CrossMark

Review

Evolution of selective-sequencing approaches for virus discovery and virome analysis

Arvind Kumar[a], Satyapramod Murthy[a], Amit Kapoor[a,b,*]

[a] Center for Vaccines and Immunity, The Research Institute at Nationwide Children's Hospital, Columbus, OH 43205, USA
[b] Department of Pediatrics, College of Medicine and Public Health, Ohio State University, Columbus, OH 43210, USA

ABSTRACT

Recent advances in sequencing technologies have transformed the field of virus discovery and virome analysis. Once mostly confined to the traditional Sanger sequencing based individual virus discovery, is now entirely replaced by high throughput sequencing (HTS) based virus metagenomics that can be used to characterize the nature and composition of entire viromes. To better harness the potential of HTS for the study of viromes, sample preparation methodologies use different approaches to exclude amplification of non-viral components that can overshadow low-titer viruses. These virus-sequence enrichment approaches mostly focus on the sample preparation methods, like enzymatic digestion of non-viral nucleic acids and size exclusion of non-viral constituents by column filtration, ultrafiltration or density gradient centrifugation. However, recently a new approach of virus-sequence enrichment called virome-capture sequencing, focused on the amplification or HTS library preparation stage, was developed to increase the ability of virome characterization. This new approach has the potential to further transform the field of virus discovery and virome analysis, but its technical complexity and sequence-dependence warrants further improvements. In this review we discuss the different methods, their applications and evolution, for selective sequencing based virome analysis and also propose refinements needed to harness the full potential of HTS for virome analysis.

1. Introduction

Despite being the simplest of biological entity, viruses play an enormously important and complex role in human and animal health, environmental ecology and are also known to shape the evolution of their hosts (Greenbaum and Ghedin, 2015; Koonin and Dolja, 2013; Koonin et al., 2015; Koonin and Wolf, 2012; Mager and Stoye, 2015). Virus discovery during most of the previous century followed a traditional approach of virus isolation and amplification in cell culture or animal models (Leland and Ginocchio, 2007; Palacios and Oberste, 2005). These virus isolates were then classified based on their morphological and serological properties (Gelderblom, 1996; Muir et al., 1998). The widely used cell culture isolation of viruses failed to identify viruses that were refractory to grow in vitro, like hepatitis C virus (HCV) (Choo et al., 1989).,Moreover, traditional approaches including serological assays and PCR were very specific or focused on a group of viruses. The inherent dependence of these approaches on the biological properties or sequences of known viruses resulted in our limited knowledge of the virus world (Koonin, 2010; Shi et al., 2015).

Use of sequence dependent (i.e; generic PCR assays and microarray) and sequence independent (i.e; single primer amplification (SISPA) and random priming) approaches for nucleic acid amplification combined with Sanger sequencing or HTS allowed the rapid identification of new viruses after 1980 (Bishop-Lilly et al., 2010; Chang et al., 1994; Day et al., 2010; Grard et al., 2012; Kapoor et al., 2015; Ladner et al., 2016; Linnen et al., 1996; Matsui et al., 1991; Mokili et al., 2012; Muerhoff et al., 1997; Nichol et al., 1993; Qin et al., 2014; Quan et al., 2010; Simons et al., 1995b) (Fig. 1). In SISPA method, a primer-binding sequence(adapter) is ligated to both ends of restriction enzyme digested DNA fragments allowing their subsequent amplification using a single primer. Random PCR uses a slightly different approach where the primer-binding sequence is annealed to the templates using an oligo containing mixture of all four nucleotides at terminal 6-9 positions of the 3' end. However, the comprehensive identification of all viruses in a given sample poses unique challenges due to the tiny and complex nature of viral genomes. Additionally, absence of conserved anchor sequences similar to the bacterial or eukaryotic ribosomal sequences prohibit pan-amplification and sequencing of viruses (Kapoor and Lipkin, 2014). Considering these challenges, an ideal sequencing based virus identification approach should use a selection method to increase
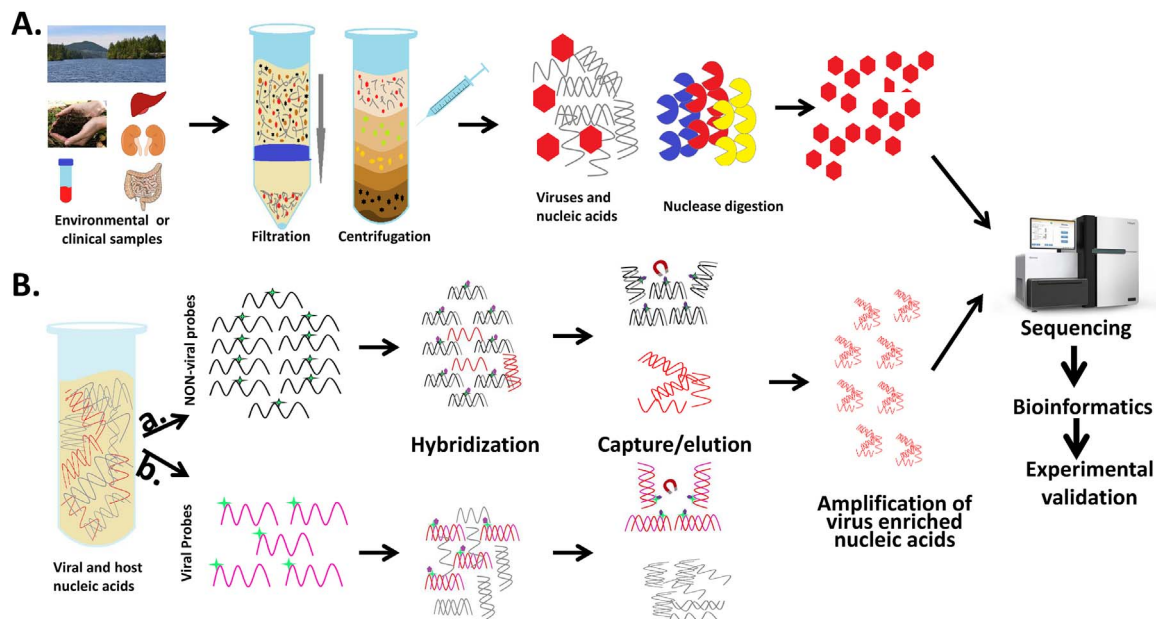
Fig. 1. Experimental outline of virus enrichment and selective sequencing approaches. (A). Details of the pre-extraction virus enrichment based selective sequencing. Briefly, samples are filtered and/or centrifuged before treatment with nucleases (DNAse, RNAse, Benzonase etc.), followed by their amplification and sequencing. (B). Outline of negative (non-viral probes) and positive (viral probes) selection method using hybridization based capture and selective enrichment of virus nucleic acids, followed by amplification and sequencing. Sequencing data analysis requires bioinformatics based virus discovery or virome analysis and experimental validation of results.

