patients. An *in vitro* study indicated that this substitution had more growth advantages for replication in neuronal cells than in intestinal cells. Thus, we demonstrated that the increased prevalence of VP1-31 G might alter viral tropism and aid central nervous system invasion. Similar positive selection processes between patients were observed in other enteric viruses. In norovirus, deep sequencing performed by van Beek et al. (2017) also indicated that the viral mutation rates vary among patients. Amino acid sequences continuously changed at epitopes because the majority of VP1 amino acid mutations were located on the virion surface, which indicated positive selection among human hosts in enteric virus infection.

4. Application of deep sequencing for viral evolution under drug treatment

The deep sequencing approach was applied to identify viral drug resistance, and the NA-H274Y substitution was defined as one of the important drug-resistant mutations of influenza virus A (H3N2) for oseltamivir (Wu et al., 2013). With the application of deep sequencing and reverse genetics systems, some additional mutations appeared with NA-H274Y and coexisted in the population. These mutations allow viruses to restore viral fitness previously reduced by drug-resistant mutations. Recently, drug resistance was reported in immunosuppressed patients infected with the influenza A virus a few days after antiviral drug prophylaxis was initiated (Eshaghi et al., 2014). The patients received multiple antiviral drugs, including oseltamivir, zanamivir, and peramivir, during the course of infection. Although no drug-resistant mutation was identified in the initial infection, a combination of multiple mutations became dominant in the viral population and reduced the response to the three antiviral drugs used.

Recent studies regarding HBV indicated a number of variations in pre-core and basal core promoters associated with peginterferon-alfa and adefovir combination therapy (Jansen et al., 2017), and they determined that the variants located in pre-core and basal core promoters correlated with alanine transaminase, age, lower HBV-DNA level, and previous interferon (IFN) therapy in peginterferon-alfa and adefovir in HBeAg-negative patients. In a study of patients who had chronic HBV infection and were receiving long-term treatment with nucleotide analogs, a reverse transcriptase S78 T mutation in the polymerase protein resulted in a premature stop codon at S protein C69 that truncated the small HBV surface protein. This mutation was found to be associated with enhanced replication, and it was resistant to nucleotide analog treatment (Shirvani-Dastgerdi et al., 2017).

5. Application of deep sequencing for viral evolution selection under humoral immunity

In addition to antiviral drugs, humoral immunity against viruses confers another selection pressure in viral evolution. As viruses face the pressure, the variants that contain antibody escape mutations located in the immunodominant structural proteins will become the predominant strains. Influenza virus hemagglutinin frequently accumulated multiple mutations that account for antigenic changes to escape antibody recognition. A recent study utilized deep sequencing to examine the viral quasispecies of influenza viruses isolated from a cohort of patients who included vaccinated or nonvaccinated individuals. Although purifying selection is the dominant selection pressure, it did not change protein residues in vaccinated or nonvaccinated subjects, and a low frequency of two substitutions, HA-G158 K and HA-N159 K, were located in or close to antigenic sites and receptor binding sites. Thus, these two mutations are potentially associated with immune escape from specific antibodies against influenza viruses. The variants containing the two mutations might increase intra- or interhost antigenic diversity because they face strong immune selection pressures. Intrahost selection pressures also appear in acute human infections of dengue virus (Parameswaran et al., 2017). Parameswaran et al. (2017) analyzed the intrahost diversity of DENV-3 from dengue patients. The virus population was shaped by immune pressures and replication sites. Most of the variants are likely immune escape variants that might be evolutionarily constrained by replication defects. A recent study identified a G145R immune escape mutant that appeared in a case of HBV with coexisting HBsAg and anti-HBs. As analyzed by deep sequencing, the complexity in the S and RT regions dynamically changed with disease progression (Xue et al., 2017).

