

Laboratory Methods in Molecular Epidemiology: Viral Infections*

RICARDO PARREIRA¹

¹Unidade de Microbiologia Médica/Global Health and Tropical Medicine (GHTM) Research Centre, Instituto de Higiene e Medicina Tropical (IHMT), Universidade Nova de Lisboa (UNL), 1349-008 Lisboa, Portugal

ABSTRACT Viruses, which are the most abundant biological entities on the planet, have been regarded as the “dark matter” of biology in the sense that despite their ubiquity and frequent presence in large numbers, their detection and analysis are not always straightforward. The majority of them are very small (falling under the limit of 0.5 µm), and collectively, they are extraordinarily diverse. In fact, the majority of the genetic diversity on the planet is found in the so-called virosphere, or the world of viruses. Furthermore, the most frequent viral agents of disease in humans display an RNA genome, and frequently evolve very fast, due to the fact that most of their polymerases are devoid of proofreading activity. Therefore, their detection, genetic characterization, and epidemiological surveillance are rather challenging. This review (part of the Curated Collection on Advances in Molecular Epidemiology of Infectious Diseases) describes many of the methods that, throughout the last few decades, have been used for viral detection and analysis. Despite the challenge of having to deal with high genetic diversity, the majority of these methods still depend on the amplification of viral genomic sequences, using sequence-specific or sequence-independent approaches, exploring thermal profiles or a single nucleic acid amplification temperature. Furthermore, viral populations, and especially those with RNA genomes, are not usually genetically uniform but encompass swarms of genetically related, though distinct, viral genomes known as viral quasispecies. Therefore, sequence analysis of viral amplicons needs to take this fact into consideration, as it constitutes a potential analytic problem. Possible technical approaches to deal with it are also described here. *This article is part of a curated collection.

INTRODUCTION

The study of viruses as the focus of molecular epidemiology investigations involves addressing their possible roles as etiological agents of disease and how they

influence disease dynamics in a given population. The focal point of the intended analyses is, most of the time, viral genomes and their variation. Their detection and analysis allow us to (i) correctly identify viral agents (taxonomy), (ii) investigate their virulence by tentative association of viral genetic features to phenotypes, or even (iii) analyze their evolution and dispersal in time and space (1). In turn, these provide invaluable information on the structure of the viral populations and evidence for their etiological role in disease, which may support the design and implementation of adequate public health policies and disease control programs.

The task at hand is somewhat daunting, as viruses are the most abundant biological entities on the planet (2). Their estimated numbers may reach up to 10^{31} viral particles in any given environment, where their numbers top at least 1 log that of cellular organisms (3). In the majority of the cases, their small size, rapid evolution (especially in the case of the RNA viruses), genomic

Received: 16 March 2018, **Accepted:** 30 April 2018,

Published: 2 November 2018

Editors: Lee W. Riley, Divisions of Infectious Diseases and Vaccinology, School of Public Health, University of California, Berkeley, Berkeley, CA; Ronald E. Blanton, Center for Global Health & Diseases, Case Western Reserve University, Cleveland, OH

Citation: Parreira R. 2018. Laboratory methods in molecular epidemiology: viral infections. *Microbiol Spectrum* 6(6):AME-0003-2018. [doi:10.1128/microbiolspec.AME-0003-2018](https://doi.org/10.1128/microbiolspec.AME-0003-2018).

Correspondence: Ricardo Parreira, ricardo@ihmt.unl.pt

© 2018 American Society for Microbiology. All rights reserved.

Curated Collection: [Advances in Molecular Epidemiology of Infectious Diseases](#).

flexibility and size, high turnover rate, and diversity may render them relatively difficult to study. Their analysis is complicated by the lack of a shared genetic marker, such as ribosomal DNA sequences of all prokaryotes and eukaryotes, that may be universally targeted for detection and analysis (4). Furthermore, many viruses (especially those with RNA genomes) exist not as stable genetic entities but rather as swarms of many possible, phylogenetically related genetic variants, known as viral quasispecies (5), complicating the characterization of their population structure. The task is further aggravated by the fact that, as an alternative to full-length genome analysis, the investigation of individual viral genetic markers may fail to identify either those molecular traits with epidemiological significance, or detect low-frequency relevant viral variants (6, 7), conveying only a fragmented version of both the architecture of the viral genome and the intrahost viral population structure (8).

Although the number of viruses currently known to be pathogenic to humans is relatively small (≈ 220) (9), epidemiological studies strongly suggest this number to be a rough underestimation of reality (10). This is especially true in view of the fact that many viruses that infect humans are considered emerging pathogens of zoonotic origin. Many outbreaks recently caused by the avian H5N1 (11) and H7N9 (12) influenza viruses, the severe acute respiratory syndrome (SARS-CoV) and Middle East respiratory syndrome (13) coronaviruses, hepatitis E virus (14), hantaviruses (15), ebolaviruses (16), and Zika virus (17) have clearly highlighted the need for continual surveillance of putative pathogenic viruses also among domestic and sylvatic animal populations. Even outside epidemiological settings such as viral outbreaks, many clinical syndromes, including autoimmune, degenerative, and neoplastic conditions, still have no identifiable etiology, opening the possibility that many new viruses await discovery (18). Therefore, the characterization of viruses, whether truly novel or not, transcends pure academic interest, departing from being a scientific idiosyncrasy to having indisputable epidemiological/clinical relevance.

Characterizing viruses may rely on a wide range of technical possibilities that include either molecular methods or more “classical” techniques, such as electron microscopy, serology, or isolation in cell culture, suckling mice (or other small mammals), or even embryonated eggs (19). Although successful viral isolation should always be pursued (allowing for *a posteriori* characterization of the virus in question), it may also select from the viral quasispecies those genetic forms that are better suited for replication under the labora-

tory conditions used, especially if repeated passage in culture (for successive viral enrichment) is enforced. However, many viruses defy isolation. In these cases, one may be forced to study them in the context of a complex sample (see below), where they may exist not as single viral species but as part of a viral community (virome), exclusively using molecular approaches. Moreover, their presence in variable titers in these samples frequently contrasts with the incomparably more abundant nonviral nucleic acids, compromising the sensitivity and specificity of the methods used for their detection.

Over time, the development of molecular methods for viral characterization has followed different analytic paths. These include, for example, nucleic acid cross-hybridization, targeted amplification using PCR (with specific or degenerate primers), or sequence-independent amplification methods such as random PCR (rPCR), single-primer amplification (SISPA), displacement amplification techniques based on the phi29 DNA polymerase, or even nucleic acid subtraction hybridization methods (20) (Table 1). More recently, a more generalized access to next-generation sequencing (NGS) platforms (21) has facilitated the study of viruses and resulted in the development of metagenomics, i.e., the culture-independent study of collective microbial populations in a sample via analysis of its nucleic acid content (22). In the last decade, although NGS has certainly imparted considerable momentum to the study of viruses, its use in many laboratory settings may still be limited by logistic, technical, or financial constraints.

This review in the Advances in Molecular Epidemiology of Infectious Diseases Curated Collection addresses a number of technical possibilities that may be explored for the characterization of viruses in the context of molecular epidemiology. Despite each method’s advantages and downsides, their successful implementation also depends on judicious case identification and selection, rigorous specimen collection and adequate preservation, and downstream processing under appropriate biocontainment conditions, possibly suggested by the type of symptoms involved, and suspected transmission routes (18). Furthermore, due to both inter- and intraindividual viral load fluctuations, serial collections of biological specimens from infected individuals are ideal. Last, though not least, sample collection should also not be necessarily restricted to the sites where symptoms are revealed, as the identification of a virus away from where symptoms manifest may bring about important information regarding viral tropism and transmission dynamics.

TABLE 1 Techniques used for virus discovery and genomic characterization, and their potentials/limitations^a

Method	Analytical potential	Limitations	Virus(es) detected
Virus isolation in cell culture, animals (insects or small mammals), or embryonated eggs	Isolation of infectious viruses and preparation of a viral reference stock. Allows for characterization of viral replication, definition of a potential host range, and assessment of pathogenicity. Usually not suitable in detecting pathogens in an outbreak scenario due to requirement for a long time.	Time-consuming and expensive. Viruses may not replicate or may require adaptation to under laboratory conditions. Viral diversity may be partially lost (reduction in quasispecies complexity) and/or one viral strain may dominate the viral population.	CHIKV, DENV, HHV-1 (HSV), HHV-5 (CMV), HIV, SARS-CoV, WNV, ZIKV
Electron microscopy	Direct observation of viral particles. Virus morphological features may suggest their putative identity and point towards the choice of primers for viral genome amplification. May aid virus identification in an outbreak scenario.	Labor-intensive and lacks sensitivity. May require virus concentration (e.g., filtration, PEG precipitation, and/or ultracentrifugation).	Any virus
Microarrays	Putative comparative analysis of one viral genome to many others, allowing for potential novel virus discovery, and genome characterization (including genotyping). May aid virus identification in an outbreak scenario.	Divergent/novel viruses may be missed. Spurious hybridizations may occur (false-positive identifications). Hybridization profiles may be difficult to interpret.	CardV, HAdV, HBoV, HCV, HEV, HIV, HPV, HPaV, multiple respiratory and CNS infection viruses, influenza A viruses (e.g., H1N1 and H5N1), RhiV HHV-8, TTV
Representational difference analysis (cyclic subtractive hybridization and amplification)	Allows sensitive identification of differences between complex DNA samples. Can be used for the identification of unknown pathogens and the characterization of polymorphic markers, as well as for the identification of differentially expressed genes.	Requires multiple rounds of hybridization, subtraction, and amplification. Requires for tester and driver samples to be sufficiently similar for only the differences between them to stand out.	
Virus genomic amplification using sequence-dependent amplification	Rapid (may support viral detection during an outbreak) viral genomic detection using either consensus, degenerate, or CODEHOP primers and flexible amplification protocols (heminested/nested PCR or RT-PCR). May be combined with transcription-based amplification methods (e.g., NASBA) or isothermal methods based on DNA strand displacement (e.g., RCA and LAMP). Especially suited for discovery of genotypes/variants of known viruses.	Divergent viruses may be missed due to primer/target mismatch precluding genome amplification. Primer design may be challenging due to high viral diversity or lack of reference sequences in the sequence databases.	HAdV, HantV, HHV-4, HIV, hepatitis viruses (e.g., HAV, HBV, and HCV), HMPV, NorV, RSV, influenza A virus, multiple flavi- and phleboviruses
Sequence-independent amplification of viral genomes (SISPA, VIDISCA, rPCR, whole-genome amplification)	Viral genomic detection/amplification is independent of <i>a priori</i> knowledge of its nature/sequence. May be combined with high-throughput sequencing methods for detection/characterization of multiple viral genomes in a single sample. Especially useful for analysis of viruses that fail to replicate <i>in vitro</i> , as well as for viral discovery.	May lack sensitivity due to the presence of nonviral nucleic acids. Usually requires sample pretreatment (filtration, nuclease digestion, centrifugation), which may further reduce detection sensitivity. May introduce bias towards the most frequently represented viral genome (in the case that multiple viral strains are present). Technically demanding.	AstV, HPaV, SARS-CoV and other HCoV, multiple anelloviruses and parvoviruses, NorV, RotV
Virus genome characterization using Sanger sequencing	Straightforward access to high-quality (low error rate) large sequence reads, from either direct sequencing of PCR products or recombinant DNA clones. Allows viral detection and genotyping.	Population sequencing may not detect minor sequence variants (fails to fully characterize viral quasispecies). Sequence representation biases may occur in the case that the envisaged analysis requires a cloning step.	Any DNA or RNA virus

(continued)

TABLE 1 Techniques used for virus discovery and genomic characterization, and their potentials/limitations^a (*continued*)

Method	Analytical potential	Limitations	Virus(es) detected
Virus genome characterization using high-throughput/deep-sequencing platforms	Allows for unbiased detection/characterization (including genotyping) of multiple virus genomes in any given sample, especially when combined with sequence-independent amplification methods and metagenomic analytical approaches. Third-generation sequencing platforms do not require <i>a priori</i> amplification of viral genomes. Being fast, it is very much suitable for pathogen detection in an outbreak scenario. High performance in detecting low-representation genomic variants (e.g., early detection of drug-resistant mutants).	High number of sequence reads requires high-capacity computational capacity. Sequence assembly may be challenging especially when multiple viral genomes are sequenced in parallel (possibility of chimeric assembly) and/or <i>de novo</i> sequence assembly is envisaged due to lack of a reference for comparison. Prone to sequencing errors. Virus enrichment techniques are still usually required before analysis can be performed.	EBOV, HBV, HCV, HIV, HPV, many new tick-borne (e.g., SFTSV), mosquito-borne (including DENV, YFV, and ZIKV), and sand fly-borne (e.g., Arrabida virus) viruses, influenza A viruses, MCPyV

^aAstV, astroviruses; CardV, cardioviruses; CHIKV, chikungunya virus; CMV, cytomegalovirus; CNS, central nervous system; DENV, dengue virus; EBOV, Ebola virus; HAdV, human adenoviruses; HantV, hantaviruses; HAV, hepatitis A virus; HBV, hepatitis B virus; HBoV, human bocavirus; HCV, hepatitis C virus; HCoV, human coronaviruses; HEV, human enterovirus, including hepatitis E virus; HHV, human herpesvirus; HMPV, human metapneumovirus; HPaV, human parechovirus; HPV, human papillomavirus; MCPyV, Merkel cell polyomavirus; NorV, noroviruses; RhiV, rhinoviruses; RotV, rotaviruses; RSV, respiratory syncytial virus; SARS-CoV, SARS coronavirus; SFTSV, severe fever with thrombocytopenia syndrome virus; TTV, Torque Teno virus; WNV, West Nile virus; YFV, yellow fever virus; ZIKV, Zika virus.

PARTICLE-ASSOCIATED NUCLEIC ACID TARGETING: ANALYSIS OF VIRAL COMMUNITIES AND COMPLEX SAMPLES

The analysis of viromes in complex biological samples (all those that are characterized by the presence of a large amount of nonviral contaminating nucleic acids from multiple sources) calls for the use of methods that make possible the purification, and concomitant concentration, of the viral fraction prior to its analysis (23). These methods follow a general strategy known as particle-associated nucleic acid amplification (24, 25) and take advantage of the fact that the viral capsid and envelope (in case it exists) protect the viral genome by externally induced degradation with exogenous nucleases (Fig. 1). They also rely upon the significant difference in the size and density of viral particles compared to most eukaryotic/prokaryotic cells, and they include three main steps—(i) filtration, (ii) nuclease treatment, and (iii) ultracentrifugation—prior to mass parallel sequencing (ideally) using NGS platforms. When these steps are combined, the initial disproportionate high contribution of nonviral nucleic acids in a complex biological sample is minimized, with a concomitant enrichment of the latter in virus-associated nucleic acids. These steps are used after sample homogenization (in the case of solid material) usually by mechanical disintegration, sometimes combined with the use of proteases. An initial clarification of the homogenate by low-speed centrifugation is also used for removal of cellular debris (Fig. 1). For high-volume aqueous samples, an initial concentration may be achieved through the use of

tangential-flow filtration and/or centrifugal filter devices (26). Unfortunately, the identification and study of viral genomes that may exist either as naked episomes or integrated in their host cell DNA cannot be tackled using such purification approaches (27).

The filtration step is usually carried out using disposable syringe-fitted filters (0.45 or 0.22 µm) that exclude the passage of cellular fragments and subcellular organelles (e.g., mitochondria). Despite the fact that these are particularly useful when working with small sample volumes, nonetheless they also block the passage of very large viruses (giant viruses, or giruses: >0.5 µm and genomes over 300 kbp [28]), which may be removed from the filtered flowthrough along with larger cellular material. The following step, nuclease treatment, is usually carried out at 37°C for up to 1 h and calls upon the use of at least one DNase and one RNase. A recent study, however, has suggested that the use of up to 3 DNases (TurboDNase, Benzonase, and exonuclease I) in combination with RNase A significantly minimizes the nonviral sample input (29). To further increase the elimination of contaminating nucleic acids, and at the same time concentrate the viral fraction, the nuclease-treated sample is subsequently subjected to ultracentrifugation steps. Frequently, these include isopycnic CsCl, discontinuous iodixanol, or sucrose gradients (e.g. 70% to 20% run for 2 h at 130,000 × g), but the last two are usually preferred if the infectivity of the viral preparation is to be maintained. This is especially true when working with enveloped viruses, as CsCl affects the stability of the viral surface (18). Furthermore, unless the

density of the envisaged viral particles is known, while the discontinuous gradients allows collection of viral particles from the fraction interface, viral purification using CsCl would force the analysis of a much larger number of fractions.