the viral sequences or reduce non-viral sequences while remaining unbiased towards the nature and composition of the virus genome.

The term metagenomics can be defined as the study of genetic materials recovered from environmental samples (Handelsman et al., 1998). The culture-independent nature of this approach allowed discovery of unprecedented microbial diversity that remained under-estimated by culture-dependent approaches. This vastly useful information gained by bacterial metagenomics studies, ignited interest of marine virologists to adapt this approach for viral metagenomics (Breitbart et al., 2002). Soon after, several groups used this approach, or its modified forms, to identify human and animal viruses. The early use of the metagenomics approach for human and animal clinical samples was very successful. New viruses were identified from in vitro cultures supernatants, respiratory secretions, stool suspensions, urine and other body fluids (Allander et al., 2001; Allander et al., 2005; Jones et al., 2005; Kapoor et al., 2008; Victoria et al., 2008). Metagenomics remains the most widely used approach for new virus identification and virome analysis to date. Here we discuss the use and limitations of virus metagenomics; recently used selective sequencing based virome analysis and provide a perspective on refinements of virus metagenomics based approaches.

2. Evolution and use of sequencing technologies for virus metagenomics

Simple dideoxy or Sanger sequencing was not very applicable for viral metagenomics but was the only option available before 2005. The process involved plasmid cloning of fragmented or amplified nucleic acids from environmental or clinical samples followed by sequencing of individual clones. Technical and cost constrains allowed sequencing of only a few hundred clones in most of these early studies but even this small scale sequencing revealed presence of vast diversity among viral communities and also led to identification of several human and animal viruses from clinical samples (Allander et al., 2001; Allander et al., 2005; Chang et al., 1994; Choo et al., 1989; Kapoor et al., 2009; Kapoor et al., 2008; Linnen et al., 1996; Matsui et al., 1991; Muerhoff et al., 1997; Nichol et al., 1993; Nishizawa et al., 1997; van der Hoek et al., 2004; Victoria et al., 2008). These studies highlighted the importance of

adapting more efficient sequencing technologies for viral metagenomics. The section below briefly describe different sequencing platforms, their advantages and disadvantages for virus metagenomics based virus identification and virome analysis.

2.1. 454 or pyroseqeuncing

This is the first HTS technology used for metagenomics based virus discovery. The principle of 454 (later purchased by Roche) technology is sequencing-by-synthesis chemistry in which DNA molecules are amplified through an emulsion PCR, generating multiple clones of DNA using a single template. Pyrophosphates released during base incorporation are enzymatically converted to a light signal which can be detected using a charged couple devices camera (Droege and Hill, 2008). 454 sequencing was widely used to identify several new viruses and virome profiles from human and animal samples (Day et al., 2010) including arboviruses (Bishop-Lilly et al., 2010), orbiviruses (Li et al., 2014), arenaviruses (Palacios et al., 2008), Lujo virus (Briese et al., 2009), astrovirus (Quan et al., 2010), gyroviruses (Phan et al., 2012), porcine bocaviruses (Yu et al., 2013), picornaviruses (Boros et al., 2012), rhabdoviruses (Grard et al., 2012), coronaviruses (Honkavuori et al., 2014), gammapapillomavirus (Phan et al., 2013) and seadorna-virus (Reuter et al., 2013). Most of these viruses were identified from samples like serum, respiratory, and fecal samples. Some studies used this technology to identify viruses from tissue and organ samples- canine adenovirus 1(van der Heijden et al., 2012), bat hepacivirus and pegivirus (Quan et al., 2013). This method generally allows sequencing of 500,000 to 5 million individual reads that can be between 100 and 450 nucleotides long. Although, this technology offered a higher yield than Sanger sequencing at a lower cost, this technology has been supplanted by other NGS technologies due to its high cost, error rate in homopolymeric regions and low throughput.