6. Conclusions

By applying deep sequencing technologies coupled with a bioinformatics analysis, viral quasispecies evolution can be revealed, thus allowing for the identification of potential genetic determinants for viral transmission or virulence. Compared with Sanger sequencing, which only defines the consensus sequence of a population, deep sequencing reads provide high-resolution data on the pooled virus variant composition in viral populations. Compared with the labor-intensive work of clone sequencing using the Sanger sequencing method, deep sequencing methods with high throughput provide more detailed genetic information, including information on mutations and genome recombination or reassortment hidden in viral quasispecies during viral evolution; moreover, various factors, such as selected platforms or analysis algorithms, affect the viral variant findings generated from deep sequencing. Different platforms generate various average read lengths, error rates, and throughput, providing different advantages for viral diversity analyses. In general, longer read lengths are accompanied by higher error rates and lower throughput as deep sequencing data are generated; however, the ability of long read results to define more accurate genetic haplotype compositions in viral quasispecies for inter- and intrahost diversity studies is a double-edged sword. Utilizing deep sequencing techniques, these evolutionary studies initially focused on drug mutations or antibody escape mutations of chronic RNA viruses, and recently, more studies have focused on the genetic evolution of emerging or re-emerging viruses, including influenza virus, arboviruses, and enteric viruses. Deep sequencing platforms have also been utilized for rapid sequencing, such as RNA or DNA sequencing as an emerging virus outbreak occurs. Faced with the challenges of exporting specimens to laboratories or establishing conventional sequencers in remote outbreak regions, Hoenen et al. (2016), Hoenen et al. (2016) used a nanopore sequencer at a field diagnostic laboratory in Liberia during the Ebola virus outbreak. Although the nanopore sequencer platform exhibited high error rates relative to those of other short-read platforms, it is the only portable deep sequencer available capable of generating rapid sequencing results to survey viral epidemiology.

In summary, by revealing the viral population either within hosts or between hosts, viral epidemiology and transmission among various hosts from different species or virus dissemination within the host can be well defined, thereby providing valuable information for understanding inter- and intrahost viral evolution. Deep sequencing methods have advanced viral quasispecies investigations and allowed examinations of viral populations at a higher throughput than clone sequencing methods. As the average costs of sequencing each base are gradually reduced, more scientists have utilized deep sequencing methods to examine viral populations. Thus, these methods are likely to become the standard methods of investigating viral diversity in the future.

Acknowledgments

This work was financially supported by the Center of Infectious Disease and Signaling Research, National Cheng Kung University, Aim for the Top University Project, Ministry of Education; Centers of Disease Control, Ministry of Health and Welfare; National Mosquito-Borne Diseases Control Research Center and National Institute of Infectious

Diseases and Vaccinology of the National Health Research Institutes; Ministry of Science and Technology, Taiwan, grant 105-2320-B-006 -032 -MY3.

References

Alame, M., Lacourt, D., Zenagui, R., Mechin, D., Danton, F., Koenig, M., Claustres, M., Cossee, M., 2016. Implementation of a reliable next-generation sequencing strategy for molecular diagnosis of dystrophinopathies. J. Mol. Diagn. 18, 731–740. https://doi.org/10.1016/j.jmoldx.2016.05.003.

Bartolini, B., Lionetti, R., Giombini, E., Sias, C., Taibi, C., Montalbano, M., D'Offizi, G., McPhee, F., Hughes, E.A., Zhou, N., Ippolito, G., Garbuglia, A.R., Capobianchi, M.R., 2015. Dynamics of HCV genotype 4 resistance-associated variants during virologic escape with pIFN/RBV + daclatasvir: a case study using ultra deep pyrosequencing. J. Clin. Virol. 66, 38–43. https://doi.org/10.1016/j.jcv.2015.02.001.

Bednar, M.M., Hauser, B.M., Zhou, S., Jacobson, J.M., Eron Jr., J.J., Frank, I., Swanstrom, R., 2016. Diversity and tropism of HIV-1 rebound virus populations in plasma level after treatment discontinuation. J. Infect. Dis. 214, 403–407. https://doi.org/10.1093/infdis/jiw172.

Bimber, B.N., Dudley, D.M., Lauck, M., Becker, E.A., Chin, E.N., Lank, S.M., Grunenwald, H.L., Caruccio, N.C., Maffitt, M., Wilson, N.A., Reed, J.S., Sosman, J.M., Tarosso, L.F., Sanabani, S., Kallas, E.G., Hughes, A.L., O'Connor, D.H., 2010. Whole-genome characterization of human and simian immunodeficiency virus intrahost diversity by ultradeep pyrosequencing. J. Virol. 84, 12087–12092. https://doi.org/10.1128/JVI.01378-10.

Bonhoeffer, S., Nowak, M.A., 1997. Pre-existence and emergence of drug resistance in HIV-1 infection. Proc. Biol. Sci. 264, 631–637. https://doi.org/10.1098/rspb.1997.0089.

Borucki, M.K., Allen, J.E., Chen-Harris, H., Zemla, A., Vanier, G., Mabery, S., Torres, C., Hullinger, P., Slezak, T., 2013. The role of viral population diversity in adaptation of bovine coronavirus to new host environments. PLoS One 8, e52752. https://doi.org/10.1371/journal.pone.0052752.

Bourret, V., Croville, G., Mariette, J., Klopp, C., Bouchez, O., Tiley, L., Guerin, J.L., 2013. Whole-genome, deep pyrosequencing analysis of a duck influenza A virus evolution in swine cells. Infect. Genet. Evol. 18, 31–41. https://doi.org/10.1016/j.meegid.2013.04.034.