Additional viral concentration steps via pelleting with polyethylene glycol or ultracentrifugation steps may also be performed. However, these usually entail a significant loss in viral titer, especially when further nuclease treatments are performed (30). At this stage, epifluorescence microscopy may be used to check for the efficiency in the recovery of the viral fraction on either a fixed or live sample using SYBR gold (23), followed by DNA/RNA processing and amplification. These start with the extraction of nucleic acids from the sample, and many options are now commercially available. Be that as it may, frequently Trizol-based extractions are performed, as they cover the simultaneous preparation of DNA as well as RNA templates. Following extraction, biological samples may be further subjected to rRNA removal procedures (e.g., using RiboMinus and Ribo-Zero platforms, commercially available via Life Technologies and Epicentre, respectively) to further avoid the overwhelming domination of rRNA sequences in the NGS sequencing libraries prepared using template RNA (31).

In the great majority of situations, at this stage the amount of viral nucleic acids extracted from most virus-enriched preparations is usually still insufficient for direct sequencing, and even most high-throughput NGS sequencing platforms may not generate a sufficient number of reads for *de novo* assembly of new viral sequences. Therefore, the subsequent analysis of the viral genomes requires an additional step of partial (segments) or complete (full-length) viral genomic amplification. These amplification steps may be carried out using either sequence-specific or sequence-independent methods (32) based on PCR, preceded in the case of RNA viruses by reverse transcription (RT-PCR). Finally, when faced with multiple alternative choices with inherent sets of advantages and limitations (Table 1), based on the information available, one of the first questions to be asked is whether truly novel viruses are expected, or if those putatively present in a sample under analysis belong to previously defined taxonomic clusters. An answer to this question will guide one's choices of methods to use for the detection and characterization of viruses. Ideally, this analysis should not actually depend on any one method alone but should rely on methodological combinations requiring collaborative work from various disciplines (18).

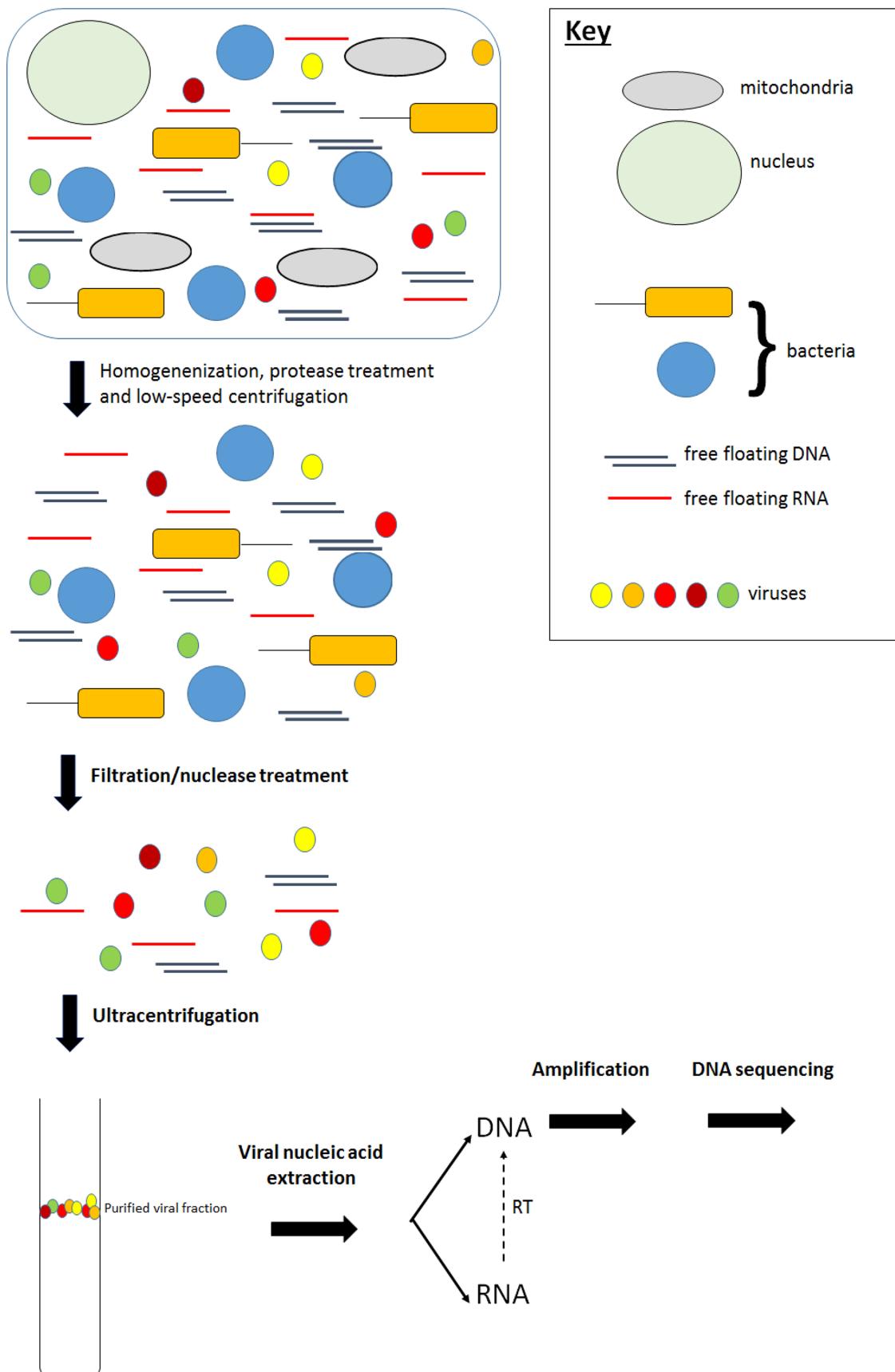
If the detection and amplification of viral nucleotide sequences may be a challenge *per se*, the analysis of the copious data generated in the context of high-throughput sequencing projects involving the analysis of viral sequence data is a challenging task on its own. This is especially true when truly novel sequences are found, as a reference sequence against which sequence assembly may be carried out may not exist. Furthermore, even after filtering out low-quality reads, contig assembly may have to be carried out using strict parameters in order to avoid sequence chimerization (3, 33, 34), i.e., assembly into a single contig different sequences from different viral genomes. Furthermore, one should not forget that viral sequence identification using homology searches with BLASTn and BLASTx and public databases (e.g., GenBank) may fail, as frequently no homologous sequences are available for comparison.

VIRUS DETECTION AND ANALYSIS USING HYBRIDIZATION-BASED METHODS

Detection of Viral Genomes Using Microarrays

The detection of viral nucleic acids using hybridization has supported the discovery of important human viral pathogens, such as hepatitis C virus (HCV). In this case, the amplification of viral sequences was carried out by expansion of a cDNA library of phage clones prepared from an infected chimpanzee and the identification of viral sequences determined by the hybridization of a few of these clones to DNA extracted from human liver tissue (35). Despite being time-consuming and highly labor-intensive, its early success paved the way for other hybridization-based techniques to be used in virology, including those that explore microarrays.

Compared to more conventional hybridization techniques, microarrays have the advantage of simplicity of use and potential for high throughput. Simply put, in a microarray, a high density of probes, usually in the form of oligonucleotides based on a panel of various viral species, are immobilized on a solid surface to which fluorescence-labeled target nucleic acids are hybridized. In the case that a hybridization event occurs, a fluorescent signal is detected and its intensity quantified (36). These signals depend on the stringency of the hybridization process, as well as types of oligonucleotides used. Whether short or long, the identification of known viral genomes or the discovery of new ones may be targeted, due to the array's sensitivity/insensitivity to single-nucleotide mismatches (37, 38), respectively. Some of these arrays have been specifically designed and used for viral genotyping and detection of drug-resistant



mutations in hepatitis B virus (HBV) (39), genotyping and polymorphism detection in SARS-CoV (40), and the identification of a plethora of enteric viruses (41) in stool samples. On the other hand, when oligonucleotides covering a wider range of viral sequences were immobilized on a solid surface, microarrays have also supported the discovery of new human and animal pathogenic viruses (Table 1), including corona- and rhinoviruses (upper respiratory tract) and cardioviruses (gastrointestinal tract) (42, 43).

Despite their potential in allowing simultaneous detection of multiple viruses in a given sample, the analysis of hybridization signals in a microarray is prone to nonspecific binding events, the detection of which is not necessarily straightforward, and calls for the use of specific software tools. Furthermore, for a specific hybridization to occur, probe and target need to match, consequently limiting its use in the detection of very divergent viral species (20). In any case, novel viruses, which may not be explicitly represented in the arrays used, may still be detected in the case that a specific hybridization pattern with phylogenetically related known viral sequences is obtained (18). The two most widely used microarray systems that cover the identification of a large number of viruses (pan-viral) are known as the GreenChip and Virochip (37, 44). They both target conserved regions of a wide range of viruses using 70-mer oligonucleotides, the sequences of which have been regularly updated to reflect our knowledge of the viral taxonomic landscape. A schematic representation of a general virus chip can be found in Fig. 2.

Subtractive Hybridization

As an alternative, methods based on subtractive hybridization usually call for the use of pairs of similar, yet nonidentical, nucleic acid samples designated the

tester (containing viral DNA/RNA) and the driver, from which their common elements may be removed by successive rounds of hybridization and subtraction (45). In the so-called representational-difference analysis (RDA), tester and driver DNA samples (or cDNA when starting from RNA) cut with restriction endonucleases, ligated to artificial adaptors, and hybridized, finally undergo PCR amplification, which occurs exponentially only in tester/driver hybrids (46). In these successive rounds of hybridization/amplification, RDA brings to light disparities in the nucleotide sequences that characterize disease-associated and normal tissue. Unfortunately, RDA relies on the use of highly similar tester and driver DNAs, which severely limits its routine application for the identification of viruses. Nonetheless, human herpesvirus 8 (Kaposi's sarcoma-associated human herpesvirus [Table 1]) ranks among the pathogenic viruses discovered via RDA (47).

All those methods of viral discovery/characterization that depend on a preliminary viral-particle enrichment step are plagued with two major downsides: a reduction in the sensitivity of the technique used and the putative introduction of amplification biases towards the most abundant sequence, frequently of nonviral (host or environmental) origin. To overcome these issues, the presence of viral genomes may still be revealed using subtraction methods that target digital (*in silico*) transcription subtraction (48). This general approach relies on the proviso that if present in a given host, viral genomes are most probably transcriptionally active and will give rise to mRNA. Therefore, instead of targeting viral genomic sequences for amplification, transcription subtraction calls for a global analysis of a samples' transcriptome in combination with mass parallel sequencing techniques (49). This generates a sufficiently large number of sequence reads that maximize the

FIGURE 1 Overview of the viral particle-associated nucleic acid amplification approach.

This strategy combines filtration, nuclease treatment, and ultracentrifugation, takes advantage of the size and density differences observed between most viruses and most eukaryotic/prokaryotic cells, and aims at the enrichment of the viral fraction when dealing with complex samples (characterized by the presence of a large amount of nonviral contaminating nucleic acids). The sample is initially homogenized (in the case of solid material), usually by mechanical disintegration and sometimes combined with the use of proteases, followed by the clarification of the homogenate by low-speed centrifugation, allowing removal of nuclei and cellular debris. The homogenate is subsequently filtrated (0.45 or 0.22 µm) to exclude smaller cellular fragments and subcellular organelles (e.g., mitochondria), followed by nuclease treatment for removal of most (though usually not all) nonencapsidated nucleic acids. The nuclease-treated sample is then subjected to ultracentrifugation/precipitation steps that concentrate and purify viruses from other contaminants that may still be present. Following a nucleic acid extraction, and a possible conversion of viral RNA to cDNA (in the case of viruses with RNA genomes) by reverse transcription, viral sequences may be amplified and sequenced.

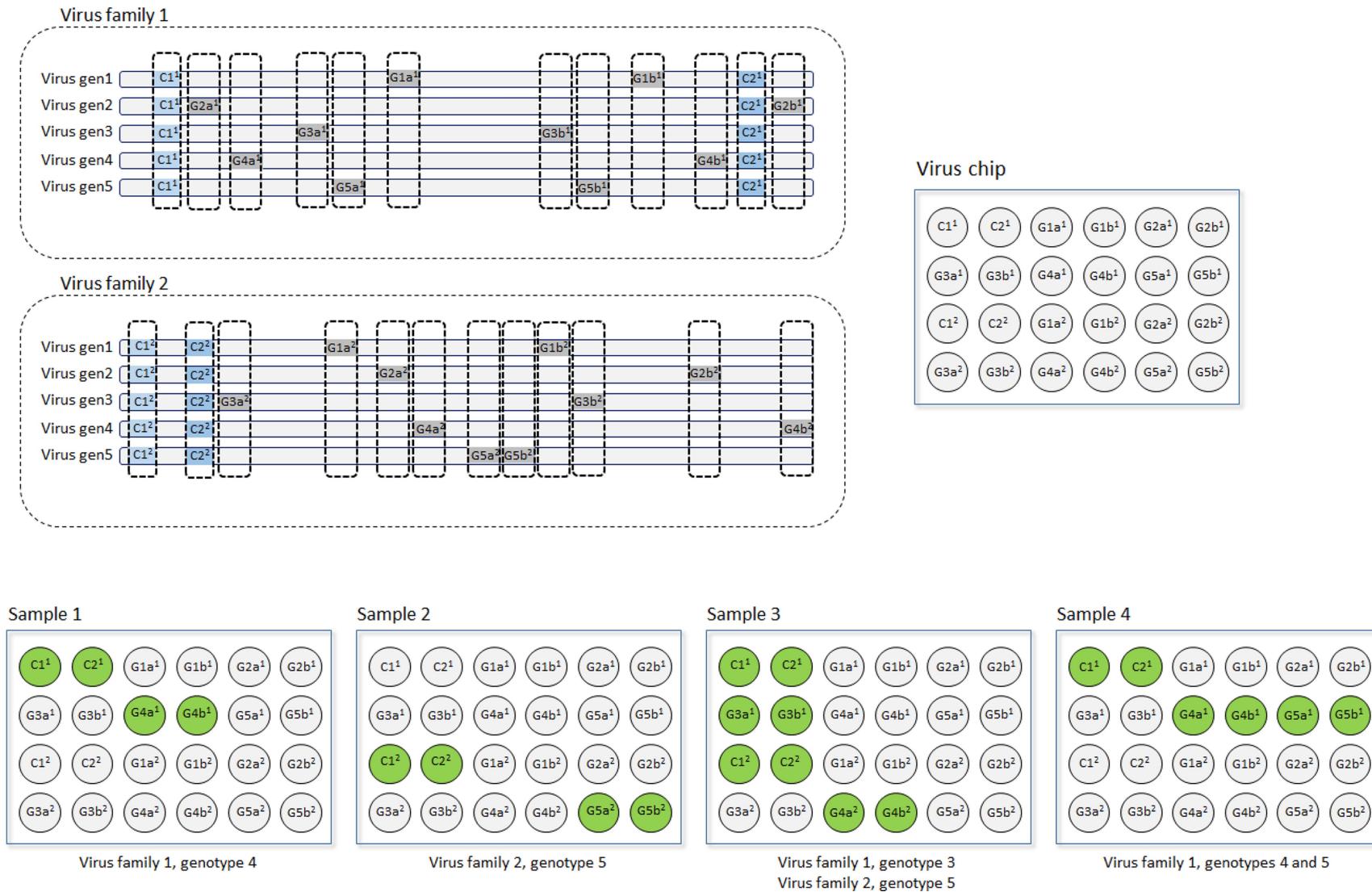


FIGURE 2 Virus chip (microarray). Alignments of multiple virus genomic sequences from different viral families allow the identification of conserved (C) family-specific (1 and 2) regions ($C1^1$, $C2^1$, $C1^2$, and $C2^2$) or genome-specific sequences within each viral group ($G1$ to $-5a^1$, $G1$ to $-5b^1$, $G1$ to $-5a^2$, and $G1$ to $-5b^2$). These sequences are represented in the virus microarray as virus-specific and genotype-specific oligonucleotides and are immobilized on a planar solid surface. The identities of the viruses possibly present in different biological samples (specific viruses, multiple viruses, or even recombinant viruses) may be revealed by the obtained pattern of hybridization signals. Both DNA and RNA viruses may be identified (after reverse transcription of their RNA genomes). Unique viruses that are not represented in the array of immobilized oligonucleotides will not be detected.

probability of detecting viral sequences even when present in the original sample in very low numbers. Sequence assembly is followed by digital subtraction of host-related sequences using database searching and high-fidelity sequences derived from serial analysis of expression libraries, leaving out unmatched virus-related sequence data. The genomes of both DNA (especially polyomaviruses [49–51]) and RNA (52) viruses have been characterized using digital subtraction (Table 1).