2.2. Ion torrent

Life Technologies/Thermo Fisher Scientific released a pH-mediated or semiconductor sequencing technology based Ion personal genomics machine (PGM) sequencer platform in 2010 followed by the Ion Proton

(2012) and Ion S5 series (2015). These platforms are conceptually similar to the 454 pyrosequencing platform in template preparation and sequencing steps. Adapter-ligated DNA fragments are amplified by emulsion-PCR on the surface of beads which are distributed into microwells where a sequencing-by-synthesis reaction occurs. Protons released during nucleotide incorporation are detected using an ion sensor, which can measure slight shifts in pH (Merriman et al., 2012; Reuter et al., 2015). Rapid sequencing runs make these sequencers particularly useful for the targeted detection of viruses in clinical samples like HIV (Archer et al., 2012; Chang et al., 2013; Gibson et al., 2014), hepatitis B virus (Yan et al., 2015), HCV (Gaspareto et al., 2016; Marascio et al., 2016), and the rapid genome sequencing of several viruses including the toscana virus (Nougairede et al., 2013), polyomavirus (Anthony et al., 2013), porcine reproductive and respiratory syndrome virus (Kvisgaard et al., 2013), orthoreovirus (Steyer et al., 2013), bluetongue virus (Lorusso et al., 2014), rotavirus (Ndze et al., 2014; Nyaga et al., 2014), influenza virus (Van den Hoecke et al., 2015) etc. Although, some studies used this technology to study viromes in skin (Bzhalava et al., 2013), ticks (Tokarz et al., 2014; Xia et al., 2015), gut virome in piglets (Karlsson et al., 2016) and seals (Kluge et al., 2016), this platform is not the ideal choice for virome study in human clinical samples due to lower outputs.

2.3. Illumina

Solexa/Illumina introduces a high throughput platform and is less expensive in cost than the Genome analyzer II platform in 2006 followed by MiSeq (2011), NextSeq 500 (2014) and the HiSeq series (2012–2014). The principle of Illumina sequencing involves reversible-termination sequencing by synthesis (SBS) with fluorescently labelled nucleotides (Liu et al., 2012). Introduction of these platforms accelerated the rate of virus discovery in humans and animals like severe fever with thrombocytopenia virus (Xu et al., 2011), bas-congo virus (Grard et al., 2012), titi monkey adenovirus (Chen et al., 2011), canine bocavirus 3 (Li et al., 2013), snake arenaviruses (Stenglein et al., 2012), human polyomavirus 9 (Sauvage et al., 2011), simian adenovirus c (Chiu et al., 2013), theiler's disease associated virus (Chandriani et al., 2013), human hepegivirus 1 (Kapoor et al., 2015), jingmen tick virus (Qin et al., 2014), guaico culex virus (Ladner et al., 2016), proto-parvovirus (Phan et al., 2016), marmota himalayana hepatovirus (Yu et al., 2016) etc. High throughput and low error rates (below 1% and mainly substitutions) are the main reasons that Illumina technologies have dominated the viral discovery field in the past several years.

2.4. PacBio sequencing

Pacific Biosciences commercialized a single-molecule real-time (SMRT) sequencing platform known as RS II in 2010 followed by Sequel (2015) for generation of long read (average 10–50 kb) without clonal amplification bias. Hairpin adapters are ligated on the end of template DNA molecules for generation of capped templates (SMRT-bell). Zero-mode waveguides (ZMW) with a single molecule of DNA and fluorescent nucleotides are affixed with a single DNA polymerase enzyme at the bottom. During polymerization, a fluorescent tag of nucleotides are cleaved off and diffuse out of the observation area of the ZMW and are detected by the detector in real time (Reuter et al., 2015; Rhoads and Au, 2015). Single molecule sequencing is extremely well suited for virus metagenomics studies, because of the very long reads. It is otherwise nearly impossible to obtain long contigs in a biologically diverse sample, (Brinzevich et al., 2014; Tombacz et al., 2015; Archer et al., 2012; Bergfors et al., 2016; Schleiss et al., 2014; Tombacz et al., 2014; Wittmann et al., 2014) but lower throughput, higher costs per base sequencing and higher error rates, currently limit the scope for the viral metagenomics study of low titer viruses.

2.5. Nanopore sequencing

Oxford Nanopore Technologies released a nanopore based portable sequencer MinION in 2014. The principle of this technology is measuring information about the characteristic changes that are induced as the biological molecules (DNA/RNA) are passing through the nanopore by a molecular motor protein. This is the first portable sequencer with the capability of RNA/DNA sequencing, longer read length (approximate 300 kb) in fewer hours, and real time sequence analysis. The nanopore platform has been used for viral detection in human clinical samples (Greninger et al., 2015; Hoenen et al., 2016) and genome sequencing (Karamitros et al., 2016; Kilianski et al., 2016). Although this platform looks promising for epidemiological investigation during an outbreak, low outputs and high error rates (around 10%) (McGinn et al., 2016) are a concern for virus discovery.