Bourret, V., Croville, G., Mansuy, J.M., Mengelle, C., Mariette, J., Klopp, C., Genthon, C., Izopet, J., Guerin, J.L., 2015. Intra-host viral variability in children clinically infected with H1N1 (2009) pandemic influenza. Infect. Genet. Evol. 33, 47–54. https://doi.org/10.1016/j.meegid.2015.04.009.

Bull, R.A., Eltahla, A.A., Rodrigo, C., Koekkoek, S.M., Walker, M., Pirozyan, M.R., Betz-Stablein, B., Toepfer, A., Laird, M., Oh, S., Heiner, C., Maher, L., Schinkel, J., Lloyd, A.R., Luciani, F., 2016. A method for near full-length amplification and sequencing for six hepatitis C virus genotypes. BMC Genomics 17, 247. https://doi.org/10.1186/s12864-016-2575-8.

Bunnik, E.M., Swenson, L.C., Edo-Matas, D., Huang, W., Dong, W., Frantzell, A., Petropoulos, C.J., Coakley, E., Schuitemaker, H., Harrigan, P.R., van' t Wout, A.B., 2011. Detection of inferred CCR5- and CXCR4-using HIV-1 variants and evolutionary intermediates using ultra-deep pyrosequencing. PLoS Pathog. 7, e1002106. https://doi.org/10.1371/journal.ppat.1002106.

Cannon, N.A., Donlin, M.J., Fan, X., Aurora, R., Tavis, J.E., Virahep, C.S.G., 2008. Hepatitis C virus diversity and evolution in the full open-reading frame during antiviral therapy. PLoS One 3, e2123. https://doi.org/10.1371/journal.pone.0002123.

Chabria, S.B., Gupta, S., Kozal, M.J., 2014. Deep sequencing of HIV: clinical and research applications. Annu. Rev. Genomics Hum. Genet. 15, 295–325. https://doi.org/10.1146/annurev-genom-091212-153406.

Cortes, K.C., Zagordi, O., Perlejewski, K., Laskus, T., Maroszek, K., Bukowska-Osko, I., Pawelczyk, A., Ploski, R., Berak, H., Horban, A., Radkowski, M., 2014. Deep sequencing of hepatitis C virus hypervariable region 1 reveals no correlation between genetic heterogeneity and antiviral treatment outcome. BMC Infect. Dis. 14, 389. https://doi.org/10.1186/1471-2334-14-389.

Cuevas, J.M., Domingo-Calap, P., Sanjuan, R., 2012. The fitness effects of synonymous mutations in DNA and RNA viruses. Mol. Biol. Evol. 29, 17–20. https://doi.org/10.1093/molbev/msr179.

Dialdestoro, K., Sibbesen, J.A., Maretty, L., Raghwani, J., Gall, A., Kellam, P., Pybus, O.G., Hein, J., Jenkins, P.A., 2016. Coalescent inference using serially sampled, high-throughput sequencing data from intrahost HIV infection. Genetics 202, 1449–1472. https://doi.org/10.1534/genetics.115.177931.

Do, L.A., Wilm, A., Van Doorn, H.R., Lam, H.M., Sim, S., Sukumaran, R., Tran, A.T., Nguyen, B.H., Tran, T.T., Tran, Q.H., Vo, Q.B., Dac, N.A., Trinh, H.N., Nguyen, T.T., Binh, B.T., Le, K., Nguyen, M.T., Thai, Q.T., Vo, T.V., Ngo, N.Q., Dang, T.K., Cao, N.H., Tran, T.V., Ho, L.V., Farrar, J., Jong, M., Chen, S., Nagarajan, N., Bryant, J.E., Hibberd, M.L., 2015. Direct whole-genome deep-sequencing of human respiratory syncytial virus A and B from Vietnamese children identifies distinct patterns of inter- and intra-host evolution. J. Gen. Virol. 96, 3470–3483. https://doi.org/10.1099/jgv.0.000298.

Domingo, E., Holland, J.J., 1997. RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. 51, 151–178. https://doi.org/10.1146/annurev.micro.51.1.151.

Domingo, E., Escarmis, C., Lazaro, E., Manrubia, S.C., 2005. Quasispecies dynamics and RNA virus extinction. Virus Res. 107, 129–139. https://doi.org/10.1016/j.virusres.2004.11.003.

Drake, J.W., 1993. Rates of spontaneous mutation among RNA viruses. Proc. Natl. Acad. Sci. U. S. A. 90, 4171–4175.