PCR-BASED AMPLIFICATION OF VIRAL NUCLEIC ACIDS

Conventional PCR versus Quantitative PCR

Targeted amplification of viral sequences using PCR (or RT-PCR in the case of RNA viruses), has been proven, time and again, as a robust approach for viral genomic detection and analysis. Conventional PCR relies on the use of pairs of oligonucleotides (primers) which target complementary sequences on a DNA template to which they hybridize. These hybridized primers are extended by a thermostable polymerase in successive rounds of thermal cycling, giving rise to the amplification of a DNA segment limited by the primers used. This feature alone precludes the use of conventional PCR for the detection of truly novel viruses, as it calls for *a priori* knowledge (at least partial) of the genomic segments targeted for amplification. Furthermore, especially in the case of viruses that exist as complex quasispecies (such as HIV or HCV), the high genetic variability of the potentially large number of genomic variants present at any given time in an infected individual requires the use of primers with degenerate nucleotide positions (degenerate oligonucleotide-primed PCR). By doing so, it becomes possible to anticipate a high range of genetically diverse or genetically divergent possibilities among the population of targets that may negatively impact the binding of the primers to their targets. Although these degenerate primers are designed to hybridize to conserved genomic regions (identified when the genomes of viral variants are compared in the form of multiple-sequence alignments), their successful use depends on a narrow balance between (i) the necessity to cover a wide range of genetic possibilities and (ii) the use of highly degenerated oligonucleotide mixtures, where only a very small fraction of primers may hybridize to potential targets. The degeneracy level of a primer mixture may be reduced by the introduction of inosine containing nucleotide residues at specific nucleotide positions (53), but it is usually set at a value below the accepted maximum of 128 (18). Furthermore, degeneracy at the 3' end of the

primer should be avoided, as it severely compromises primer extension during PCR. Alternatively, amplifications may also be carried out using either taxon-specific consensus primers (also defined from multiple-sequence alignments) or consensus-degenerate hybrid primers (also known as CODEHOP), where a more degenerate segment at the 3' half follows a constant region (clamp) at the primer 5' core (54, 55). Over time, the use of degenerate primers and conventional PCR protocols has supported the identification of numerous viruses, including human herpesvirus 4 (Epstein-Barr virus) (56), paramyxoviruses (57), corona- and thoroviruses (58), multiple flavi- and phleboviruses (59, 60), and the highly pathogenic Sin Nombre hantavirus (61) (Table 1).

Nowadays, virus detection and characterization based on partial amplification of genomic segments are frequently carried out in a real-time format (also known as quantitative PCR [qPCR]), adding to the sensitivity and specificity of conventional PCR, where the amplification step is followed by amplicon detection using gel electrophoresis. In qPCR, not only are the amplification and detection steps carried out simultaneously (therefore saving time and increasing throughput), but also they occur in a closed tube, effectively sealing the contamination risk by reducing the possibility of accidentally dispersing previously amplified products in the laboratory environment. Moreover, amplicons are frequently (but not necessarily) detected with specific fluorescent probes with defined emission spectra that not only increase the specificity of the intended detections but also support the development of multiplex assays. Consequently, over time, a multitude of reports have described multiplex real-time protocols for the detection of an array of genetically different viruses (for recent examples, see references 62 to 66) that take advantage of the possibility of detecting a few of them (usually no more than 3 or 4) in a single tube.

Viral Genomic Characterization Using Multiregion Hybridization Assays

Although qPCR is frequently used for viral genome detection *per se* (67), it can also be explored for virus genomic characterization when combining type/subtype-specific amplification primers and/or detection probes. Such methods, collectively known as multiregion hybridization assays (MHAs), have been used for epidemiological surveillance of HIV-1, being especially convenient for the characterization of viral strains in epidemiological settings where multiple viral genotypes cocirculate (68–71). Due to the high recombination activity of the retroviral reverse transcriptase, the circula-

tion of genetically distinct viruses frequently gives rise to virions with mosaic genomes. As a result, genetic typing based only on the analysis of a section (or a small number of sections) of the viral genome does not disclose a real image of its genetic makeup, which is best achieved by full-length genomic sequencing followed by phylogenetic analysis. As an alternative, MHAs allow for the simultaneous characterization of multiple segments of the viral genome without requiring any of the latter. In an initial step, 4 to 6 short regions scattered along the HIV-1 genome are amplified in separate first-round PCRs using generic primers. The obtained amplicons are then split into a series of second-round real-time amplification reactions in which they are used as templates and during which the signal emitted by clade-specific fluorescent probes allows the selective identification of specific viral subtypes or subsubtypes.

Detection of Viral Genomes Using Specific Primers and Isothermal Amplification Methods

While conventional PCR and qPCR have been well established as reference tools for detection/amplification of viral nucleic acids, over the last decade many isothermal amplification methods using virus type/subtype-specific primers have also been developed, especially for diagnostic purposes (72). Since the amplification reactions occur at a single temperature, they are usually simpler to implement and require less complex equipment, being especially adapted for laboratory settings with limited resources (67). During the last decade, some of these methods have been used for viral detection and quantification. These include nucleic acid sequence based amplification (NASBA), transcription-mediated amplification (TMA), helicase-dependent amplification (HAD), recombinase polymerase amplification (RPA), and loop-mediated isothermal amplification (LAMP). NASBA and TMA are similar in the sense that they combine the activities of a reverse transcriptase, a DNA polymerase, and an RNA polymerase (usually that of phage T7) for qualitative/quantitative detection of both RNA and DNA viruses (73). HAD differs from RPA in the way primers are annealed to their targets and extended (using a helicase/DNA polymerase combination in the first case and a recombinase/loading factor/DNA polymerase combination in the second) and the temperatures at which amplification reactions occur (67). Regardless of their technical specificities, both have been used for viral analysis and can be combined with a reverse transcription step to study viruses with RNA genomes. HAD has, for example, been used for detection of high-risk papillomaviruses (74) or genetic typing of

herpesviruses 1 and 2 (75), while RPA has been extensively used for detection of chikungunya virus (76), Crimean-Congo hemorrhagic fever virus (77), influenza virus (78), HIV-1 in the form of proviral DNA (79), and Middle East respiratory syndrome coronavirus (80). However, probably the most commonly used isothermal amplification method for the study of viruses is LAMP (81). This technique explores the use of three pairs of primers (external, internal, and loop primers) and gives rise to alternatively oriented repeats of the target sequence. It is extremely efficient and generates large amounts of DNA that, in turn, support the use of a multitude of detection methods, including device-free detection, using DNA-intercalating dyes, fluorescent probes, or cationic polymers (52, 66). Its utilization in the field of virology includes the detection and subtyping of enteroviruses (82), ebolaviruses (83), influenza viruses (84), HIV-1 (85), high-risk papillomaviruses (86), Zika virus (87), and adenoviruses (88).

Despite their unequivocal performance, qPCR and isothermal amplification methods still rely on the hybridization of oligonucleotides to a target sequence; i.e., they are both limited by the hybridization capacity of both amplification primers and detection probes (in the case that these are used for detection of an amplification product) to the nucleic acid templates. Therefore, the possibility exists for a false-negative detection to occur as viral genome sequence divergence increases, regardless of the amplification chemistries used. In other words, there is a limit to the degree of viral genetic variation that these methods can deal with. Frequently, truly novel (unknown) or divergent viruses, as well as those with polymorphisms at the primers' and/or probe's binding sequences (especially RNA viruses, which usually mutate fast), are missed, calling for the development of alternative approaches to meet this challenge. Over the last two decades, multiple systems have been developed for viral genomic sequence detection and characterization that do not entail *a priori* knowledge of their nature and composition. For this reason, they have come to be collectively known as sequence-independent genome amplification methods (32).

Sequence-Independent Viral Genome Amplification

Based on theory, sequence-independent genome amplification methods relieve a few of the constraints previously identified when considering doing viral genetic detection using microarrays (failure to detect bona fide specific hybridizations between a target nucleic acid

preparation and the microarray itself), genome subtraction methods such as RDA (need to use pairs of almost identical target/driver nucleic acid samples), or conventional PCR (relaxing the need for *a priori* knowledge of the sequences to be identified [Table 1]). They are usually easy to implement, and some are already available in commercial formats, especially those making use of the highly efficient phi29 DNA polymerase.

Sequence-independent single-primer amplification (SISPA) (Fig. 3) was one of the first sequence-independent genome amplification methods to be successfully employed for the detection of viruses that defy isolation in cell culture (89), including enteric viruses such as astro-, noro-, and rotaviruses (90–92) (Table 1). Although these viruses all have RNA genomes, SISPA may also be used to cover the detection of viruses with DNA viruses. In its original format, SISPA entails the preparation of cDNA (in the form of double-stranded DNA [dsDNA]), its digestion with frequently cutting restriction endonucleases (e.g., HinP1I or MseI) and ligation of artificial linkers (also known as adaptors), followed by cycles of PCR amplification using a single primer, specifically targeting complementary sequences in the artificial adaptors used. A technical variation of SISPA that came to be known as VIDISCA (virus discovery-cDNA-AFLP) was successfully used in 2004 for the characterization of a novel human coronavirus (NL63) (93) and, more recently, for the identification of human parechovirus type 1, previously isolated from a stool sample of an individual with acute enteritis (94). As opposed to SISPA, VIDISCA (Fig. 3) explores the use of two different adaptors for the preparation of a cDNA adaptor library, as well as the possibility for selective DNA amplification using two primers, extended at their 3' ends by one of the 4 possible nucleotides, generating 16 different possible combinations (32). However, SISPA and VIDISCA are not impervious to biases and may introduce distortions in viral sequence coverage. Certain regions of the viral genome may be absent, while others are covered many times from the collection of sequence reads. Furthermore, SISPA may also shift the amplification results towards the dominant genome in the assembly (95–97).

Due to the nonrestrictive nature of the desired DNA amplifications, sequence-independent amplification methods are negatively affected by spurious targeting of contaminating nonviral sequences by the primers used. To overcome this potential problem, selective enrichment of the viral fraction by combining filtration, nuclease treatments, and ultracentrifugation rounds for discriminatory selection of particle-associated

nucleic acids may be coupled (in analysis of viral RNA) with reverse transcription carried out with nonribosomal hexanucleotides that fail to hybridize to rRNA the analyzed sample may still contain (98, 99).

Viral sequences may alternatively be targeted for amplification with random hexamers and protocols that, unlike SISPA and VIDISCA, skip the ligation of artificial adaptors to dsDNA. These rPCRs combine two amplification steps (100). Nucleic acids are first targeted with chimeric primers that, starting from the 5' end, consist of a defined sequence followed by 6 to 8 degenerate haptamer sequences. In the second round, the use of a specific primer complementary to the defined sequence found in the primer used in the first round allows the exponential amplification of DNA. Throughout the years, rPCR has been extensively used for analysis of a profusion of both human RNA (101, 102) and DNA (103–105) viruses. Still today, rPCR is one of the most frequently used approaches for identifying new viruses, especially when used as a viral genome preamplification step (DNA enrichment) in protocols where it is combined with NGS (106–110).

Methods such as multiply primed rolling-circle amplification (MPRCA) (111) and multiple-displacement amplification (MDA) (112) explore the polymerization/displacement activities of phi29 DNA polymerase that, when used in combination with randomly hybridized primers, support unbiased whole-genome amplification of both circular and linear genomes, respectively (Fig. 4). Unlike those amplification methods that make use of conventional thermostable *Taq*-like enzymes with relatively high polymerization-associated errors, and that allow only specific regions of the viral genome to be analyzed (incomplete genomic coverage), MPRCA and MDA explore the high processivity and fidelity of the phi29 DNA polymerase. As a result, these methods support highly uniform isothermal (up to 10⁶-fold) amplifications of targets with average lengths of >10 kb (113, 114). Furthermore, these techniques may be used starting either from high-quality DNA preparations or directly from biological samples, including whole blood, plasma, serum, or tissue culture cells (112), and give rise to dsDNA molecules that can be further downward processed. Although the general experimental procedures involved are technically more demanding than other more conventional genome amplification methods, MPRCA and MDA have been previously used for the genetic characterization of several viral agents, especially those with circular genomes, including papillomaviruses (115), polyomaviruses (116, 117), anelloviruses (118), and hepadnaviruses (119).

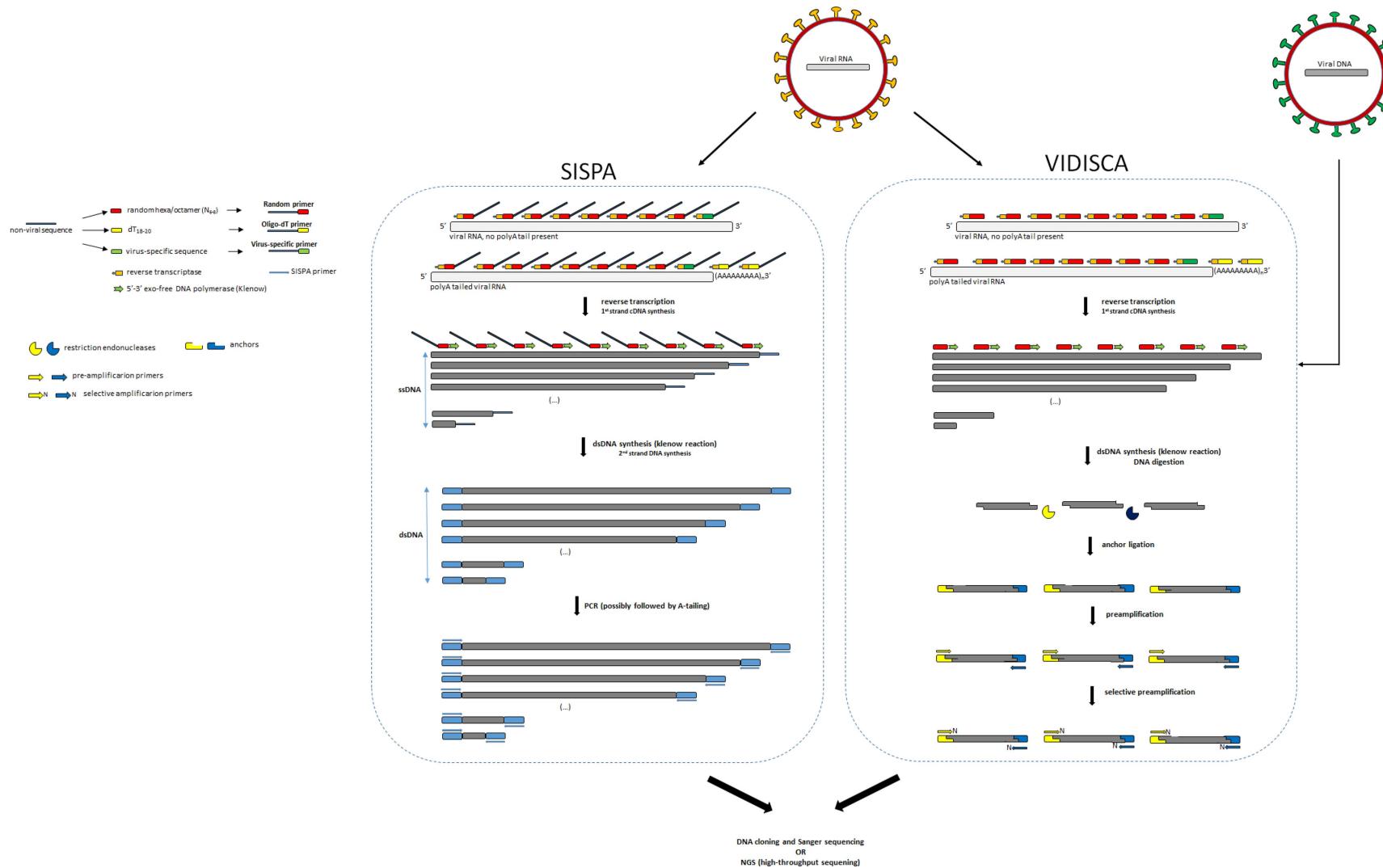


FIGURE 3 Outline of the SISPA and VIDISCA methods. Viral genomic RNA is converted to cDNA using either tagged (SISPA) or nontagged (VIDISCA) random (6 to 8 nucleotides), poly(A), or virus-specific primers. Second-strand DNA is then synthesized using an 5'-to-3' exonuclease-free DNA polymerase (usually the Klenow fragment of *E. coli* DNA polymerase I), in the presence of random tagged (SISPA) or untagged random hexa/octamers. dsDNA is amplified by PCR using a single tag-specific primer (SISPA) or digested with restriction endonucleases, followed by the ligation of anchors to the ends of the cleaved dsDNA (VIDISCA). In SISPA, amplified DNA fragments may be processed for either blunt or TA cloning in a vector (followed by Sanger sequencing) or directly sequenced using one of the possible NGS platforms. In the case of VIDISCA, anchor-ligated DNA follows two steps of amplification before it may be used for cloning/Sanger sequencing or NGS analysis. Viruses with DNA genomes may also be readily analyzed by SISPA or VIDISCA. Adapted from references 93 and 107.