3. Selective sequencing approaches

Currently high throughput sequencing technologies are capable of producing of millions of sequence reads, but the detection of viruses in clinical samples is still challenging due to the presence of an extremely low amount of viral nucleic acids compared to a high background of host, bacterial and other contaminating genetic material. Therefore, the efficient and sensitive metagenomics study of viruses in clinical or environmental samples requires the removal of non-viral nucleic acids and/or the enrichment of virus-derived nucleic acids (Capobianchi et al., 2013; Conceicao-Neto et al., 2015; Hall et al., 2014). These virus enrichment methods can be broadly classified as sample preparation methods and sequencing library preparation methods Fig. 1, or alternatively can be also called pre-extraction (Fig. 1A) and post-extraction virus enrichment methods (Fig. 1B), respectively.

Choice of pre-extraction virus enrichment methods largely depends on the nature of the sample to be analyzed. Generally, the virus metagenomics samples are contaminated with host (bacteria or eukaryote) and environmental nucleic acids. For the virome analysis of environmental samples, several methods including ultra-filtration, iron chloride precipitation and density gradient centrifugation (polyethylene glycol, sucrose cushion) are available for pre-extraction virus enrichment (Andrews-Pfannkoch et al., 2010; John et al., 2011). For the virome analysis of clinical samples with a low abundance of host cells, like cerebrospinal fluid, respiratory samples, serum, urine or stools, filtration and nuclease digestion or density centrifugation can be used as a method of pre-extraction virus enrichment (Batty et al., 2013; Conceicao-Neto et al., 2015; Daly et al., 2011; Hall et al., 2014; Kohl et al., 2015; Rosseel et al., 2015). For the virome analysis of clinical samples with an abundance of host cells, like blood or tissues, pre-extraction based enrichment is not appropriate as the virus genome itself can be present in its non-capsided or transcribed form.

3.1. Virus particle enrichment methods

Most of the pre-extraction virus enrichment methods are based on the physical properties of virions and their differences from the other living forms. Virus particles or virions are the encapsidated forms of virus genomes. Even the first recognition of virus as an entity in the early 19th century, then recognized as liquid poison, was done on their ability to pass through the Cahmberland filters that can retain most bacteria (Lecoq, 2001). Except for a few large recently identified viruses (Halary et al., 2016), most animal viruses are less than 200–300 nm in diameter and therefore can pass through filters with 0.2–0.45 μm pore size. Filtration is therefore the most commonly used method for selectively sequencing the viruses from environmental or clinical samples. Notably, several metagenomic based virus identification studies used the 0.45 μm filter and not the 0.2 μm filter because the pore diameters of commercial filters are quite variable and using a filter with pore diameter closer to the virus size can reduce the amount of viruses in the

filtrate (Conceicao-Neto et al., 2015).

The compact nature of virions allows the use of density gradient centrifugation (sucrose, cesium chloride, poly ethylene glycol) as a method to enrich samples for virus derived nucleic acids. Density centrifugation is commonly used in molecular virology labs and was also used in the first virus metagenomics study for enriching viruses in sea water samples. This method was also used to study viromes of lakes, soils and from extreme environments (Brum et al., 2013; Kleiner et al., 2015; Thurber et al., 2009). The major limitation of the density centrifugation method is its applicability for analysis of clinical samples. Moreover, the differences in the density of enveloped and non-enveloped viruses and other variables requires the analysis of several gradient fractions. The presence of contaminating nucleic acids in the density gradient solution can further contaminate the samples with environmental nucleic acids.

Most viral capsids can protect their genome from degradation by nucleases. This property of virions was frequently used by plant and animal virologists for identification and transmission of viruses (Rosseel et al., 2015; Thurber et al., 2009). However, the nuclease treatment or nuclease mediated removal of non-capsidated nucleic acids to enrich samples for viral nucleic acids was first described by Allander et al. (Allander et al., 2001). The study showed that DNAse treatment of virus containing serum samples can help enrich the metagenomic libraries from virus derived sequences. This work described the identification of two novel parvovirus species in bovine serum. Soon after, this group and then several others used a similar approach to identify a plethora of human and animal viruses(add REFS). Although it was very helpful, nuclease treatment is not very effective in degrading unprotected nucleic acids and often fails to enrich samples enough for the identification of low-titer viruses. Moreover, nuclease treatment cannot be used to find viruses in cells or tissue samples where most virus nucleic acids are not in their encapsidated form.