Duffy, S., Shackelton, L.A., Holmes, E.C., 2008. Rates of evolutionary change in viruses: patterns and determinants. Nat. Rev. Genet. 9, 267–276. https://doi.org/10.1038/nrg2323.

Eigen, M., Schuster, P., 1977. The hypercycle. A principle of natural self-organization. Part A: emergence of the hypercycle. Naturwissenschaften 64, 541–565.

Eshaghi, A., Shalhoub, S., Rosenfeld, P., Li, A., Higgins, R.R., Stogios, P.J., Savchenko, A., Bastien, N., Li, Y., Rotstein, C., Gubbay, J.B., 2014. Multiple influenza A (H3N2) mutations conferring resistance to neuraminidase inhibitors in a bone marrow transplant recipient. Antimicrob. Agents Chemother. 58, 7188–7197. https://doi.org/10.1128/AAC.03667-14.

Fischer, W., Ganusov, V.V., Giorgi, E.E., Hraber, P.T., Keele, B.F., Leitner, T., Han, C.S., Gleasner, C.D., Green, L., Lo, C.C., Nag, A., Wallstrom, T.C., Wang, S., McMichael, A.J., Haynes, B.F., Hahn, B.H., Perelson, A.S., Borrow, P., Shaw, G.M., Bhattacharya, T., Korber, B.T., 2010. Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. PLoS One 5, e12303. https://doi.org/10.1371/journal.pone.0012303.

Fisher, K.E., Zhang, L., Wang, J., Smith, G.H., Newman, S., Schneider, T.M., Pillai, R.N., Kudchadkar, R.R., Owonikoko, T.K., Ramalingam, S.S., Lawson, D.H., Delman, K.A., El-Rayes, B.F., Wilson, M.M., Sullivan, H.C., Morrison, A.S., Balci, S., Adsay, N.V., Gal, A.A., Sica, G.L., Saxe, D.F., Mann, K.P., Hill, C.E., Khuri, F.R., Rossi, M.R., 2016. Clinical validation and implementation of a targeted next-generation sequencing assay to detect somatic variants in non-small cell lung, melanoma, and gastrointestinal malignancies. J. Mol. Diagn. 18, 299–315. https://doi.org/10.1016/j.jmoldx.2015.11.006.

Fordyce, S.L., Bragstad, K., Pedersen, S.S., Jensen, T.G., Gahrn-Hansen, B., Daniels, R., Hay, A., Kampmann, M.L., Bruhn, C.A., Moreno-Mayar, J.V., Avila-Arcos, M.C., Gilbert, M.T., Nielsen, L.P., 2013. Genetic diversity among pandemic 2009 influenza viruses isolated from a transmission chain. Virol. J. 10, 116. https://doi.org/10.1186/1743-422X-10-116.

Gall, A., Kaye, S., Hue, S., Bonsall, D., Rance, R., Baillie, G.J., Fidler, S.J., Weber, J.N., McClure, M.O., Kellam, P., Investigators, S.T., 2013. Restriction of V3 region sequence divergence in the HIV-1 envelope gene during antiretroviral treatment in a cohort of recent seroconverters. Retrovirology 10, 8. https://doi.org/10.1186/1742-4690-10-8.

Gaschen, B., Taylor, J., Yusim, K., Foley, B., Gao, F., Lang, D., Novitsky, V., Haynes, B., Hahn, B.H., Bhattacharya, T., Korber, B., 2002. Diversity considerations in HIV-1 vaccine selection. Science 296, 2354–2360. https://doi.org/10.1126/science.1070441.

Hanada, K., Suzuki, Y., Gojobori, T., 2004. A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes. Mol. Biol. Evol. 21, 1074–1080. https://doi.org/10.1093/molbev/msh109.

Hansen, T.A., Mollerup, S., Nguyen, N.P., White, N.E., Coghlan, M., Alquezar-Planas, D.E., Joshi, T., Jensen, R.H., Fridholm, H., Kjartansdottir, K.R., Mourier, T., Warnow, T., Belsham, G.J., Bunce, M., Willerslev, E., Nielsen, L.P., Vinner, L., Hansen, A.J., 2016. High diversity of picornaviruses in rats from different continents revealed by deep sequencing. Emerg. Microbes Infect. 5, e90. https://doi.org/10.1038/emi.2016.90.

Hedegaard, D.L., Tully, D.C., Rowe, I.A., Reynolds, G.M., Bean, D.J., Hu, K., Davis, C., Wilhelm, A., Ogilvie, C.B., Power, K.A., Tarr, A.W., Kelly, D., Allen, T.M., Balfe, P., McKeating, J.A., 2017. High resolution sequencing of hepatitis C virus reveals limited intra-hepatic compartmentalization in end-stage liver disease. J. Hepatol. 66, 28–38. https://doi.org/10.1016/j.jhep.2016.07.048.