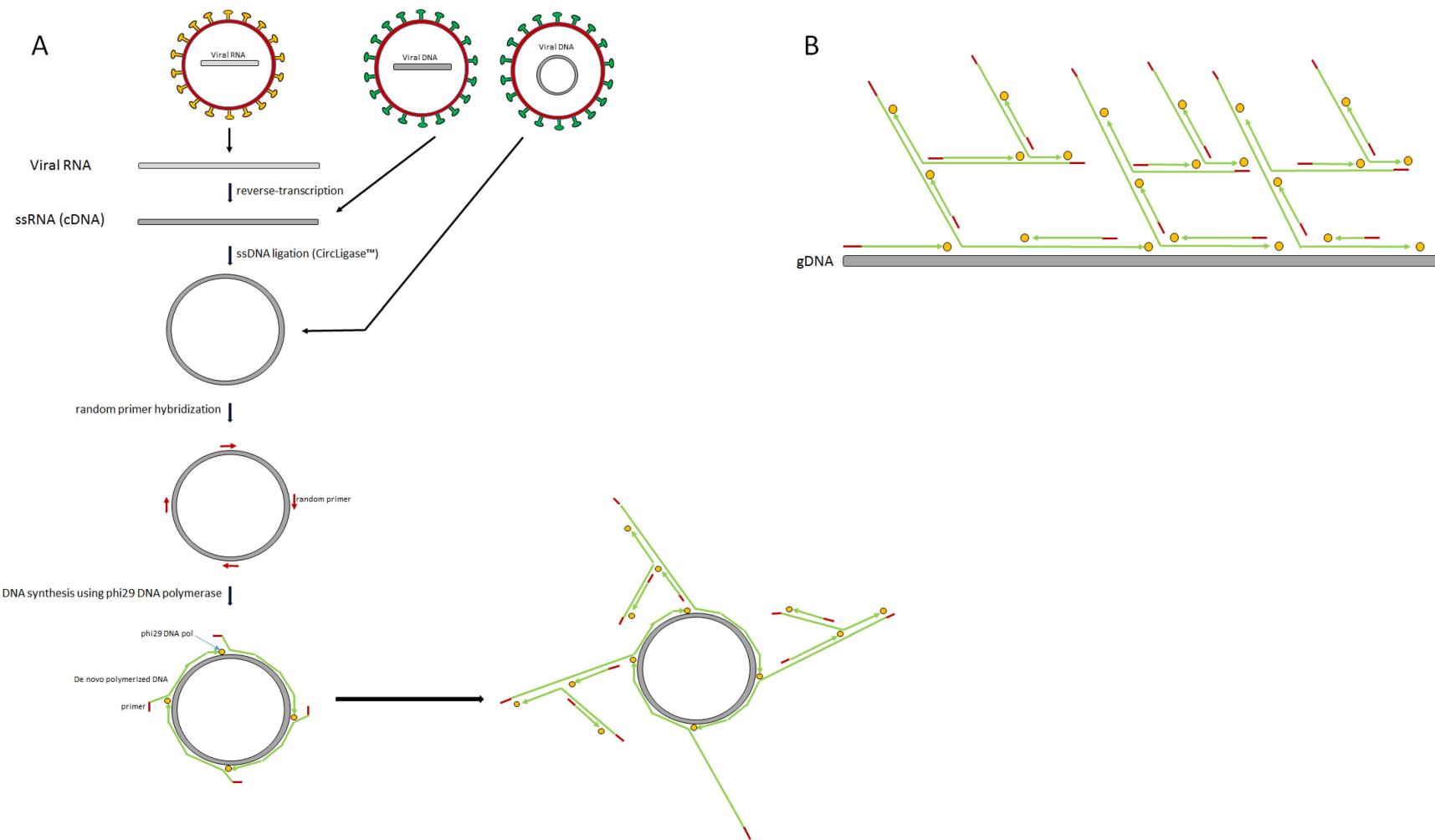


FIGURE 4 MPRCA (also known as rolling-circle amplification [RCA]) (**A**) and MDA (**B**) are isothermal DNA amplification methods that explore the polymerization as well as the displacement activities of bacteriophage phi29 DNA polymerase. When used in combination with randomly hybridized primers, both methods support unbiased whole-genome amplification. Viral genomes may be directly amplified (DNA viruses) or first converted to ssDNA (RNA viruses) using reverse transcriptase. Both linear and circular molecules may be amplified by MDA and MPRCA, respectively. Linear DNA molecules (viral ssDNA genomes or cDNA) may also be converted to circular molecules (substrates for MPRCA) using an ssDNA ligase.

Virus Surveillance, Identification, and Genetic Characterization Using Mass Spectrometry

While virus detection/characterization approaches based on the amplification of viral genomes are sensitive, their use is challenging because viral genomic sequences, especially those of RNA viruses, rapidly change in time and space in a process known as genetic drift. In the case of viruses with segmented genomes, this problem is further complicated by the possibility of exchange of segment between viral genomes in a coinfecting cell, which may give rise to viral genomes with unanticipated genetic makeups. These phenomena may compromise the ability of the PCR primers used to bind to their targets on the viral genomes. Furthermore, molecular methods based on nucleic acid amplification may be time-consuming, as they entail extraction/purification and amplification/detection steps that, in a conventional format (PCR/RT-PCR followed by gel electrophoresis), may take up to 8 to 10 h. This time frame will be considerably extended if an unambiguous identification of the amplified DNA depends on subsequent nucleic acid sequencing (sometimes preceded by DNA cloning) and phylogenetic analysis.

Though initially limited to chemical sciences, technical improvements in mass spectrometry, including the development of electron spray ionization and matrix-assisted laser desorption ionization, have brought this technique to the domain of biological sciences. The accuracy of a mass spectrometer is very high, results are obtained within minutes, and the time for analysis is essentially limited to sample processing. Until recently, the ability to identify biomolecules based on the mass/charge ratio of the ions generated after sample processing/ionization has been mostly focused on the analysis of proteins, identified by a characteristic peptide mass fingerprint. Mass spectrometry has been extensively used in bacteriology, in which the analysis of the abundant ribosomal proteins usually allows the identification of bacteria in clinical/environmental samples or bacterial isolates. However, the method can also be used to detect toxins and spores (e.g., for detection of biological warfare agents), enzymatic reaction products (e.g., antibiotic degradation products generated by the presence of β -lactamases), or nucleic acids (mostly DNA; for a review, see reference 120).

As far as viruses are concerned, mass spectrometry supports identification, typing/subtyping, and even mutation detection using two approaches: the identification of so-called signature peptides derived from the proteolytic digestion of viral antigens (in a process currently known as viral proteotyping [121, 122] or the analysis

of genomic sequences amplified by PCR/RT-PCR. Both of these approaches have been used for the detection, genetic characterization, and epidemiological surveillance of influenza viruses (123) with very high sensitivity and throughput (85, 86), for the detection of enteric viruses (124), arboviruses (125), papillomaviruses (126), polioviruses (127), and respiratory viruses (128), for genotyping of hepatitis viruses (129), and for detection of mutations in the genomes of herpesviruses (130).

Combining Nucleic Acid Amplification and Its Detection Using Nanotechnologies for Detection of Multiple Viruses by Increased Multiplexing

Although NGS does have the potential to identify viral genomes without any prior knowledge of their sequence (see following sections), it is not the most practical approach for direct viral screening in large numbers of samples. As an alternative, a plethora of possible combinations of conventional PCR and qPCR that explore the use of either conserved or degenerate primers followed by DNA sequencing have been extensively used for virus identification, genomic characterization (including mutation detection), and epidemiological surveillance.

While the throughput of qPCR may be increased by multiplexing, if the number of viruses to be detected rises above the resolution limits of each protocol (usually no more than 3 or 4), different qPCRs need to be set up independently. However, a considerably higher multiplexing level may be achieved, for example, if nucleic acid amplification is combined with a detection system based on micro- or nanoparticles. These include polystyrene microspheres with covalently linked tags, magnetic nanoparticles covered with capture probes that bind to biotinylated viral amplification products, or gold nanoprobe as the endpoint of a method that allows both virus detection and phenotyping. In all cases, the viral genomic regions of interest are first amplified with a multiplex PCR (or RT-PCR) protocol.

In the so-called Luminex xTAG technology (proprietary name belonging to the Luminex Corporation), the amplification step is followed by a linear asymmetric amplification of viral sequences by target-specific primer extension (TSPE), during which the amplified molecules are labeled with a biotinylated nucleotide (for a summary of the process, see reference 67). The presence of one of multiple 24-nucleotide-long tag sequences, located at the 5' ends of the primers used for TSPE, allows the subsequent capture of the amplified products by

complementary anti-tags, which are covalently coupled to MagPlex-tag microspheres. Therefore, the hybridization of the amplified/labeled products with a particular microsphere sets the stage for their identification using a fluorescent molecule to detect the incorporated biotin. This method offers a high degree of customization and may be expanded to the detection of complex array of new and emerging viruses, and its potential has been demonstrated in the detection of the highly pathogenic Hendra and Nipah paramyxoviruses (131) and multiple viruses associated with central nervous system infections, including herpesviruses, polyomaviruses, and enteroviruses (132).

An alternative method has been used to simultaneously detect HBV and HCV as well as HIV-1 (133). In this case, silica-modified magnetic nanoparticles are used for extraction of viral nucleic acids (both RNA and DNA) that then act as templates for the amplification of specific viral sequences by multiplex PCR (or RT-PCR), again labeling DNA with a biotinylated nucleotide. Finally, the amplicons are selectively captured on the surface of amino acid-modified carboxyl-coated magnetic nanospheres, and the addition of streptavidin-modified alkaline phosphatase (which binds to biotin) supports a detection system based on chemiluminescence.

Finally, a recently described new method has been used for detection and phenotyping of ebolaviruses (134). This method (i) enables the detection of very small amounts of material using hairpin oligonucleotides that polymerize into a long nicked dsDNA molecule after the addition of a single-stranded DNA (ssDNA) initiator (amplified by RT-PCR), (ii) takes advantage of the physical properties of gold-labeled nanoparticles (AuNPs), and (iii) explores the performance of an enzyme-free process called hybridization chain reaction (135). This method is characterized by a remarkable signal amplification capacity, allowing results to be inspected visually on a microtiter plate, as a result of the shift in the surface plasmon band of the gold nanoprobes. Furthermore, as different initiator viral sequences may be labeled by hybridization with one of many possible oligonucleotides, single AuNPs may be rapidly reprogrammed for each virus subtype or group of viruses.

ANALYSIS OF VIRAL AMPLIFICATION PRODUCTS: SINGLE MOLECULE VERSUS POPULATION SEQUENCING

Methodological decisions as to what sequencing approach should be followed for the characterization of

viral sequences may be determined by the complexity of the viral population under analysis, i.e., whether it is amenable to single molecule vs. population analysis via Sanger sequencing. A more thorough assessment of the diversity of viral genomes may be achieved not by direct sequencing of PCR products (population sequencing) but via the characterization of viral genomes based on the analysis of multiple individual clones, obtained by cloning the amplified viral sequences in a cloning vector (see below).

As much as unbiased massive parallel sequencing may be regarded as the most powerful method to identify truly novel potential pathogens or reveal the true genetic structure of a given viral population (Table 1), Sanger sequencing may still be the method of choice, especially if viral characterization builds from the analyses of individual PCR products. However, the specificity of any amplification experiment is determined by the stringency of the amplification conditions used (itself impacted by the extent of primer-template matches), and more frequently than not, especially when degenerate primers are used, multiple amplicons are obtained. In this case, securing nonambiguous sequence data sets may depend on the preliminary purification of the desired amplicons, or even their stabilization in a vector, usually in the form of a plasmid. However, as cloning experiments make use of living hosts (frequently *Escherichia coli*), an unexpected number of possible clones may be missed, especially in the case that they impart any kind of stress or toxicity to the cloning host, therefore distorting the apparent genetic composition of the viral population. Although this may be partially overcome by analyzing multiple clones from each cloning experiment, the workload involved frequently pushes a decision towards direct sequencing of PCR products, known as population sequencing. While this brings to light the features of the most frequently amplified viral sequence, minor variants associated with drug resistance mutations fail to be detected in such a way if their frequency falls below 20%, especially in the case of those viruses (e.g., HIV) that are characterized by potentially complex quasispecies (136). In other words, a thorough investigation of the true composition of the viral quasispecies becomes relevant when we take into account the fact that its diffuse nature, equivalent to a cloud of genetic possibilities generated during viral replication, provides fertile ground from which highly fit viral genomes may be selected from via natural/artificial selection. In this regard, certain viral features may not be revealed if only a consensus sequence obtained via Sanger (population) sequencing becomes available (137).

VIRAL QUASISPECIES CHARACTERIZATION AND DETECTION OF MINOR VIRAL GENOMIC VARIANTS

As shown above, in contrast to the level of genetic resolution brought about by deep sequencing via NGS, Sanger sequencing is time-consuming and very labor-intensive, involves the analysis of multiple individual clones, and may suffer from biases that distort the true representation of sequences in a given sample. This may be the consequence of preferential amplification of some sequences over others or cloning bottlenecks (Table 1).

The ensemble of techniques that allow a thorough inspection of the composition of a viral genome assembly via deep sequencing allows detection of low-frequency variants and assessment of their relative frequencies (138). As stated above, while Sanger sequencing is usually incapable of detecting genomic variants underrepresented in the viral quasispecies (<20%), ultradeep sequencing based on NGS may detect viral variants at levels as low as 0.1% (139, 140). Through time, experience has shown that when virus replication cannot be halted, the minor variants are continuously generated. This results from the combination of high virus mutation and turnover rates, and the frequency of these minor variants may increase if they present higher fitness than the wild-type viruses under selective pressure (for example, antiretroviral treatment), to the point of compromising the effectiveness of therapeutic interventions and vaccine design (141, 142).