3.2. Positive selection based virus or virome capture methods

In positive selection methods, samples are enriched for viral nucleic acids directly using probes targeting the viruses like in PCR assays, microarray or virus capture (in solution based hybridization) approaches. The simplest examples of the positive selection approach are generic PCR assays which use degenerated primers to target several related viruses or their variants (Irving et al., 2014; Kwok and Chiang, 2016; Simons et al., 1995a). The PCR based approach is restricted by its ability to detect only a limited number of viruses due to an issue of multiplexity. To avoid this, several groups developed a positive selection approach for enriching samples for viruses of a defined taxonomic group (family, genus and species) based on DNA microarray (Gardner et al., 2010; Palacios et al., 2007; Wang et al., 2002). DNA microarray has been employed to characterize and discover a number of novel or variant viruses including the human cardioviruses, porcine circovirus, rhinovirus and adenovirus respectively (Chen et al., 2011; Chiu et al., 2008; Kistler et al., 2007; Palacios et al., 2007; Rota et al., 2003). However, due to their limited specificity, none of these methods were suitable for the comprehensive characterization of vertebrate viruses or for the analysis of viromes. Most recently, two very comprehensive probe sets were developed for positive selection based sequencing of all known vertebrate viruses and their variants (Briese et al., 2015; Wylie et al., 2015). These methods use virus-specific probes in a liquid-phase hybridization to capture viral sequences from metagenomics libraries. One of these approaches was termed Virome Capture Sequencing (VirCapSeq) (Briese et al., 2015). Basically, VirCapSeq is similar to any other DNA hybridization based enrichment method, including the well-known exon capture method used for transcriptomics. In principle, VirCapSeq is made of set of specific oligonucleotides (or probes), like used on viral microarrays, that when mixed with samples can hybridize or capture complementary viral nucleic acids (Fig. 1B-b). VirCapSeq includes probes for all viruses known to infect vertebrates by targeting

their protein coding regions. However, the design of VirCapSeq was challenging due to the constant influx of genome data from newly identified viruses and the biased abundance of virus genomes of pathogenic viruses, like HIV and HCV in sequence databases. The design of VirCapSeq, therefore required the use of a unique probe set containing 2 million probes or oligonucleotides that differed by at least 10% to reduce sequence redundancy and widen the coverage of virus sequences. Stringent hybridization conditions of VirCapSeq allow specific enrichment of virus-like sequences in libraries generated from different sample types. Compared to common metagenomics and other virus enrichment approaches (filtration and nuclease digestion prior to total nucleic acid extraction and RiboZero rRNA depletion after extraction), the use of VirCapSeq gained 100–10,000-fold more virus sequences in the metagenomics libraries. The sensitivity of VirCapSeq was also compared with real time PCR for detection of viral sequence in blood and serum samples spiked with live enterovirus D68. VirCapSeq was able to detect 10 copies/ml of virus in both sample types which is comparable to the sensitivity of PCR based methods. Interestingly, the use of VirCapSeq allows the identification of viruses whose genomes are vastly different from those used to design the probes, as all viruses of a taxonomic group share some highly conserved stretches of nucleotides. Another similar positive selection approach was termed ViroCap and was simultaneously developed by Wylie et.al. (Wylie et al., 2015). Like VirCapSeq, ViroCap also targeted most virus species (of 34 virus families) that are known to infect vertebrates and excluded known endogenous retroviruses. To design ViroCap, approximately 1 billion bases of sequence data was condensed to ∼200 million bases of probe sequences. Use of ViroCap allowed 296–674–fold increases in the number of virus reads compared to simple metagenomics sequencing.

Although both these positive selection approaches described above are biased towards sequences of known viruses, they can be efficiently used to study mixed infections and provide a sensitive characterization of all viruses present in clinical samples. As the appreciation of the role that viromes play in human health is increasing, these platforms can provide useful tools to study the dynamics of human virome in longitudinal samples collected before and after appearance of diseases.

3.3. Negative selection based virus metagenomics

An ideal negative subtraction strategy would allow the removal of all non-viral nucleic acids from samples, including the nucleic acids derived from the host, the reagents and the environment. A negative selection approach can use the principle of suppression subtractive hybridization (SSH) or representational difference analysis (RDA) which allows the comparison of two DNA populations (tester and driver) and enrichment of differentially distributed molecules (Diatchenko et al., 1996; Lisitsyn et al., 1993). The genomic DNA sample that contains the sequences of interest is referred to as tester and the reference sample is referred to as driver. Tester and driver DNAs are hybridized in excess of driver and the hybrid sequences are then removed. Consequently, the remaining unhybridized DNAs represent DNA molecules that are present in the tester yet absent from the driver DNA. Although these approaches have been usually employed to identify genes with a differential expression pattern (Yin et al., 2013), but successfully have been used to identify human herpes virus-8 (Chang et al., 1994), GB virus (Simons et al., 1995b), TT virus (Nishizawa et al., 1997), Walrus Calicivirus (Ganova-Raeva et al., 2004) and a new strain of the murine hepatitis virus (Islam et al., 2015). Another simpler approach can be based on the simple subtraction of non-target nucleic acid using biotinylated probes (Fig. 1B-a). An example of negative selection approach is subtraction of ribosomal RNA (rRNA) from extracted ribonucleic acids. This method is very efficient in finding RNA viruses in samples, like serum and cerebrospinal fluid, where a majority of host RNA is rRNA (Matranga et al., 2014; Rosseel et al., 2015). Several commercial kits including Illumina's Ribo-Zero rRna removal kit, and Ambion's GLOBINclear kit provide sequence

specific (human/mouse/rat/bacteria) biotinylated oligos designed to capture host rRNA or highly expressed gene. These hybrids can be captured on a streptavidin magnetic bead and removed from the sample, enriching these samples for.virus-derived nucleic acids or amplification products. New England Biolab's NEBNext rRNA depletion kit uses a slightly different approach where a DNA oligo binds to rRNA sequences in the samples followed by RNase H mediated selective removal of hybridized rRNA. Recently, a new approach known as DASH (depletion of abundant sequences by hybridization) was described by Gu et al. for depletion of rRNA with Cas9 protein complexed with a library of rRNA (Gu et al., 2016). However a major limitation is the unavailability of rRNA subtraction reagents for many vertebrate and invertebrate species. Moreover, rRNA subtraction is not promising when looking for viruses in cells and tissue samples that usually contain an abundance of genomic DNA and transcripts (Li et al., 2016).