Hoenen, T., 2016. Sequencing of ebola virus genomes using nanopore technology. Bio. 6. https://doi.org/10.21769/BioProtoc.1998.

Hoenen, T., Groseth, A., Rosenke, K., Fischer, R.J., Hoenen, A., Judson, S.D., Martellaro, C., Falzarano, D., Marzi, A., Squires, R.B., Wollenberg, K.R., de Wit, E., Prescott, J., Safronetz, D., van Doremalen, N., Bushmaker, T., Feldmann, F., McNally, K., Bolay, F.K., Fields, B., Sealy, T., Rayfield, M., Nichol, S.T., Zoon, K.C., Massaquoi, M., Munster, V.J., Feldmann, H., 2016. Nanopore sequencing as a rapidly deployable ebola outbreak tool. Emerg. Infect. Dis. 22, 331–334. https://doi.org/10.3201/eid2202.151796.

Huang, S.W., Huang, Y.H., Tsai, H.P., Kuo, P.H., Wang, S.M., Liu, C.C., Wang, J.R., 2017. A selective bottleneck shapes the evolutionary mutant spectra of enterovirus A71 during viral dissemination in humans. J. Virol. 91. https://doi.org/10.1128/JVI.01062-17.

Hughes, A.L., O'Connor, S., Dudley, D.M., Burwitz, B.J., Bimber, B.N., O'Connor, D., 2010. Dynamics of haplotype frequency change in a CD8 + TL epitope of simian immunodeficiency virus. Infect. Genet. Evol. 10, 555–560. https://doi.org/10.1016/j.meegid.2010.02.001.

Jabara, C.B., Jones, C.D., Roach, J., Anderson, J.A., Swanstrom, R., 2011. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. Proc. Natl. Acad. Sci. U. S. A. 108, 20166–20171. https://doi.org/10.1073/pnas.1110064108.

Jansen, L., Welkers, M.R.A., van Dort, K.A., Takkenberg, R.B., Lopatin, U., Zaaijer, H.L., de Jong, M.D., Reesink, H.W., Kootstra, N.A., 2017. Viral minority variants in the core promoter and precore region identified by deep sequencing are associated with response to peginterferon and adefovir in HBeAg negative chronic hepatitis B patients. Antiviral Res. 145, 87–95. https://doi.org/10.1016/j.antiviral.2017.07.013.

Jenkins, G.M., Rambaut, A., Pybus, O.G., Holmes, E.C., 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J. Mol. Evol. 54, 156–165. https://doi.org/10.1007/s00239-001-0064-3.

Johnson, J.A., Li, J.F., Wei, X., Lipscomb, J., Irlbeck, D., Craig, C., Smith, A., Bennett, D.E., Monsour, M., Sandstrom, P., Lanier, E.R., Heneine, W., 2008. Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naive populations and associate with reduced treatment efficacy. PLoS Med. 5, e158. https://doi.org/10.1371/journal.pmed.0050158.

Jonges, M., Welkers, M.R., Jeeninga, R.E., Meijer, A., Schneeberger, P., Fouchier, R.A., de

Jong, M.D., Koopmans, M., 2014. Emergence of the virulence-associated PB2 E627K substitution in a fatal human case of highly pathogenic avian influenza virus A(H7N7) infection as determined by Illumina ultra-deep sequencing. J. Virol. 88, 1694–1702. https://doi.org/10.1128/JVI.02044-13.

Kluk, M.J., Lindsley, R.C., Aster, J.C., Lindeman, N.I., Szeto, D., Hall, D., Kuo, F.C., 2016. Validation and implementation of a custom next-generation sequencing clinical assay for hematologic malignancies. J. Mol. Diagn. 18, 507–515. https://doi.org/10.1016/j.jmoldx.2016.02.003.

Kuroda, M., Katano, H., Nakajima, N., Tobiume, M., Ainai, A., Sekizuka, T., Hasegawa, H., Tashiro, M., Sasaki, Y., Arakawa, Y., Hata, S., Watanabe, M., Sata, T., 2010. Characterization of quasispecies of pandemic 2009 influenza A virus (A/H1N1/2009) by de novo sequencing using a next-generation DNA sequencer. PLoS One 5, e10256. https://doi.org/10.1371/journal.pone.0010256.

Lauring, A.S., Andino, R., 2010. Quasispecies theory and the behavior of RNA viruses. PLoS Pathog. 6, e1001005. https://doi.org/10.1371/journal.ppat.1001005.