The resolution power of deep sequencing and its practical implications in virology have already been recognized for over a decade. Initially, deep sequencing was explored with the purpose of analyzing HIV drug resistance-associated mutations, prediction of coreceptor usage, and the early detection of emergence of CXCR4-tropic viruses (143–146). The high resolution of deep sequencing clearly contrasts with that of other experimental approaches (including line probe assay-type methods [147, 148]) that are limited in their potential to detect minor variants, including, for example, those associated with antiviral resistance. Indeed, in the early stages of the infection process, and before therapy is initiated, these drug-resistant viruses are usually part of the collection of circulating minor sequence variants and cannot easily be detected by genotyping methods based on population sequencing (149). In these situations, deep sequencing using NGS can support the reconstruction of the true structure of the viral population to the point where it becomes possible to distinguish random mutants (the result of the viral error-prone replication) from those variants adapted to therapeutic pressure. Its impact

on clinical practice is evident, as it bestows medical doctors with rigorous information that allows them to best predict treatment outcomes (150). Analyses of viral diversity using deep sequencing have also led to the identification of superinfection events in HIV-infected individuals (151) and the elucidation of how HIV evolves in different human compartments (152).

Many aspects related to the interaction of many other viruses with their hosts (e.g., HBV, HCV, and dengue viruses [DENV]) have also been addressed through high-throughput genomic sequencing projects. These have included the study of how HBV evolves resistance to reverse transcriptase inhibitors (153), the recognition of immune-mediated bottlenecks associated with transmission of HCV and DENV (154–156), the identification of nonuniform patterns of genetic variation across the genome and phylogenetic lineages of DENV-2 (157), the recognition of host genetic factors in controlling the breadth of within-host genetic diversity in DENV- and West Nile virus-infected individuals (155, 158), and the identification of noncoding small RNAs that correlate with hepatocellular carcinoma in HCV-infected patients (159). Since the implementation of new anti-HCV therapeutic regimens using inhibitors of the polymerase (NS5B), protease (NS2A/3), or a multifunctional regulator of cellular pathways and virus replication (NS5A), the hindsight gained from drug resistance mutation detection/analysis and quasispecies characterization achieved with HIV has been translated to HCV. In this regard, tools that allow deep-sequencing analysis of viral population structure and detection of minor variants give invaluable insights into the dynamics of antiviral resistance acquisition by HCV (160).

Since viruses with RNA genomes account for the majority of those that infect eukaryotic cells (161), not surprisingly, they also comprise the majority of those analyzed using the technical approaches described above. However, human ssDNA and dsDNA viruses are also common, and all these technologies may equally be put to use in order to explore their identities, genomic features, and population dynamics (Table 1). Despite the fact that populations of DNA viruses are normally less complex than those of RNA viruses, two of the DNA viruses that pose health threats, especially to the health of immunosuppressed individuals, such as human herpesvirus 5 (HHV-5; also known as human cytomegalovirus) and human papillomaviruses (HPV) have also been recently investigated using NGS-generated deep-sequencing data. One study has shown high HHV-5 genetic variability in lung transplant recipients (162), and another has revealed that in transplant patients, mixed infections with multiple

HHV-5 strains appear to be associated with more severe disease outcomes (163). Curiously, the same does not seem to hold true in the case of newborns, for whom ultradeep sequencing of urine samples rarely disclosed the presence of multiple viral genotypes (164). Furthermore, one single study based on the analysis of amplifiers from HPV consensus PCR of skin samples has (i) recently shown that deep sequencing appears to be very reliable for detection of cutaneous viruses and (ii) identified more than 390 HPV types (types are defined as assemblages differing by $\geq 10\%$ in their respective L1 coding sequences [165]), demonstrating that cutaneous papillomaviruses are extremely diverse and far more numerous than previously known (166). Finally, a new method based on multiplex PCR and deep sequencing has been validated for (i) the characterization of the E6/E7 region of high-risk papillomaviruses and (ii) the identification of intra-type variants (differing at between 2 and 10% of their sequences; also known as subtypes or lineages [167]). Importantly, this assay may be used in combination with the detection of other markers (including E6/E7 mRNA and p16INK4/cyclin-dependent kinase inhibitor expression levels and viral genomic methylation patterns) to define causative roles of HPV in cancer (168).

Finally, deep-sequencing methods have also been used to explore the dynamics of influenza viruses in nature, including the detection of genome segment swapping (genetic reassortment) between human and nonhuman viruses, investigation of antigenic stability associated with the virus-encoded hemagglutinin and neuramidase, detection and analysis of transmission of neuramidase (e.g., oseltamivir) mutants, identification of virulence signatures, and estimation of viral pandemic potential (169–175). As a result of the combination of deep sequencing with robust methods of phylogenetic reconstruction, the origin of the recent pandemics (e.g., the 2009 outbreak caused by H1N1 viruses) has been clarified (176) and the potential contribution of host factors in the expansion of the highly pathogenic avian H5N1 virus inferred (177).

CONCLUDING REMARKS

During the last decade, the study of viruses has been boosted by the development of a series of technical approaches that rely on a refined isolation and analysis of viruses and viral genomes (more frequently than viral proteins), with different levels of performance and limitations. Multiple protocols are now available for the purification of viral particles, and the amplification, high-throughput sequencing, and bioinformatic analyses

of viral genomes. In multiple possible combinations, these techniques grant us the possibility to detect unexpected new viruses in clinical samples, investigate viral diversity, assess the stability of their genomes (including the detection of recombinant and reassorted genomes) and the complexity of their populations, disclose patterns of within- and between-host viral evolution, identify low-abundance mutations associated with pathogenesis, host adaptation, and drug resistance, even characterize viral communities in specific human body niches, in both healthy and disease-conditioned individuals, and ensure vaccine safety (137). Indeed, the intrinsic genetic instability of many viruses may translate into the accumulation of virulent strains during the manufacturing of live attenuated vaccines, and the techniques that allow the characterization of viral genomes with high resolution have been used for adequate vaccine quality control and to ensure vaccine safety (137). Although the characterization of viral genomes has been extensively used for the analysis of stocks of oral polio vaccine (178), it has already been used for the examination of other live attenuated viral vaccines (e.g., yellow fever, rubella, and rotavirus), including one against DNA viruses (varicella-zoster virus). In this context, deep sequencing not only allows detection of, e.g., neurovirulence-associated mutations that may escape identification with other methods, but also warrants the identification of contaminating viruses from the vaccine producer cells (179).

Last, though not least, when the techniques mentioned above are used to identify new viruses as possible etiological agents of disease, detection of viral genomes in a given setting does not stand as proof of causation of disease, which is defined through the application of Koch's postulates or other sets of postulates. As a matter of fact, the molecular era has set new challenges for definition of proof of causation, especially when viruses are concerned. Just as an example, over the last years many new viruses for which no pathology has been associated have been found using metagenomics approaches. This frequently contrasts with the "traditional" way of searching for etiological agents of disease. However, as in the past, Koch's postulates have been regularly modified, and the most recent modification (metagenomics Koch's postulates) focuses on the use of genetic information to segregate healthy versus diseased metagenomes (3, 180). Alternatively, a new integrative approach to causation has also been proposed which is less dependent on the types of data available (e.g., antibody detection, PCR amplification, *in situ* hybridization, or genome sequencing) but focuses on the strength

of associations inferred from the data, defining the identified causal relationships as possible, probable, and confirmed (181). In many situations, it is the epidemiologic observation linked to the data generated by these viral detection and sequencing methods that provides evidence of causation.

As the technologies that deal with the detection and characterization of nucleic acids improve, so will our knowledge of virus diversity, viral infectious disease epidemiology, and how they impact human health. However, the implementation of high-throughput technologies for virus detection and characterization via genome sequencing has generated a secondary obstacle. Already today, it is not so much the acquisition of data that strictly defines the tempo of scientific progress in this field but the challenges brought about by the rapid accumulation of overwhelming amounts of sequence data. The available algorithms for sequence analysis, including sequence read cleanup, assembly, and amplicon analysis, present limitations and sometimes require the development of in-house bioinformatics tools to tackle specific problems (182) and the acquisition of bioinformatics expertise.

CURATED COLLECTION

[Click here to read other articles in this collection.](#)

REFERENCES

1. Cruz-Rivera M, Forbi JC, Yamasaki LHT, Vazquez-Chacon CA, Martinez-Guarneros A, Carpio-Pedroza JC, Escobar-Gutiérrez A, Ruiz-Tovar K, Fonseca-Coronado S, Vaughan G. 2013. Molecular epidemiology of viral diseases in the era of next generation sequencing. *J Clin Virol* 57:378–380. <http://dx.doi.org/10.1016/j.jcv.2013.04.021>.
2. Suttle CA. 2007. Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol* 5:801–812. <http://dx.doi.org/10.1038/nrmicro1750>.
3. Mokili JL, Rohwer F, Dutilh BE. 2012. Metagenomics and future perspectives in virus discovery. *Curr Opin Virol* 2:63–77. <http://dx.doi.org/10.1016/j.coviro.2011.12.004>.
4. Roux S, Hallam SJ, Woyke T, Sullivan MB. 2015. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *Elife* 4:1–20. <http://dx.doi.org/10.7554/elife.08490>.
5. Domingo E, Sheldon J, Perales C. 2012. Viral quasispecies evolution. *Microbiol Mol Biol Rev* 76:159–216. <http://dx.doi.org/10.1128/MMBR.05023-11>.
6. Ramachandran S, Xia GL, Ganova-Raeva LM, Nainan OV, Khudyakov Y. 2008. End-point limiting-dilution real-time PCR assay for evaluation of hepatitis C virus quasispecies in serum: performance under optimal and suboptimal conditions. *J Virol Methods* 151:217–224. <http://dx.doi.org/10.1016/j.jviromet.2008.05.005>.
7. Henn MR, Boutwell CL, Charlebois P, Lennon NJ, Power KA, Macalalad AR, Berlin AM, Malboeuf CM, Ryan EM, Gnerre S, Zody MC, Erlich RL, Green LM, Berical A, Wang Y, Casali M, Streeck H, Bloom AK, Dudek T, Tully D, Newman R, Axten KL, Gladden AD, Battis L, Kemper M, Zeng Q, Shea TP, Guija S, Zedlack C, Gasser O, Brander C, Hess C, Günthard HF, Brumme ZL, Brumme CJ, Bazner S, Rychert J, Tinsley JP, Mayer KH, Rosenberg E, Pereyra F, Levin JZ, Young SK, Jessen H, Altfeld M, Birren BW, Walker BD, Allen TM. 2012. Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. *PLoS Pathog* 8:e1002529. <http://dx.doi.org/10.1371/journal.ppat.1002529>.
8. Ramachandran S, Campo DS, Dimitrova ZE, Xia GL, Purdy MA, Khudyakov YE. 2011. Temporal variations in the hepatitis C virus intrahost population during chronic infection. *J Virol* 85:6369–6380. <http://dx.doi.org/10.1128/JVI.02204-10>.
9. Woolhouse M, Scott F, Hudson Z, Howey R, Chase-Topping M. 2012. Human viruses: discovery and emergence. *Philos Trans R Soc Lond B Biol Sci* 367:2864–2871. <http://dx.doi.org/10.1098/rstb.2011.0354>.
10. Woolhouse MEJ, Howey R, Gaunt E, Reilly L, Chase-Topping M, Savill N. 2008. Temporal trends in the discovery of human viruses. *Proc Biol Sci* 275:2111–2115. <http://dx.doi.org/10.1098/rspb.2008.0294>.
11. Itoh Y, Shinya K, Kiso M, Watanabe T, Sakoda Y, Hatta M, Muramoto Y, Tamura D, Sakai-Tagawa Y, Noda T, Sakabe S, Imai M, Hatta Y, Watanabe S, Li C, Yamada S, Fujii K, Murakami S, Imai H, Kakugawa S, Ito M, Takano R, Iwatsuki-Horimoto K, Shimojima M, Horimoto T, Goto H, Takahashi K, Makino A, Ishigaki H, Nakayama M, Okamatsu M, Takahashi K, Warshauer D, Shultz PA, Saito R, Suzuki H, Furuta Y, Yamashita M, Mitamura K, Nakano K, Nakamura M, Brockman-Schneider R, Mitamura H, Yamazaki M, Sugaya N, Suresh M, Ozawa M, Neumann G, Gern J, Kida H, Ogasawara K, Kawaoka Y. 2009. In vitro and in vivo characterization of new swine-origin H1N1 influenza viruses. *Nature* 460:1021–1025. <http://dx.doi.org/10.1038/nature08260>.
12. Zhu H, Lam TTY, Smith DK, Guan Y. 2016. Emergence and development of H7N9 influenza viruses in China. *Curr Opin Virol* 16:106–113. <http://dx.doi.org/10.1016/j.coviro.2016.01.020>.
13. Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, Liu W, Bi Y, Gao GF. 2016. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol* 24:490–502. <http://dx.doi.org/10.1016/j.tim.2016.03.003>.
14. Sayed IM, Vercooter A-S, Abdelwahab SF, Vercauteren K, Meuleman P. 2015. Is HEV an emerging problem in industrialized countries? *Hepatology* 62:1883–1892.
15. Kruger DH, Figueiredo LTM, Song JW, Klempa B. 2015. Hantaviruses—globally emerging pathogens. *J Clin Virol* 64:128–136. <http://dx.doi.org/10.1016/j.jcv.2014.08.033>.
16. Alexander KA, Sanderson CE, Marathe M, Lewis BL, Rivers CM, Shaman J, Drake JM, Lofgren E, Dato VM, Eisenberg MC, Eubank S. 2015. What factors might have led to the emergence of Ebola in West Africa? *PLoS Negl Trop Dis* 9:e0003652. <http://dx.doi.org/10.1371/journal.pntd.0003652>.
17. Faye O, Freire CCM, Iamarino A, Faye O, de Oliveira JVC, Diallo M, Zanotto PMA, Sall AA. 2014. Molecular evolution of Zika virus during its emergence in the 20(th) century. *PLoS Negl Trop Dis* 8:e2636. <http://dx.doi.org/10.1371/journal.pntd.0002636>.
18. Sridhar S, To KK, Chan JFW, Lau SKP, Woo PCY, Yuen KY. 2015. A systematic approach to novel virus discovery in emerging infectious disease outbreaks. *J Mol Diagn* 17:230–241. <http://dx.doi.org/10.1016/j.jmoldx.2014.12.002>.
19. Reller LB, Weinstein MP, Editors S, Storch GA. 2000. Special section: medical microbiology—diagnostic virology. *Clin Infect Dis* 31:739–751.
20. Bexfield N, Kellam P. 2011. Metagenomics and the molecular identification of novel viruses. *Vet J* 190:191–198. <http://dx.doi.org/10.1016/j.tvjl.2010.10.014>.
21. Datta S, Budhauliya R, Das B, Chatterjee S, Vanlalhmuaka, Veer V. 2015. Next-generation sequencing in clinical virology: discovery of new viruses. *World J Virol* 4:265–276. <http://dx.doi.org/10.5501/wjv.v4.i3.265>.
22. Petrosino JF, Highlander S, Luna RA, Gibbs RA, Versalovic J. 2009. Metagenomic pyrosequencing and microbial identification. *Clin Chem* 55:856–866. <http://dx.doi.org/10.1373/clinchem.2008.107565>.
23. Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F. 2009. Laboratory procedures to generate viral metagenomes. *Nat Protoc* 4:470–483. <http://dx.doi.org/10.1038/nprot.2009.10>.