Although not yet available, the impact of an ideal negative selection approach will be far reaching as it will enable not only the sensitive analysis of viromes but also the comparative analysis of changes in viromes within a host or between, or to compare viromes in samples obtained at different time points.

4. Conclusions and perspectives

Recent advances in sequencing technologies were successfully exploited for virus discovery and virome analysis expanding our knowledge of viruses that live with and around us. However, these approaches are not sufficient to study the virome of samples where the host and environmental nucleic acids are typically present in orders of magnitude higher than the viral nucleic acids. Viruses residing within cells and tissues represent such a scenario where conventional viral metagenomics fail to identify low-titer viruses; however this problem can be solved by increasing the depth of sequencing. However, the vast amount of sequence data generated using a metagenomics approach can pose bioinformatics challenges for the comparative analysis of viromes. Use of appropriate selective sequencing approaches, therefore can not only allow more sensitive identification of viruses but can also help in reducing bioinformatics of non-viral data.

Although in positive selection approaches, like VirCapSeq or ViroCap, a very broad selection of probes for all known vertebrate viruses allow a comprehensive analysis of related viruses in clinical samples. The major limitation of this technique is its inherent bias of enriching samples for only known and targeted viruses. Even the most basic virus enrichment approaches like centrifugation and filtration were biased and therefore giant viruses remained unidentified (Halary et al., 2016). Some constituents of viromes are either fluctuating or continuously evolving and therefore cannot be studied using a constant probe set. In conclusion, these methods are more appropriate for a selected group of viruses and to study the relative presence and dynamic changes in their populations. In the future, a more focused VirCapSeq assay to study viromes and its fluctuations in specific diseases or body sites like blood, respiratory system, liver or gastrointestinatl tract, holds an enormous potential to gain new insights into the role of viromes in human health and development of chronic diseases. However, we believe that only a simple and efficient negative selection approach will become an ideal selective sequencing approach for the discovery of new viruses and also for meaningful analysis of viromes.

Acknowledgements

References

Allander, T., Emerson, S.U., Engle, R.E., Purcell, R.H., Bukh, J., 2001. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. Proc. Natl. Acad. Sci. U. S. A. 98, 11609–11614.

Allander, T., Tammi, M.T., 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.

Andrews-Pfannkoch, C., Fadrosh, D.W., Thorpe, J., Williamson, S.J., 2010. Hydroxyapatite-mediated separation of double-stranded DNA, single-stranded DNA, and RNA genomes from natural viral assemblages. Appl. Environ. Microbiol. 76, 5039–5045.

Anthony, S.J., St Leger, J.A., Navarrete-Macias, I., Nilson, E., Sanchez-Leon, M., Liang, E., Seimon, T., Jain, K., Karesh, W., Daszak, P., Briese, T., Lipkin, W.I., 2013. Identification of a novel cetacean polyomavirus from a common dolphin (Delphinus delphis) with Tracheobronchitis. PLoS One 8, e68239.

Archer, J., Weber, J., Henry, K., Winner, D., Gibson, R., Lee, L., Paxinos, E., Arts, E.J., Robertson, D.L., Mimms, L., Quinones-Mateu, M.E., 2012. Use of four next-generation sequencing platforms to determine HIV-1 coreceptor tropism. PLoS One 7, e49602.

Batty, E.M., Wong, T.H., Trebes, A., Argoud, K., Attar, M., Buck, D., Ip, C.L., Golubchik, T., Cule, M., Bowden, R., Manganis, C., Klenerman, P., Barnes, E., Walker, A.S., Wyllie, D.H., Wilson, D.J., Dingle, K.E., Peto, T.E., Crook, D.W., Piazza, P., 2013. A modified RNA-Seq approach for whole genome sequencing of RNA viruses from faecal and blood samples. PLoS One 8, e66129.

Bergfors, A., Leenheer, D., Bergqvist, A., Ameur, A., Lennerstrand, J., 2016. Analysis of hepatitis C NS5A resistance associated polymorphisms using ultra deep single molecule real time (SMRT) sequencing. Antiviral Res. 126, 81–89.

Bishop-Lilly, K.A., Turell, M.J., Willner, K.M., Butani, A., Nolan, N.M., Lentz, S.M., Akmal, A., Mateczun, A., Brahmbhatt, T.N., Sozhamannan, S., Whitehouse, C.A., Read, T.D., 2010. Arbovirus detection in insect vectors by rapid, high-throughput pyrosequencing. PLoS Negl. Trop. Dis. 4, e878.

Boros, A., Nemes, C., Pankovics, P., Kapusinszky, B., Delwart, E., Reuter, G., 2012. Identification and complete genome characterization of a novel picornavirus in turkey (Meleagris gallopavo). J. Gen. Virol. 93, 2171–2182.

Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., Azam, F., Rohwer, F., 2002. Genomic analysis of uncultured marine viral communities. Proc. Natl. Acad. Sci. U. S. A. 99, 14250–14255.

Briese, T., Paweska, J.T., McMullan, L.K., Hutchison, S.K., Street, C., Palacios, G., Khristova, M.L., Weyer, J., Swanepoel, R., Egholm, M., Nichol, S.T., Lipkin, W.I., 2009. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. PLoS Pathog. 5, e1000455.

Briese, T., Kapoor, A., Mishra, N., Jain, K., Kumar, A., Jabado, O.J., Lipkin, W.I., 2015. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. mBio 6, e01491–01415.

Brinzevich, D., Young, G.R., Sebra, R., Ayllon, J., Maio, S.M., Deikus, G., Chen, B.K., Fernandez-Sesma, A., Simon, V., Mulder, L.C., 2014. HIV-1 interacts with human endogenous retrovirus K (HML-2) envelopes derived from human primary lymphocytes. J. Virol. 88, 6213–6223.

Brum, J.R., Culley, A.I., Steward, G.F., 2013. Assembly of a marine viral metagenome after physical fractionation. PLoS One 8, e60604.

Bzhalava, D., Johansson, H., Ekstrom, J., Faust, H., Moller, B., Eklund, C., Nordin, P., Stenquist, B., Paoli, J., Persson, B., Forslund, O., Dillner, J., 2013. Unbiased approach for virus detection in skin lesions. PLoS One 8, e65953.

Capobianchi, M.R., Giombini, E., Rozera, G., 2013. Next-generation sequencing technology in clinical virology. Clin. Microbiol. Infect. 19, 15–22.

Chandriani, S., Skewes-Cox, P., Zhong, W., Ganem, D.E., Divers, T.J., Van Blaricum, A.J., Tennant, B.C., Kistler, A.L., 2013. Identification of a previously undescribed divergent virus from the Flaviviridae family in an outbreak of equine serum hepatitis. Proc. Natl. Acad. Sci. U. S. A. 110, E1407–1415.

Chang, Y., Cesarman, E., Pessin, M.S., Lee, F., Culpepper, J., Knowles, D.M., Moore, P.S., 1994. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. Science 266, 1865–1869.

Chang, M.W., Oliveira, G., Yuan, J., Okulicz, J.F., Levy, S., Torbett, B.E., 2013. Rapid deep sequencing of patient-derived HIV with ion semiconductor technology. J. Virol. Methods 189, 232–234.

Chen, E.C., Yagi, S., Kelly, K.R., Mendoza, S.P., Tarara, R.P., Canfield, D.R., Maninger, N., Rosenthal, A., Spinner, A., Bales, K.L., Schnurr, D.P., Lerche, N.W., Chiu, C.Y., 2011. Cross-species transmission of a novel adenovirus associated with a fulminant pneumonia outbreak in a new world monkey colony. PLoS Pathog. 7, e1002155.

Chiu, C.Y., Greninger, A.L., Kanada, K., Kwok, T., Fischer, K.F., Runckel, C., Louie, J.K., Glaser, C.A., Yagi, S., Schnurr, D.P., Haggerty, T.D., Parsonnet, J., Ganem, D., DeRisi, J.L., 2008. Identification of cardioviruses related to Theiler's murine encephalomyelitis virus in human infections. Proc. Natl. Acad. Sci. U. S. A. 105, 14124–14129.

Chiu, C.Y., Yagi, S., Lu, X., Yu, G., Chen, E.C., Liu, M., Dick Jr., E.J., Carey, K.D., Erdman, D.D., Leland, M.M., Patterson, J.L., 2013. A novel adenovirus species associated with an acute respiratory outbreak in a baboon colony and evidence of coincident human infection. mBio 4, e00084.

Choo, Q.L., Kuo, G., Weiner, A.J., Overby, L.R., Bradley, D.W., Houghton, M., 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 244, 359–362.

Conceicao-Neto, N., Zeller, M., Lefrere, H., De Bruyn, P., Beller, L., Deboutte, W., Yinda, C.K., Lavigne, R., Maes, P., Van Ranst, M., Heylen, E., Matthijnssens, J., 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. Sci. Rep. 5, 16532.

Daly, G.M., Bexfield, N., Heaney, J., Stubbs, S., Mayer, A.P., Palser, A., Kellam, P., Drou, N., Caccamo, M., Tiley, L., Alexander, G.J., Bernal, W., Heeney, J.L., 2011. A viral discovery methodology for clinical biopsy samples utilising massively parallel next generation sequencing. PLoS One 6, e28879.

Day, J.M., Ballard, L.L., Duke, M.V., Scheffler, B.E., Zsak, L., 2010. Metagenomic analysis of the turkey gut RNA virus community. Virol. J. 7, 313.

Diatchenko, L., Lau, Y.F., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E.D., Siebert, P.D., 1996. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. U. S. A. 93, 6025–6030.

Droege, M., Hill, B., 2008. The Genome Sequencer FLX System–longer reads, more applications, straight forward bioinformatics and more complete data sets. J. Biotechnol. 136, 3–10.