Lequime, S., Fontaine, A., Ar Gouilh, M., Moltini-Conclois, I., Lambrechts, L., 2016. Genetic Drift, purifying selection and vector genotype shape dengue virus intra-host genetic diversity in mosquitoes. PLoS Genet. 12, e1006111. https://doi.org/10.1371/journal.pgen.1006111.

Lowe, C.F., Merrick, L., Harrigan, P.R., Mazzulli, T., Sherlock, C.H., Ritchie, G., 2016. Implementation of next-generation sequencing for hepatitis B virus resistance testing and genotyping in a clinical microbiology laboratory. J. Clin. Microbiol. 54, 127–133. https://doi.org/10.1128/JCM.02229-15.

Moldovan, N., Balazs, Z., Tombacz, D., Csabai, Z., Szucs, A., Snyder, M., Boldogkoi, Z., 2017a. Multi-platform analysis reveals a complex transcriptome architecture of a circovirus. Virus Res. 237, 37–46. https://doi.org/10.1016/j.virusres.2017.05.010.

Moldovan, N., Tombacz, D., Szucs, A., Csabai, Z., Snyder, M., Boldogkoi, Z., 2017b. Multi-platform sequencing approach reveals a novel transcriptome profile in pseudorabies virus. Front. Microbiol. 8, 2708. https://doi.org/10.3389/fmicb.2017.02708.

Moncla, L.H., Zhong, G., Nelson, C.W., Dinis, J.M., Mutschler, J., Hughes, A.L., Watanabe, T., Kawaoka, Y., Friedrich, T.C., 2016. Selective bottlenecks shape evolutionary pathways taken during mammalian adaptation of a 1918-like avian influenza virus. Cell Host Microbe 19, 169–180. https://doi.org/10.1016/j.chom.2016.01.011.

Nowak, M.A., 1992. What is a quasispecies? Trends Ecol. Evol. (Amst.) 7, 118–121. https://doi.org/10.1016/0169-5347(92)90145-2.

Nowak, M.A., Anderson, R.M., McLean, A.R., Wolfs, T.F., Goudsmit, J., May, R.M., 1991. Antigenic diversity thresholds and the development of AIDS. Science 254, 963–969.

Ode, H., Matsuda, M., Matsuoka, K., Hachiya, A., Hattori, J., Kito, Y., Yokomaku, Y., Iwatani, Y., Sugiura, W., 2015. Quasispecies analyses of the HIV-1 near-full-length genome with illumina MiSeq. Front. Microbiol. 6, 1258. https://doi.org/10.3389/fmicb.2015.01258.

Palmer, S., Kearney, M., Maldarelli, F., Halvas, E.K., Bixby, C.J., Bazmi, H., Rock, D., Falloon, J., Davey Jr., R.T., Dewar, R.L., Metcalf, J.A., Hammer, S., Mellors, J.W., Coffin, J.M., 2005. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. J. Clin. Microbiol. 43, 406–413. https://doi.org/10.1128/JCM.43.1.406-413.2005.

Parameswaran, P., Charlebois, P., Tellez, Y., Nunez, A., Ryan, E.M., Malboeuf, C.M., Levin, J.Z., Lennon, N.J., Balmaseda, A., Harris, E., Henn, M.R., 2012. Genome-wide patterns of intrahuman dengue virus diversity reveal associations with viral phylogenetic clade and interhost diversity. J. Virol. 86, 8546–8558. https://doi.org/10.1128/JVI.00736-12.

Parameswaran, P., Wang, C., Trivedi, S.B., Eswarappa, M., Montoya, M., Balmaseda, A., Harris, E., 2017. Intrahost selection pressures drive rapid dengue virus microevolution in acute human infections. Cell Host Microbe 22, 400–410. https://doi.org/10.1016/j.chom.2017.08.003. e5.

Poon, L.L., Song, T., Rosenfeld, R., Lin, X., Rogers, M.B., Zhou, B., Sebra, R., Halpin, R.A., Guan, Y., Twaddle, A., DePasse, J.V., Stockwell, T.B., Wentworth, D.E., Holmes, E.C., Greenbaum, B., Peiris, J.S., Cowling, B.J., Ghedin, E., 2016. Quantifying influenza virus diversity and transmission in humans. Nat. Genet. 48, 195–200. https://doi.org/10.1038/ng.3479.

Prabhakara, S., Malhotra, R., Acharya, R., Poss, M., 2013. Mutant-bin: unsupervised haplotype estimation of viral population diversity without reference genome. J. Comput. Biol. 20, 453–463. https://doi.org/10.1089/cmb.2012.0174.