24. Stang A, Korn K, Wildner O, Überla K. 2005. Characterization of virus isolates by particle-associated nucleic acid PCR. *J Clin Microbiol* 43:716–720. <http://dx.doi.org/10.1128/JCM.43.2.716-720.2005>.
25. Breitbart M, Rohwer F. 2005. Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. *Biotechniques* 39:729–736. <http://dx.doi.org/10.2144/000112019>.
26. Suttle CA, Chan AM, Cottrell MT. 1991. Use of ultrafiltration to isolate viruses from seawater which are pathogens of marine phytoplankton. *Appl Environ Microbiol* 57:721–726.
27. Tang P, Chiu C. 2010. Metagenomics for the discovery of novel human viruses. *Future Microbiol* 5:177–189. <http://dx.doi.org/10.2217/fmb.09.120>.
28. Van Etten JL, Lane LC, Dunigan DD. 2010. DNA viruses: the really big ones (giruses). *Annu Rev Microbiol* 64:83–99. <http://dx.doi.org/10.1146/annurev.micro.112408.134338>.
29. Temmam S, Monteil-Bouchard S, Robert C, Pascalis H, Michelle C, Jardot P, Charrel R, Raoult D, Desnues C. 2015. Host-associated metagenomics: a guide to generating infectious RNA viromes. *PLoS One* 10: e0139810. <http://dx.doi.org/10.1371/journal.pone.0139810>.
30. Kohl C, Brinkmann A, Dabrowski PW, Radonić A, Nitsche A, Kurth A. 2015. Protocol for metagenomic virus detection in clinical specimens. *Emerg Infect Dis* 21:48–57. <http://dx.doi.org/10.3201/eid2101.140766>.
31. He S, Wurtzel O, Singh K, Froula JL, Yilmaz S, Tringe SG, Wang Z, Chen F, Lindquist EA, Sorek R, Hugenholtz P. 2010. Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. *Nat Methods* 7:807–812. <http://dx.doi.org/10.1038/nmeth.1507>.
32. Ambrose HE, Clewley JP. 2006. Virus discovery by sequence-independent genome amplification. *Rev Med Virol* 16:365–383. <http://dx.doi.org/10.1002/rmv.515>.
33. Mangul S, Wu NC, Mancuso N, Zelikovsky A, Sun R, Eskin E. 2014. Accurate viral population assembly from ultra-deep sequencing data. *Bioinformatics* 30:i329–i337. <http://dx.doi.org/10.1093/bioinformatics/btu295>.
34. Namiki T, Hachiya T, Tanaka H, Sakakibara Y. 2012. MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res* 40:e155. <http://dx.doi.org/10.1093/nar/gks678>.
35. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244:359–362. <http://dx.doi.org/10.1126/science.2523562>.
36. Miller MB, Tang YW. 2009. Basic concepts of microarrays and potential applications in clinical microbiology. *Clin Microbiol Rev* 22: 611–633. <http://dx.doi.org/10.1128/CMR.00019-09>.
37. Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, DeRisi JL. 2002. Microarray-based detection and genotyping of viral pathogens. *Proc Natl Acad Sci U S A* 99:15687–15692. <http://dx.doi.org/10.1073/pnas.242579699>.
38. Schrenzel J, Kostic T, Bodrossy L, Francois P. 2009. Introduction to microarray-based detection methods, p 1–34. In Kostic T, Butaye P, Schrenzel J (ed), *Detection of Highly Dangerous Pathogens: Microarray Methods for BSL 3 and BSL 4 Agents*. Wiley-Blackwell, Berlin, Germany.
39. Hua W, Zhang G, Guo S, Li W, Sun L, Xiang G. 2015. Microarray-based genotyping and detection of drug-resistant HBV mutations from 620 Chinese patients with chronic HBV infection. *Braz J Infect Dis* 19:291–295. <http://dx.doi.org/10.1016/j.bjid.2015.03.012>.
40. Guo X, Geng P, Wang Q, Cao B, Liu B. 2014. Development of a single nucleotide polymorphism DNA microarray for the detection and genotyping of the SARS coronavirus. *J Microbiol Biotechnol* 24:1445–1454. <http://dx.doi.org/10.4014/jmb.1404.04024>.
41. Martínez MA, Soto-Del Río ML, Gutiérrez RM, Chiu CY, Greninger AL, Contreras JF, López S, Arias CF, Isa P. 2015. DNA microarray for detection of gastrointestinal viruses. *J Clin Microbiol* 53:136–145. <http://dx.doi.org/10.1128/JCM.01317-14>.
42. Chiu CY, Greninger AL, Kanada K, Kwok T, Fischer KF, Runckel C, Louie JK, Glaser CA, Yagi S, Schnurr DP, Haggerty TD, Parsonnet J, Ganem D, DeRisi JL. 2008. Identification of cardioviruses related to Theiler's murine encephalomyelitis virus in human infections. *Proc Natl Acad Sci U S A* 105:14124–14129. <http://dx.doi.org/10.1073/pnas.0805968105>.
43. Kistler A, Avila PC, Rouskin S, Wang D, Ward T, Yagi S, Schnurr D, Ganem D, DeRisi JL, Boushey HA. 2007. Pan-viral screening of respiratory tract infections in adults with and without asthma reveals unexpected human coronavirus and human rhinovirus diversity. *J Infect Dis* 196:817–825. <http://dx.doi.org/10.1086/520816>.
44. Palacios G, Quan PL, Jabado OJ, Conlan S, Hirschberg DL, Liu Y, Zhai J, Renwick N, Hui J, Hegyi H, Grolla A, Strong JE, Towner JS, Geisbert TW, Jahrling PB, Büchen-Osmond C, Ellerbrok H, Sanchez-Seco MP, Lussier Y, Formenty P, Nichol MST, Feldmann H, Briese T, Lipkin WI. 2007. Panmicrobial oligonucleotide array for diagnosis of infectious diseases. *Emerg Infect Dis* 13:73–81. http://dx.doi.org/10.3201/eid1301_060837.
45. Muerhoff AS, Leary TP, Desai SM, Mushahwar IK. 1997. Amplification and subtraction methods and their application to the discovery of novel human viruses. *J Med Virol* 53:96–103. [http://dx.doi.org/10.1002/\(SICI\)1096-9071\(199709\)53:1<96::AID-JMV16>3.0.CO;2-V](http://dx.doi.org/10.1002/(SICI)1096-9071(199709)53:1<96::AID-JMV16>3.0.CO;2-V).
46. Lisitsyn NA, Lisitsyn NM, Wigler M. 1993. Cloning the differences between two genomes. *Science* 259:946–951.
47. Chang Y, Cesarmen E, Pessin MS, Lee F, Culpepper J, Knowles DM, Moore PS. 1994. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 266:1865–1869. <http://dx.doi.org/10.1126/science.7997879>.
48. Weber G, Shendure J, Tanenbaum DM, Church GM, Meyerson M. 2002. Identification of foreign gene sequences by transcript filtering against the human genome. *Nat Genet* 30:141–142. <http://dx.doi.org/10.1038/ng818>.
49. Mirvish ED, Shuda M. 2016. Strategies for human tumor virus discoveries: from microscopic observation to digital transcriptome subtraction. *Front Microbiol* 7:676. <http://dx.doi.org/10.3389/fmicb.2016.00676>.
50. Feng H, Shuda M, Chang Y, Moore PS. 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* 319:1096–1100.
51. Feng H, Taylor JL, Benos PV, Newton R, Waddell K, Lucas SB, Chang Y, Moore PS. 2007. Human transcriptome subtraction by using short sequence tags to search for tumor viruses in conjunctival carcinoma. *J Virol* 81:11332–11340.
52. Palacios G, Druce J, Du L, Tran T, Birch C, Briese T, Conlan S, Quan PL, Hui J, Marshall J, Simons JF, Egholm M, Paddock CD, Shieh WJ, Goldsmith CS, Zaki SR, Catton M, Lipkin WI. 2008. A new arenavirus in a cluster of fatal transplant-associated diseases. *N Engl J Med* 358:991–998. <http://dx.doi.org/10.1056/NEJMoa073785>.
53. Zheng L, Gibbs MJ, Rodoni BC. 2008. Quantitative PCR measurements of the effects of introducing inosines into primers provides guidelines for improved degenerate primer design. *J Virol Methods* 153:97–103. <http://dx.doi.org/10.1016/j.jviromet.2008.07.029>.
54. Rose TM, Henikoff JG, Henikoff S. 2003. CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. *Nucleic Acids Res* 31:3763–3766. <http://dx.doi.org/10.1093/nar/gkg524>.
55. Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM, Henikoff S. 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res* 26:1628–1635. <http://dx.doi.org/10.1093/nar/26.7.1628>.
56. Phaneuf CR, Oh K, Pak N, Saunders DC, Conrardy C, Landers JP, Tong S, Forest CR. 2013. Sensitive, microliter PCR with consensus degenerate primers for Epstein Barr virus amplification. *Biomed Micro-devices* 15:221–231. <http://dx.doi.org/10.1007/s10544-012-9720-1>.
57. van Boheemen S, Bestebroer TM, Verhagen JH, Osterhaus ADME, Pas SD, Herfst S, Fouchier RAM. 2012. A family-wide RT-PCR assay

- for detection of paramyxoviruses and application to a large-scale surveillance study. *PLoS One* 7:e34961. <http://dx.doi.org/10.1371/journal.pone.0034961>.
58. Zlateva KT, Crusio KM, Leontovich AM, Lauber C, Claas E, Kravchenko AA, Spaan WJM, Gorbalenya AE. 2011. Design and validation of consensus-degenerate hybrid oligonucleotide primers for broad and sensitive detection of corona- and toroviruses. *J Virol Methods* 177:174–183. <http://dx.doi.org/10.1016/j.jviromet.2011.08.005>.
59. Vázquez A, Sánchez-Seco M-P, Palacios G, Molero F, Reyes N, Ruiz S, Aranda C, Marqués E, Escosa R, Moreno J, Figuerola J, Tenorio A. 2012. Novel flaviviruses detected in different species of mosquitoes in Spain. *Vector Borne Zoonotic Dis* 12:223–229. <http://dx.doi.org/10.1089/vbz.2011.0687>.
60. Matsuno K, Weisend C, Kajihara M, Matysiak C, Williamson BN, Simunza M, Mweene AS, Takada A, Tesh RB, Ebihara H. 2015. Comprehensive molecular detection of tick-borne phleboviruses leads to the retrospective identification of taxonomically unassigned bunyaviruses and the discovery of a novel member of the genus phlebovirus. *J Virol* 89:594–604. <http://dx.doi.org/10.1128/JVI.02704-14>.
61. Nichol ST, Spiropoulou CF, Morzunov S, Rollin PE, Ksiazek TG, Feldmann H, Sanchez A, Childs J, Zaki S, Peters CJ. 1993. Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness. *Science* 262:914–917. <http://dx.doi.org/10.1126/science.8235615>.
62. Dedkov VG, Magassouba N, Safonova MV, Bodnev SA, Pyankov OV, Camara J, Sylla B, Agafonov AP, Maleev VV, Shipulin GA. 2018. Sensitive multiplex real-time RT-qPCR assay for the detection of filoviruses. *Heal Secur* 16:14–21.
63. Hyun J, Ko DH, Lee SK, Kim HS, Kim JS, Song W, Kim HS. 2018. Evaluation of a new multiplex real-time PCR assay for detecting gastro-enteritis-causing viruses in stool samples. *Ann Lab Med* 38:220–225.
64. Yang F, Wu H, Liu F, Lu X, Peng X, Wu N. 2018. Establishment of a multiplex real-time RT-PCR assay for rapid identification of H6 subtype avian influenza viruses. *Arch Virol* 163:1671–1675. <http://dx.doi.org/10.1007/s00705-018-3773-2>.
65. Qin M, Wang DY, Huang F, Nie K, Qu M, Wang M, Shu YL, Ma XJ. 2010. Detection of pandemic influenza A H1N1 virus by multiplex reverse transcription-PCR with a GeXP analyzer. *J Virol Methods* 168:255–258. <http://dx.doi.org/10.1016/j.jviromet.2010.04.031>.
66. Sas MA, Vina-Rodriguez A, Mertens M, Eiden M, Emmerich P, Chaintoutis SC, Mirazimi A, Groschup MH. 2018. A one-step multiplex real-time RT-PCR for the universal detection of all currently known CCHFV genotypes. *J Virol Methods* 255:38–43. <http://dx.doi.org/10.1016/j.jviromet.2018.01.013>.
67. Boonham N, Kreuze J, Winter S, van der Vlugt R, Bergervoet J, Tomlinson J, Mumford R. 2014. Methods in virus diagnostics: from ELISA to next generation sequencing. *Virus Res* 186:20–31. <http://dx.doi.org/10.1016/j.virusres.2013.12.007>.
68. Hoelscher M, Dowling WE, Sanders-Buell E, Carr JK, Harris ME, Thomschke A, Robb ML, Birx DL, McCutchan FE. 2002. Detection of HIV-1 subtypes, recombinants, and dual infections in east Africa by a multi-region hybridization assay. *AIDS* 16:2055–2064. <http://dx.doi.org/10.1097/00002030-200210180-00011>.
69. Kijak GH, Tovanabutra S, Sanders-Buell E, Watanaveeradej V, de Souza MS, Nelson KE, Ketsararat V, Gulgolarn V, Wera-arpachai M, Sriplienchan S, Khamboonrueng C, Birx DL, Robb ML, McCutchan FE. 2007. Distinguishing molecular forms of HIV-1 in Asia with a high-throughput, fluorescent genotyping assay, MHAbe v.2. *Virology* 358:178–191. <http://dx.doi.org/10.1016/j.virol.2006.07.055>.
70. Freitas FB, Esteves A, Piedade J, Parreira R. 2013. Novel multiregion hybridization assay for the identification of the most prevalent genetic forms of the human immunodeficiency virus type 1 circulating in Portugal. *AIDS Res Hum Retroviruses* 29:318–328.
71. Vidal N, Diop H, Montavon C, Butel C, Bosch S, Ngole EM, Touré-Kane C, Mboup S, Delaporte E, Peeters M. 2013. A novel multiregion hybridization assay reveals high frequency of dual inter-subtype infections among HIV-positive individuals in Cameroon, West Central Africa. *Infect Genet Evol* 14:73–82. <http://dx.doi.org/10.1016/j.meegid.2012.11.017>.
72. Yan L, Zhou J, Zheng Y, Gamson AS, Roembke BT, Nakayama S, Sintim HO. 2014. Isothermal amplified detection of DNA and RNA. *Mol Biosyst* 10:970–1003. <http://dx.doi.org/10.1039/c3mb70304e>.
73. Deiman B, van Aarle P, Sillekens P. 2002. Characteristics and applications of nucleic acid sequence-based amplification (NASBA). *Mol Biotechnol* 20:163–179. <http://dx.doi.org/10.1385/MB:20:2:163>.
74. Barbieri D, Venturoli S, Rösl F, Rincon-Orozco B. 2014. Detection of high-risk human papillomavirus type 16 and 18 using isothermal helicase-dependent amplification. *Diagn Microbiol Infect Dis* 79:178–182. <http://dx.doi.org/10.1016/j.diagmicrobio.2014.02.012>.
75. Tong Y, McCarthy K, Kong H, Lemieux B. 2012. Development and comparison of a rapid isothermal nucleic acid amplification test for typing of herpes simplex virus types 1 and 2 on a portable fluorescence detector. *J Mol Diagn* 14:569–576. <http://dx.doi.org/10.1016/j.jmoldx.2012.05.005>.
76. Patel P, Abd El Wahed A, Faye O, Prüger P, Kaiser M, Thaloengsok S, Ubol S, Sakuntabhai A, Leparc-Goffart I, Hufert FT, Sall AA, Weidmann M, Niedrig M. 2016. A field-deployable reverse transcription recombinase polymerase amplification assay for rapid detection of the Chikungunya virus. *PLoS Negl Trop Dis* 10:e0004953. <http://dx.doi.org/10.1371/journal.pntd.0004953>.
77. Bonney LC, Watson RJ, Afrough B, Mullojonova M, Dzhuraeva V, Tishkova F, Hewson R. 2017. A recombinase polymerase amplification assay for rapid detection of Crimean-Congo haemorrhagic fever virus infection. *PLoS Negl Trop Dis* 11:e0006013. <http://dx.doi.org/10.1371/journal.pntd.0006013>.
78. Yehia N, Arafa AS, Abd El Wahed A, El-Sanousi AA, Weidmann M, Shalaby MA. 2015. Development of reverse transcription recombinase polymerase amplification assay for avian influenza H5N1 HA gene detection. *J Virol Methods* 223:45–49. <http://dx.doi.org/10.1016/j.jviromet.2015.07.011>.
79. Lillis L, Lehman D, Singhal MC, Cantera J, Singleton J, Labarre P, Toyama A, Piepenburg O, Parker M, Wood R, Overbaugh J, Boyle DS. 2014. Non-instrumented incubation of a recombinase polymerase amplification assay for the rapid and sensitive detection of proviral HIV-1 DNA. *PLoS One* 9:e108189. <http://dx.doi.org/10.1371/journal.pone.0108189>.
80. Abd El Wahed A, Patel P, Heidenreich D, Hufert FT, Weidmann M. 2013. Reverse transcription recombinase polymerase amplification assay for the detection of Middle East respiratory syndrome coronavirus. *PLoS Curr* 5:ecurrents.outbreaks.62df1c7c75fffc96cd59034531e2e8364.
81. Notomi T, Mori Y, Tomita N, Kanda H. 2015. Loop-mediated isothermal amplification (LAMP): principle, features, and future prospects. *J Microbiol* 53:1–5. <http://dx.doi.org/10.1007/s12275-015-4656-9>.
82. Yaqing H, Wenping Z, Zhiyi Y, Xionghu W, Shouyi Y, Hong Y, Yingchun D, Guifang H. 2012. Detection of human enterovirus 71 reverse transcription loop-mediated isothermal amplification (RT-LAMP). *Lett Appl Microbiol* 54:233–239. <http://dx.doi.org/10.1111/j.1472-765X.2011.03198.x>.
83. Oloniniyi OK, Kurosaki Y, Miyamoto H, Takada A, Yasuda J. 2017. Rapid detection of all known ebolavirus species by reverse transcription-loop-mediated isothermal amplification (RT-LAMP). *J Virol Methods* 246:8–14. <http://dx.doi.org/10.1016/j.jviromet.2017.03.011>.
84. Nakauchi M, Takayama I, Takahashi H, Tashiro M, Kageyama T. 2014. Development of a reverse transcription loop-mediated isothermal amplification assay for the rapid diagnosis of avian influenza A (H7N9) virus infection. *J Virol Methods* 204:101–104. <http://dx.doi.org/10.1016/j.jviromet.2014.03.028>.
85. Curtis KA, Morrison D, Rudolph DL, Shankar A, Bloomfield LSP, Switzer WM, Owen SM. 2018. A multiplexed RT-LAMP assay for detection of group M HIV-1 in plasma or whole blood. *J Virol Methods* 255:91–97. <http://dx.doi.org/10.1016/j.jviromet.2018.02.012>.