Redd, A.D., Collinson-Streng, A.N., Chatziandreou, N., Mullis, C.E., Laeyendecker, O., Martens, C., Ricklefs, S., Kiwanuka, N., Nyein, P.H., Lutalo, T., Grabowski, M.K., Kong, X., Manucci, J., Sewankambo, N., Wawer, M.J., Gray, R.H., Porcella, S.F., Fauci, A.S., Sagar, M., Serwadda, D., Quinn, T.C., 2012. Previously transmitted HIV-1 strains are preferentially selected during subsequent sexual transmissions. J. Infect. Dis. 206, 1433–1442. https://doi.org/10.1093/infdis/jis503.

Rodriguez-Roche, R., Blanc, H., Borderia, A.V., Diaz, G., Henningsson, R., Gonzalez, D., Santana, E., Alvarez, M., Castro, O., Fontes, M., Vignuzzi, M., Guzman, M.G., 2016. Increasing clinical severity during a dengue virus type 3 cuban epidemic: deep sequencing of evolving viral populations. J. Virol. 90, 4320–4333. https://doi.org/10.1128/JVI.02647-15.

Rogers, M.B., Song, T., Sebra, R., Greenbaum, B.D., Hamelin, M.E., Fitch, A., Twaddle, A., Cui, L., Holmes, E.C., Boivin, G., Ghedin, E., 2015. Intrahost dynamics of antiviral resistance in influenza A virus reflect complex patterns of segment linkage, reassortment, and natural selection. MBio 6. https://doi.org/10.1128/mBio.02464-14.

Romano, C.M., Lauck, M., Salvador, F.S., Lima, C.R., Villas-Boas, L.S., Araujo, E.S., Levi, J.E., Pannuti, C.S., O'Connor, D., Kallas, E.G., 2013. Inter- and intra-host viral diversity in a large seasonal DENV2 outbreak. PLoS One 8, e70318. https://doi.org/10.1371/journal.pone.0070318.

Roy, S., LaFramboise, W.A., Nikiforov, Y.E., Nikiforova, M.N., Routbort, M.J., Pfeifer, J., Nagarajan, R., Carter, A.B., Pantanowitz, L., 2016. Next-generation sequencing informatics: challenges and strategies for implementation in a clinical environment. Arch. Pathol. Lab. Med. 140, 958–975. https://doi.org/10.5858/arpa.2015-0507-RA.

Rozera, G., Abbate, I., Vlassi, C., Giombini, E., Lionetti, R., Selleri, M., Zaccaro, P., Bartolini, B., Corpolongo, A., D'Offizi, G., Baiocchini, A., Del Nonno, F., Ippolito, G., Capobianchi, M.R., 2014. Quasispecies tropism and compartmentalization in gut and peripheral blood during early and chronic phases of HIV-1 infection: possible correlation with immune activation markers. Clin. Microbiol. Infect. 20, O157–O166. https://doi.org/10.1111/1469-0691.12367.

Sede, M.M., Moretti, F.A., Laufer, N.L., Jones, L.R., Quarleri, J.F., 2014. HIV-1 tropism dynamics and phylogenetic analysis from longitudinal ultra-deep sequencing data of CCR5- and CXCR4-using variants. PLoS One 9, e102857. https://doi.org/10.1371/journal.pone.0102857.

Shirvani-Dastgerdi, E., Winer, B.Y., Celia-Terrassa, T., Kang, Y., Tabernero, D., Yagmur, E., Rodriguez-Frias, F., Gregori, J., Luedde, T., Trautwein, C., Ploss, A., Tacke, F., 2017. Selection of the highly replicative and partially multidrug resistant rtS78T HBV polymerase mutation during TDF-ETV combination therapy. J. Hepatol. 67, 246–254. https://doi.org/10.1016/j.jhep.2017.03.027.

Sim, S., Aw, P.P., Wilm, A., Teoh, G., Hue, K.D., Nguyen, N.M., Nagarajan, N., Simmons, C.P., Hibberd, M.L., 2015. Tracking dengue virus intra-host genetic diversity during human-to-mosquito transmission. PLoS Negl. Trop. Dis. 9, e0004052. https://doi.org/10.1371/journal.pntd.0004052.

Sobel Leonard, A., McClain, M.T., Smith, G.J., Wentworth, D.E., Halpin, R.A., Lin, X., Ransier, A., Stockwell, T.B., Das, S.R., Gilbert, A.S., Lambkin-Williams, R., Ginsburg, G.S., Woods, C.W., Koelle, K., 2016. Deep sequencing of influenza A virus from a human challenge study reveals a selective bottleneck and only limited intrahost genetic diversification. J. Virol. 90, 11247–11258. https://doi.org/10.1128/JVI.01657-16.