86. Rohatensky MG, Livingstone DM, Mintchev P, Barnes HK, Nakoneshny SC, Demetrick DJ, Dort JC, van Marle G. 2018. Assessing the performance of a loop mediated isothermal amplification (LAMP) assay for the detection and subtyping of high-risk subtypes [sic] of human papilloma virus (HPV) for oropharyngeal squamous cell carcinoma (OPSCC) without DNA purification. *BMC Cancer* 18:166. <http://dx.doi.org/10.1186/s12885-018-4087-1>.
87. Sabalza M, Yasmin R, Barber CA, Castro T, Malamud D, Kim BJ, Zhu H, Montagna RA, Abrams WR. 2018. Detection of Zika virus using reverse-transcription LAMP coupled with reverse dot blot analysis in saliva. *PLoS One* 13:e0192398. <http://dx.doi.org/10.1371/journal.pone.0192398>.
88. Ziros PG, Kokkinos PA, Allard A, Vantarakis A. 2015. Development and evaluation of a loop-mediated isothermal amplification assay for the detection of adenovirus 40 and 41. *Food Environ Virol* 7:276–285. <http://dx.doi.org/10.1007/s12560-015-9182-8>.
89. Reyes GR, Kim JP. 1991. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. *Mol Cell Probes* 5:473–481. [http://dx.doi.org/10.1016/S0890-8508\(05\)80020-9](http://dx.doi.org/10.1016/S0890-8508(05)80020-9).
90. Lambden PR, Cooke SJ, Caul EO, Clarke IN. 1992. Cloning of noncultivable human rotavirus by single primer amplification. *J Virol* 66:1817–1822.
91. Matsui SM, Kim JP, Greenberg HB, Su W, Sun Q, Johnson PC, DuPont HL, Oshiro LS, Reyes GR. 1991. The isolation and characterization of a Norwalk virus-specific cDNA. *J Clin Invest* 87:1456–1461. <http://dx.doi.org/10.1172/JCI11512>.
92. Matsui SM, Kim JP, Greenberg HB, Young LM, Smith LS, Lewis TL, Herrmann JE, Blacklow NR, Dupuis K, Reyes GR. 1993. Cloning and characterization of human astrovirus immunoreactive epitopes. *J Virol* 67: 1712–1715.
93. van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJM, Wolthers KC, Wertheim-van Dillen PME, Kaandorp J, Spaargaren J, Berkhout B. 2004. Identification of a new human coronavirus. *Nat Med* 10:368–373. <http://dx.doi.org/10.1038/nm1024>.
94. de Souza Luna LK, Baumgarte S, Grywna K, Panning M, Drexler JF, Drosten C. 2008. Identification of a contemporary human parechovirus type 1 by VIDISCA and characterisation of its full genome. *Virol J* 5:26. <http://dx.doi.org/10.1186/1743-422X-5-26>.
95. Kim KH, Bae JW. 2011. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. *Appl Environ Microbiol* 77:7663–7668. <http://dx.doi.org/10.1128/AEM.00289-11>.
96. Rosseel T, Van Borm S, Vandebussche F, Hoffmann B, van den Berg T, Beer M, Höper D. 2013. The origin of biased sequence depth in sequence-independent nucleic acid amplification and optimization for efficient massive parallel sequencing. *PLoS One* 8:e76144. <http://dx.doi.org/10.1371/journal.pone.0076144>.
97. Karlsson OE, Belák S, Granberg F. 2013. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. *Biosecur Bioterror* 11(Suppl 1):S227–S234. <http://dx.doi.org/10.1089/bst.2013.0008>.
98. Endoh D, Mizutani T, Kirisawa R, Maki Y, Saito H, Kon Y, Morikawa S, Hayashi M. 2005. Species-independent detection of RNA virus by representational difference analysis using non-ribosomal hexanucleotides for reverse transcription. *Nucleic Acids Res* 33:e65. <http://dx.doi.org/10.1093/nar/gni064>.
99. de Vries M, Deijns M, Canuti M, van Schaik BDC, Faria NR, van de Garde MDB, Jachimowski LCM, Jebbink MF, Jakobs M, Luyf ACM, Coenjaerts FEJ, Claas ECJ, Molenkamp R, Koekkoek SM, Lammens C, Leus F, Goossens H, Ieven M, Baas F, van der Hoek L. 2011. A sensitive assay for virus discovery in respiratory clinical samples. *PLoS One* 6: e16118. <http://dx.doi.org/10.1371/journal.pone.0016118>.
100. Froussard P. 1992. A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Res* 20:2900. <http://dx.doi.org/10.1093/nar/20.11.2900>.
101. Fouchier RAM, Hartwig NG, Bestebroer TM, Niemeyer B, de Jong JC, Simon JH, Osterhaus ADME. 2004. A previously undescribed coronavirus associated with respiratory disease in humans. *Proc Natl Acad Sci U S A* 101:6212–6216. <http://dx.doi.org/10.1073/pnas.0400762101>.
102. Li K, Brownley A, Stockwell TB, Beeson K, McIntosh TC, Busam D, Ferriera S, Murphy S, Levy S. 2008. Novel computational methods for increasing PCR primer design effectiveness in directed sequencing. *BMC Bioinformatics* 9:191. <http://dx.doi.org/10.1186/1471-2105-9-191>.
103. Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, Andersson B. 2005. Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proc Natl Acad Sci U S A* 102: 12891–12896. <http://dx.doi.org/10.1073/pnas.0504666102>.
104. Allander T, Andreasson K, Gupta S, Bjerkner A, Bogdanovic G, Persson MAA, Dalianis T, Ramqvist T, Andersson B. 2007. Identification of a third human polyomavirus. *J Virol* 81:4130–4136. <http://dx.doi.org/10.1128/JVI.00028-07>.
105. Li L, Barry P, Yeh E, Glaser C, Schnurr D, Delwart E. 2009. Identification of a novel human gammapapillomavirus species. *J Gen Virol* 90:2413–2417. <http://dx.doi.org/10.1099/vir.0.012344-0>.
106. Tan V, Van Doorn HR, Van der Hoek L, Minh Hien V, Jebbink MF, Quang Ha D, Farrar J, Van Vinh Chau N, de Jong MD. 2011. Random PCR and ultracentrifugation increases sensitivity and throughput of VIDISCA for screening of pathogens in clinical specimens. *J Infect Dev Ctries* 5:142–148. <http://dx.doi.org/10.3855/jidc.1087>.
107. Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, Afonso C, Zhang X, Anderson NG, Ghedin E, Spiro DJ. 2008. Viral genome sequencing by random priming methods. *BMC Genomics* 9:5. <http://dx.doi.org/10.1186/1471-2164-9-5>.
108. Daly GM, Bexfield N, Heaney J, Stubbs S, Mayer AP, Palser A, Kellam P, Drou N, Caccamo M, Tiley L, Alexander GJM, Bernal W, Heaney JL. 2011. A viral discovery methodology for clinical biopsy samples utilising massively parallel next generation sequencing. *PLoS One* 6:e28879. <http://dx.doi.org/10.1371/journal.pone.0028879>.
109. Hang J, Forshey BM, Kochel TJ, Li T, Solórzano VF, Halsey ES, Kuschner RA. 2012. Random amplification and pyrosequencing for identification of novel viral genome sequences. *J Biomol Tech* 23:4–10. <http://dx.doi.org/10.7171/jbt.12-2301-001>.
110. Neill JD, Bayles DO, Ridpath JF. 2014. Simultaneous rapid sequencing of multiple RNA virus genomes. *J Virol Methods* 201:68–72. <http://dx.doi.org/10.1016/j.jviromet.2014.02.016>.
111. Johne R, Müller H, Rector A, van Ranst M, Stevens H. 2009. Rolling-circle amplification of viral DNA genomes using phi29 polymerase. *Trends Microbiol* 17:205–211. <http://dx.doi.org/10.1016/j.tim.2009.02.004>.
112. Silander K, Saarela J. 2008. Whole genome amplification with Phi29 DNA polymerase to enable genetic or genomic analysis of samples of low DNA yield. *Methods Mol Biol* 439:1–18. http://dx.doi.org/10.1007/978-1-59745-188-8_1.
113. Allen LZ, Ishoey T, Novotny MA, McLean JS, Lasken RS, Williamson SJ. 2011. Single virus genomics: a new tool for virus discovery. *PLoS One* 6:e17722. <http://dx.doi.org/10.1371/journal.pone.0017722>.
114. Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore SF, Egholm M, Lasken RS. 2002. Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci U S A* 99:5261–5266. <http://dx.doi.org/10.1073/pnas.082089499>.
115. Rector A, Tachezy R, Van Ranst M. 2004. A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification. *J Virol* 78:4993–4998. <http://dx.doi.org/10.1128/JVI.78.10.4993-4998.2004>.
116. Li L, Shan T, Soji OB, Alam MM, Kunz TH, Zaidi SZ, Delwart E. 2011. Possible cross-species transmission of circoviruses and cycloviruses among farm animals. *J Gen Virol* 92:768–772. <http://dx.doi.org/10.1099/vir.0.028704-0>.

117. Rockett R, Barracough KA, Isbel NM, Dudley KJ, Nissen MD, Sloots TP, Bialasiewicz S. 2015. Specific rolling circle amplification of low-copy human polyomaviruses BKV, HPyV6, HPyV7, TSPyV, and STLPyV. *J Virol Methods* 215-216:17-21. <http://dx.doi.org/10.1016/j.jviromet.2015.02.004>.
118. Macera L, Cortey M, Maggi F, Segalés J, Kekarainen T. 2011. A novel rolling circle amplification assay to detect members of the family Anelloviridae in pigs and humans. *Virus Res* 160:424-427. <http://dx.doi.org/10.1016/j.virusres.2011.06.025>.
119. Zhong Y, Hu S, Xu C, Zhao Y, Xu D, Zhao Y, Zhao J, Li Z, Zhang X, Zhang H, Li J. 2014. A novel method for detection of HBVcccDNA in hepatocytes using rolling circle amplification combined with in situ PCR. *BMC Infect Dis* 14:608. <http://dx.doi.org/10.1186/s12879-014-0608-y>.
120. Singhal N, Kumar M, Kanaujia PK, Virdi JS. 2015. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front Microbiol* 6:791. <http://dx.doi.org/10.3389/fmicb.2015.00791>.
121. Schwahn AB, Wong JW, Downard KM. 2009. Subtyping of the influenza virus by high resolution mass spectrometry. *Anal Chem* 81: 3500-3506. <http://dx.doi.org/10.1021/ac900026f>.
122. Downard KM. 2013. Proteotyping for the rapid identification of influenza virus and other biopathogens. *Chem Soc Rev* 42:8584-8595. <http://dx.doi.org/10.1039/c3cs60081e>.
123. Sampath R, Russell KL, Massire C, Eshoo MW, Harpin V, Blyn LB, Melton R, Ivy C, Pennella T, Li F, Levene H, Hall TA, Libby B, Fan N, Walcott DJ, Ranken R, Pear M, Schink A, Gutierrez J, Drader J, Moore D, Metzgar D, Addington L, Rothman R, Gaydos CA, Yang S, St George K, Fuschino ME, Dean AB, Stallknecht DE, Goekjian G, Yingst S, Monteville M, Saad MD, Whitehouse CA, Baldwin C, Rudnick KH, Hofstadler SA, Lemon SM, Ecker DJ. 2007. Global surveillance of emerging influenza virus genotypes by mass spectrometry. *PLoS One* 2:e489. <http://dx.doi.org/10.1371/journal.pone.0000489>.
124. Piao J, Jiang J, Xu B, Wang X, Guan Y, Wu W, Liu L, Zhang Y, Huang X, Wang P, Zhao J, Kang X, Jiang H, Cao Y, Zheng Y, Jiang Y, Li Y, Yang Y, Chen W. 2012. Simultaneous detection and identification of enteric viruses by PCR-mass assay. *PLoS One* 7:e42251. <http://dx.doi.org/10.1371/journal.pone.0042251>.
125. Grant-Klein RJ, Baldwin CD, Turell MJ, Rossi CA, Li F, Lovari R, Crowder CD, Matthews HE, Rounds MA, Eshoo MW, Blyn LB, Ecker DJ, Sampath R, Whitehouse CA. 2010. Rapid identification of vector-borne flaviviruses by mass spectrometry. *Mol Cell Probes* 24:219-228. <http://dx.doi.org/10.1016/j.mcp.2010.04.003>.
126. Kriegsmann M, Wandernoth P, Lisenko K, Casadonte R, Longuespée R, Arens N, Kriegsmann J. 2017. Detection of HPV subtypes by mass spectrometry in FFPE tissue specimens: a reliable tool for routine diagnostics. *J Clin Pathol* 70:417-423. <http://dx.doi.org/10.1136/clinpath-2016-204017>.
127. Calderaro A, Arcangeletti M-C, Rodighiero I, Buttrini M, Gorrini C, Motta F, Germi D, Medici M-C, Chezzi C, De Conto F. 2014. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry applied to virus identification. *Sci Rep* 4:6803. <http://dx.doi.org/10.1038/srep06803>.
128. Calderaro A, Arcangeletti MC, Rodighiero I, Buttrini M, Montecchini S, Vasile Simone R, Medici MC, Chezzi C, De Conto F. 2016. Identification of different respiratory viruses, after a cell culture step, by matrix assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS). *Sci Rep* 6:36082. <http://dx.doi.org/10.1038/srep36082>.
129. Ganova-Raeva LM, Dimitrova ZE, Campo DS, Khudyakov Y. 2012. Application of mass spectrometry to molecular surveillance of hepatitis B and C viral infections. *Antivir Ther* 17(7 Part B):1477-1482. <http://dx.doi.org/10.3851/IMP2466>.
130. Zürcher S, Mooser C, Lüthi AU, Mühlmann K, Barbani MT, Mohacs P, Garzoni C, Gorgievski-Hrisoho M, Schaller A, Flatz L. 2012. Sensitive and rapid detection of ganciclovir resistance by PCR based MALDI-TOF analysis. *J Clin Virol* 54:359-363. <http://dx.doi.org/10.1016/j.jcv.2012.04.019>.
131. Foord AJ, White JR, Colling A, Heine HG. 2013. Microsphere suspension array assays for detection and differentiation of Hendra and Nipah viruses. *BioMed Res Int* 2013:289295. <http://dx.doi.org/10.1155/2013/289295>.
132. Shi X, Wu R, Shi M, Zhou L, Wu M, Yang Y, An X, Dai W, Tian L, Zhang C, Ma X, Zhao G. 2016. Simultaneous detection of 13 viruses involved in meningoencephalitis using a newly developed multiplex PCR Mag-array system. *Int J Infect Dis* 49:80-86. <http://dx.doi.org/10.1016/j.ijid.2016.05.023>.
133. Ali Z, Wang J, Tang Y, Liu B, He N, Li Z. 2016. Simultaneous detection of multiple viruses based on chemiluminescence and magnetic separation. *Biomater Sci* 5:57-66. <http://dx.doi.org/10.1039/C6BM00527F>.
134. Balciooglu M, Rana M, Hizir MS, Robertson NM, Haque K, Yigit MV. 2017. Rapid visual screening and programmable subtype classification of Ebola virus biomarkers. *Adv Healthc Mater* 6:1600739.
135. Dirks RM, Pierce NA. 2004. Triggered amplification by hybridization chain reaction. *Proc Natl Acad Sci U S A* 101:15275-15278. <http://dx.doi.org/10.1073/pnas.0407024101>.
136. Mohamed S, Penaranda G, Gonzalez D, Camus C, Khiri H, Boulmé R, Sayada C, Philibert P, Olive D, Halfon P. 2014. Comparison of ultra-deep versus Sanger sequencing detection of minority mutations on the HIV-1 drug resistance interpretations after virological failure. *AIDS* 28: 1315-1324. <http://dx.doi.org/10.1097/QAD.0000000000000267>.
137. Barzon L, Lavezzo E, Militello V, Toppo S, Palù G. 2011. Applications of next-generation sequencing technologies to diagnostic virology. *Int J Mol Sci* 12:7861-7884. <http://dx.doi.org/10.3390/ijms12117861>.
138. Beerenwinkel N, Zagordi O. 2011. Ultra-deep sequencing for the analysis of viral populations. *Curr Opin Virol* 1:413-418. <http://dx.doi.org/10.1016/j.coviro.2011.07.008>.
139. Fisher R, van Zyl GU, Travers SAA, Kosakovsky Pond SL, Engelbrech S, Murrell B, Scheffler K, Smith D. 2012. Deep sequencing reveals minor protease resistance mutations in patients failing a protease inhibitor regimen. *J Virol* 86:6231-6237. <http://dx.doi.org/10.1128/JVI.06541-11>.
140. Gibson RM, Meyer AM, Winner D, Archer J, Feyertag F, Ruiz-Mateos E, Leal M, Robertson DL, Schmotzer CL, Quinones-Mateu ME. 2014. Sensitive deep-sequencing-based HIV-1 genotyping assay to simultaneously determine susceptibility to protease, reverse transcriptase, integrase, and maturation inhibitors, as well as HIV-1 coreceptor tropism. *Antimicrob Agents Chemother* 58:2167-2185. <http://dx.doi.org/10.1128/AAC.02710-13>.
141. Van Laethem K, Theys K, Vandamme AM. 2015. HIV-1 genotypic drug resistance testing: digging deep, reaching wide? *Curr Opin Virol* 14:16-23. <http://dx.doi.org/10.1016/j.coviro.2015.06.001>.
142. Vrancken B, Lequime S, Theys K, Lemey P. 2010. Covering all bases in HIV research: unveiling a hidden world of viral evolution. *AIDS Rev* 12:89-102.
143. Tsibris AMN, Korber B, Arnaout R, Russ C, Lo CC, Leitner T, Gaschen B, Theiler J, Paredes R, Su Z, Hughes MD, Gulick RM, Greaves W, Coakley E, Flexner C, Nusbaum C, Kuritzkes DR. 2009. Quantitative deep sequencing reveals dynamic HIV-1 escape and large population shifts during CCR5 antagonist therapy in vivo. *PLoS One* 4:e5683. <http://dx.doi.org/10.1371/journal.pone.0005683>.
144. Wang C, Mitsuya Y, Gharizadeh B, Ronagh M, Shafer RW. 2007. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *Genome Res* 17:1195-1201. <http://dx.doi.org/10.1101/gr.646830>.
145. Dybowski JN, Heider D, Hoffmann D. 2010. Structure of HIV-1 quasi-species as early indicator for switches of co-receptor tropism. *AIDS Res Ther* 7:41. <http://dx.doi.org/10.1186/1742-6405-7-41>.
146. Bachem CWB, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RGF, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RGF. 1996. Visualization of differential gene

- expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. *Plant J* 9:745–753. <http://dx.doi.org/10.1046/j.1365-313X.1996.9050745.x>.
147. Koch N, Yahi N, Colson P, Fantini J, Tamalet C. 1999. Genetic polymorphism near HIV-1 reverse transcriptase resistance-associated codons is a major obstacle for the line probe assay as an alternative method to sequence analysis. *J Virol Methods* 80:25–31. [http://dx.doi.org/10.1016/S0166-0934\(99\)00030-0](http://dx.doi.org/10.1016/S0166-0934(99)00030-0).
148. Alvarez M, García F, Martínez NM, Hernández Quero J, Louwagie J, De Brauwer A, Maroto MC. 2004. Retrospective analysis of anti-retroviral HIV treatment success based on medical history or guided by the reverse hybridisation LiPA HIV genotyping system. *J Med Virol* 73:151–157. <http://dx.doi.org/10.1002/jmv.20069>.
149. Barzon L, Lavezzo E, Costanzi G, Franchin E, Toppo S, Palù G. 2013. Next-generation sequencing technologies in diagnostic virology. *J Clin Virol* 58:346–350. <http://dx.doi.org/10.1016/j.jcv.2013.03.003>.
150. Boltz VF, Bao Y, Lockman S, Halvas EK, Kearney MF, McIntyre JA, Schooley RT, Hughes MD, Coffin JM, Mellors JW, Zwickl B, Mutuluza CK, Kaseba C, Maponga CC, Watts H, Kuritzkes D, Campbell TB, Kidd-Freeman L, Carten M, Hitti J, Marovich M, Mugenyi PN, Rwambuya S, Sanne IM, Putnam B, Marcus C, Wester C, DiFrancesco R, Halvas E, Beddison A, Lehrman S, Aweka F, Dong B, Ziba PN, Saag MS, Holmes WC, Hammer SM, OCTANE/A5208 Team. 2014. Low-frequency nevirapine (NVP)-resistant HIV-1 variants are not associated with failure of antiretroviral therapy in women without prior exposure to single-dose NVP. *J Infect Dis* 209:703–710. <http://dx.doi.org/10.1093/infdis/jit635>.
151. Redd AD, Quinn TC, Tobian AAR. 2013. Frequency and implications of HIV superinfection. *Lancet Infect Dis* 13:622–628. [http://dx.doi.org/10.1016/S1473-3099\(13\)70066-5](http://dx.doi.org/10.1016/S1473-3099(13)70066-5).
152. Lee GQ, Swenson LC, Poon AFY, Martin JN, Hatano H, Deeks SG, Harrigan PR. 2012. Prolonged and substantial discordance in prevalence of raltegravir-resistant HIV-1 in plasma versus PBMC samples revealed by 454 “deep” sequencing. *PLoS One* 7:e46181. <http://dx.doi.org/10.1371/journal.pone.0046181>.
153. Margeridon-Thermet S, Shulman NS, Ahmed A, Shahriar R, Liu T, Wang C, Holmes SP, Babrzadeh F, Gharizadeh B, Hanczaruk B, Simen BB, Egholm M, Shafer RW. 2009. Ultra-deep pyrosequencing of hepatitis B virus quasispecies from nucleoside and nucleotide reverse-transcriptase inhibitor (NRTI)-treated patients and NRTI-naïve patients. *J Infect Dis* 199:1275–1285. <http://dx.doi.org/10.1086/597808>.
154. Wang GP, Sherrill-Mix SA, Chang K-M, Quince C, Bushman FD. 2010. Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. *J Virol* 84:6218–6228. <http://dx.doi.org/10.1128/JVI.02271-09>.
155. Grubaugh ND, Weger-Lucarelli J, Murrieta RA, Fauver JR, Garcia-Luna SM, Prasad AN, Black WC, IV, Ebel GD. 2016. Genetic drift during systemic arbovirus infection of mosquito vectors leads to decreased relative fitness during host switching. *Cell Host Microbe* 19:481–492. <http://dx.doi.org/10.1016/j.chom.2016.03.002>.
156. 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. <http://dx.doi.org/10.1371/journal.pgen.1006111>.
157. Parameswaran P, Charlebois P, Tellez Y, Nunez A, Ryan EM, Malboeuf CM, Levin JZ, Lennon NJ, Balmaseda A, Harris E, Henn MR. 2012. Genome-wide patterns of intrahuman dengue virus diversity reveal associations with viral phylogenetic clade and interhost diversity. *J Virol* 86:8546–8558. <http://dx.doi.org/10.1128/JVI.00736-12>.
158. Lambrechts L, Lequime S. 2016. Evolutionary dynamics of dengue virus populations within the mosquito vector. *Curr Opin Virol* 21:47–53. <http://dx.doi.org/10.1016/j.coviro.2016.07.013>.
159. Hou W, Bonkovsky HL. 2013. Non-coding RNAs in hepatitis C-induced hepatocellular carcinoma: dysregulation and implications for early detection, diagnosis and therapy. *World J Gastroenterol* 19:7836–7845. <http://dx.doi.org/10.3748/wjg.v19.i44.7836>.
160. Wei B, Kang J, Kibukawa M, Chen L, Qiu P, Lahser F, Marton M, Levitan D. 2016. Development and validation of a template-independent next-generation sequencing assay for detecting low-level resistance-associated variants of hepatitis C virus. *J Mol Diagn* 18:643–656. <http://dx.doi.org/10.1016/j.jmoldx.2016.04.001>.
161. Koonin EV, Dolja VV, Krupovic M. 2015. Origins and evolution of viruses of eukaryotes: the ultimate modularity. *Virology* 479:480:2–25. <http://dx.doi.org/10.1016/j.virol.2015.02.039>.
162. Renzette N, Bhattacharjee B, Jensen JD, Gibson L, Kowalik TF. 2011. Extensive genome-wide variability of human cytomegalovirus in congenitally infected infants. *PLoS Pathog* 7:e1001344. <http://dx.doi.org/10.1371/journal.ppat.1001344>.
163. Lisboa LF, Tong Y, Kumar D, Pang XL, Asberg A, Hartmann A, Rollag H, Jardine AG, Pescovitz MD, Humar A. 2012. Analysis and clinical correlation of genetic variation in cytomegalovirus. *Transpl Infect Dis* 14:132–140. <http://dx.doi.org/10.1111/j.1399-3062.2011.00685.x>.
164. Görzer I, Trajanoski S, Popow-Kraupp T, Puchhammer-Stöckl E. 2015. Analysis of human cytomegalovirus strain populations in urine samples of newborns by ultra deep sequencing. *J Clin Virol* 73:101–104. <http://dx.doi.org/10.1016/j.jcv.2015.11.003>.
165. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. 2004. Classification of papillomaviruses. *Virology* 324:17–27. <http://dx.doi.org/10.1016/j.virol.2004.03.033>.
166. Bzhalava D, Mühr LSA, Lagheden C, Ekström J, Forslund O, Dillner J, Hultin E. 2014. Deep sequencing extends the diversity of human papillomaviruses in human skin. *Sci Rep* 4:5807. <http://dx.doi.org/10.1038/srep05807>.
167. Burk RD, Harari A, Chen Z. 2013. Human papillomavirus genome variants. *Virology* 445:232–243. <http://dx.doi.org/10.1016/j.virol.2013.07.018>.
168. Ndiaye C, Mena M, Alemany L, Arbyn M, Castellsagué X, Laporte L, Bosch FX, de Sanjosé S, Trottier H. 2014. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. *Lancet Oncol* 15:1319–1331. [http://dx.doi.org/10.1016/S1470-2045\(14\)70471-1](http://dx.doi.org/10.1016/S1470-2045(14)70471-1).
169. Ghedin E, Fitch A, Boyne A, Griesemer S, DePasse J, Bera J, Zhang X, Halpin RA, Smit M, Jennings L, St George K, Holmes EC, Spiro DJ. 2009. Mixed infection and the genesis of influenza virus diversity. *J Virol* 83:8832–8841. <http://dx.doi.org/10.1128/JVI.00773-09>.
170. Tsai KN, Chen GW. 2011. Influenza genome diversity and evolution. *Microbes Infect* 13:479–488. <http://dx.doi.org/10.1016/j.micinf.2011.01.013>.
171. Crusat M, Liu J, Palma AS, Childs RA, Liu Y, Wharton SA, Lin YP, Coombs PJ, Martin SR, Matrosovich M, Chen Z, Stevens DJ, Hien VM, Thanh TT, Nhu NT, Nguyet LA, Ha Q, van Doorn HR, Hien TT, Conradt HS, Kiso M, Gamblin SJ, Chai W, Skehel JJ, Hay AJ, Farrar J, de Jong MD, Feizzi T. 2013. Changes in the hemagglutinin of H5N1 viruses during human infection–influence on receptor binding. *Virology* 447:326–337. <http://dx.doi.org/10.1016/j.virol.2013.08.010>.
172. Waybright N, Petrangolo E, Lowary P, Bogan J, Mulholland N. 2008. Detection of human virulence signatures in H5N1. *J Virol Methods* 154:200–205. <http://dx.doi.org/10.1016/j.jviromet.2008.09.013>.
173. Deng YM, Caldwell N, Barr IG. 2011. Rapid detection and subtyping of human influenza A viruses and reassortants by pyrosequencing. *PLoS One* 6:e23400. <http://dx.doi.org/10.1371/journal.pone.0023400>.
174. Yu X, Jin T, Cui Y, Pu X, Li J, Xu J, Liu G, Jia H, Liu D, Song S, Yu Y, Xie L, Huang R, Ding H, Kou Y, Zhou Y, Wang Y, Xu X, Yin Y, Wang J, Guo C, Yang X, Hu L, Wu X, Wang H, Liu J, Zhao G, Zhou J, Pan J, Gao GF, Yang R, Wang J. 2014. Influenza H7N9 and H9N2 viruses: coexistence in poultry linked to human H7N9 infection and genome characteristics. *J Virol* 88:3423–3431. <http://dx.doi.org/10.1128/JVI.02059-13>.
175. Ghedin E, Holmes EC, DePasse JV, Pinilla LT, Fitch A, Hamelin ME, Papenburg J, Boivin G. 2012. Presence of oseltamivir-resistant pandemic

- A/H1N1 minor variants before drug therapy with subsequent selection and transmission. *J Infect Dis* 206:1504–1511. <http://dx.doi.org/10.1093/infdis/jis571>.
176. Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghwanji J, Bhatt S, Peiris JSM, Guan Y, Rambaut A. 2009. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459:1122–1125. <http://dx.doi.org/10.1038/nature08182>.
177. Trovão NS, Suchard MA, Baele G, Gilbert M, Lemey P. 2015. Bayesian inference reveals host-specific contributions to the epidemic expansion of influenza A H5N1. *Mol Biol Evol* 32:3264–3275.
178. Neverov A, Chumakov K. 2010. Massively parallel sequencing for monitoring genetic consistency and quality control of live viral vaccines. *Proc Natl Acad Sci U S A* 107:20063–20068. <http://dx.doi.org/10.1073/pnas.1012537107>.
179. Victoria JG, Wang C, Jones MS, Jaing C, McLoughlin K, Gardner S, Delwart EL. 2010. Viral nucleic acids in live-attenuated vaccines: detection of minority variants and an adventitious virus. *J Virol* 84:6033–6040. <http://dx.doi.org/10.1128/JVI.02690-09>.
180. Firth C, Lipkin WI. 2013. The genomics of emerging pathogens. *Annu Rev Genomics Hum Genet* 14:281–300. <http://dx.doi.org/10.1146/annurev-genom-091212-153446>.
181. Lipkin WI. 2010. Microbe hunting. *Microbiol Mol Biol Rev* 74:363–377. <http://dx.doi.org/10.1128/MMBR.00007-10>.
182. Chiu CY. 2013. Viral pathogen discovery. *Curr Opin Microbiol* 16:468–478. <http://dx.doi.org/10.1016/j.mib.2013.05.001>.