Stapleford, K.A., Moratorio, G., Henningsson, R., Chen, R., Matheus, S., Enfissi, A., Weissglas-Volkov, D., Isakov, O., Blanc, H., Mounce, B.C., Dupont-Rouzeyrol, M., Shomron, N., Weaver, S., Fontes, M., Rousset, D., Vignuzzi, M., 2016. Whole-genome sequencing analysis from the chikungunya virus caribbean outbreak reveals novel evolutionary genomic elements. PLoS Negl. Trop. Dis. 10, e0004402. https://doi.org/10.1371/journal.pntd.0004402.

Tellez-Sosa, J., Rodriguez, M.H., Gomez-Barreto, R.E., Valdovinos-Torres, H., Hidalgo, A.C., Cruz-Hervert, P., Luna, R.S., Carrillo-Valenzo, E., Ramos, C., Garcia-Garcia, L., Martinez-Barnetche, J., 2013. Using high-throughput sequencing to leverage surveillance of genetic diversity and oseltamivir resistance: a pilot study during the 2009 influenza A(H1N1) pandemic. PLoS One 8, e67010. https://doi.org/10.1371/journal.pone.0067010.

Tsai, K.N., Chen, G.W., 2011. Influenza genome diversity and evolution. Microbes Infect. 13, 479–488. https://doi.org/10.1016/j.micinf.2011.01.013.

Tsibris, A.M., Korber, B., Arnaout, R., Russ, C., Lo, C.C., Leitner, T., Gaschen, B., Theiler, J., Paredes, R., Su, Z., Hughes, M.D., Gulick, R.M., Greaves, W., Coakley, E., Flexner, C., Nusbaum, C., Kuritzkes, D.R., 2009. Quantitative deep sequencing reveals dynamic HIV-1 escape and large population shifts during CCR5 antagonist therapy in vivo. PLoS One 4, e5683. https://doi.org/10.1371/journal.pone.0005683.

van Beek, J., de Graaf, M., Smits, S.L., Schapendonk, C.M.E., Verjans, G., Vennema, H., van der Eijk, A.A., Phan, M.V.T., Cotten, M.L., Koopmans, M.P.G., 2017. Whole genome next generation sequencing to study within-host evolution of chronic norovirus infection among immunocompromised patients. J. Infect. Dis. https://doi.org/10.1093/infdis/jix520.

Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., Andino, R., 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439, 344–348. https://doi.org/10.1038/nature04388.

Wang, G.P., Sherrill-Mix, S.A., Chang, K.M., Quince, C., Bushman, F.D., 2010. Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. J. Virol. 84, 6218–6228. https://doi.org/10.1128/JVI.02271-09.

Wilker, P.R., Dinis, J.M., Starrett, G., Imai, M., Hatta, M., Nelson, C.W., O'Connor, D.H., Hughes, A.L., Neumann, G., Kawaoka, Y., Friedrich, T.C., 2013. Selection on haemagglutinin imposes a bottleneck during mammalian transmission of reassortant H5N1 influenza viruses. Nat. Commun. 4, 2636. https://doi.org/10.1038/ncomms3636.

Woo, H.J., Reifman, J., 2012. A quantitative quasispecies theory-based model of virus escape mutation under immune selection. Proc. Natl. Acad. Sci. U. S. A. 109, 12980–12985. https://doi.org/10.1073/pnas.1117201109.

Wu, N.C., Young, A.P., Dandekar, S., Wijersuriya, H., Al-Mawsawi, L.Q., Wu, T.T., Sun, R., 2013. Systematic identification of H274Y compensatory mutations in influenza A virus neuraminidase by high-throughput screening. J. Virol. 87, 1193–1199. https://doi.org/10.1128/JVI.01658-12.

Xue, Y., Wang, M.J., Yang, Z.T., Yu, D.M., Han, Y., Huang, D., Zhang, D.H., Zhang, X.X., 2017. Clinical features and viral quasispecies characteristics associated with infection by the hepatitis B virus G145R immune escape mutant. Emerg. Microbes Infect. 6, e15. https://doi.org/10.1038/emi.2017.2.

Zanini, F., Brodin, J., Albert, J., Neher, R.A., 2016. Error rates, PCR recombination, and sampling depth in HIV-1 whole genome deep sequencing. Virus Res. https://doi.org/10.1016/j.virusres.2016.12.009.

Zhu, Y.O., Aw, P.P.K., de Sessions, P.F., Hong, S., See, L.X., Hong, L.Z., Wilm, A., Li, C.H., Hue, S., Lim, S.G., Nagarajan, N., Burkholder, W.F., Hibberd, M., 2017. Single-virion sequencing of lamivudine-treated HBV populations reveal population evolution dynamics and demographic history. BMC Genomics 18, 829. https://doi.org/10.1186/s12864-017-4217-1.