Ganova-Raeva, L., Smith, A.W., Fields, H., Khudyakov, Y., 2004. New Calicivirus isolated from walrus. Virus Res. 102, 207–213.

Gardner, S.N., Jaing, C.J., McLoughlin, K.S., Slezak, T.R., 2010. A microbial detection array (MDA) for viral and bacterial detection. BMC Genomics 11, 668.

Gaspareto, K.V., Ribeiro, R.M., Malta, F.M., Gomes-Gouvea, M.S., Muto, N.H., Romano, C.M., Mendes-Correa, M.C., Carrilho, F.J., Sabino, E.C., Pinho, J.R., 2016. Resistance-associated variants in HCV subtypes 1a and 1b detected by Ion Torrent sequencing platform. Antivir. Ther. 21 (8), 653–660. http://dx.doi.org/10.3851/IMP3057. Epub 2016 Jun 17. PubMed PMID: 27314166.

Gelderblom, H.R., 1996. Structure and classification of viruses. In: Baron, S. (Ed.), Medical Microbiology, 4th ed. Galveston, TX.

Gibson, R.M., Meyer, A.M., Winner, D., Archer, J., Feyertag, F., Ruiz-Mateos, E., Leal, M., Robertson, D.L., Schmotzer, C.L., Quinones-Mateu, M.E., 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.

Grard, G., Fair, J.N., Lee, D., Slikas, E., Steffen, I., Muyembe, J.J., Sittler, T., Veeraraghavan, N., Ruby, J.G., Wang, C., Makuwa, M., Mulembakani, P., Tesh, R.B., Mazet, J., Rimoin, A.W., Taylor, T., Schneider, B.S., Simmons, G., Delwart, E., Wolfe, N.D., Chiu, C.Y., Leroy, E.M., 2012. A novel rhabdovirus associated with acute hemorrhagic fever in central Africa. PLoS Pathog. 8, e1002924.

Greenbaum, B.D., Ghedin, E., 2015. Viral evolution: beyond drift and shift. Curr. Opin. Microbiol. 26, 109–115.

Greninger, A.L., Naccache, S.N., Federman, S., Yu, G., Mbala, P., Bres, V., Stryke, D., Bouquet, J., Somasekar, S., Linnen, J.M., Dodd, R., Mulembakani, P., Schneider, B.S., Muyembe-Tamfum, J.J., Stramer, S.L., Chiu, C.Y., 2015. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Med. 7, 99.

Gu, W., Crawford, E.D., O'Donovan, B.D., Wilson, M.R., Chow, E.D., Retallack, H., DeRisi, J.L., 2016. Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. Genome Biol. 17, 41.

Halary, S., Temmam, S., Raoult, D., Desnues, C., 2016. Viral metagenomics: are we missing the giants? Curr. Opin. Microbiol. 31, 34–43.

Hall, R.J., Wang, J., Todd, A.K., Bissielo, A.B., Yen, S., Strydom, H., Moore, N.E., Ren, X., Huang, Q.S., Carter, P.E., Peacey, M., 2014. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. J. Virol. Methods 195, 194–204.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., Goodman, R.M., 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem. Biol. 5, R245–249.

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

Honkavuori, K.S., Briese, T., Krauss, S., Sanchez, M.D., Jain, K., Hutchison, S.K., Webster, R.G., Lipkin, W.I., 2014. Novel coronavirus and astrovirus in Delaware Bay shore-birds. PLoS One 9, e93395.

Irving, W.L., Rupp, D., McClure, C.P., Than, L.M., Titman, A., Ball, J.K., Steinmann, E., Bartenschlager, R., Pietschmann, T., Brown, R.J., 2014. Development of a high-throughput pyrosequencing assay for monitoring temporal evolution and resistance associated variant emergence in the Hepatitis C virus protease coding-region. Antiviral Res. 110, 52–59.

Islam, M.M., Toohey, B., Purcell, D.F., Kannourakis, G., 2015. Suppression subtractive hybridization method for the identification of a new strain of murine hepatitis virus from xenografted SCID mice. Arch. Virol. 160, 2945–2955.

John, S.G., Mendez, C.B., Deng, L., Poulos, B., Kauffman, A.K., Kern, S., Brum, J., Polz, M.F., Boyle, E.A., Sullivan, M.B., 2011. A simple and efficient method for concentration of ocean viruses by chemical flocculation. Environ. Microbiol. Rep. 3, 195–202.

Jones, M.S., Kapoor, A., Lukashov, V.V., Simmonds, P., Hecht, F., Delwart, E., 2005. New DNA viruses identified in patients with acute viral infection syndrome. J. Virol. 79, 8230–8236.

Kapoor, A., Lipkin, W. Ian, 2014. Virus Discovery in the 21st Century. In: eLS. John Wiley & Sons Ltd, Chichester. http://www.els.net [doi: 10.1002/9780470015902. a0023621].

Kapoor, A., Victoria, J., Simmonds, P., Slikas, E., Chieochansin, T., Naeem, A., Shaukat, S., Sharif, S., Alam, M.M., Angez, M., Wang, C., Shafer, R.W., Zaidi, S., Delwart, E., 2008. A highly prevalent and genetically diversified Picornaviridae genus in South Asian children. Proc. Natl. Acad. Sci. U. S. A. 105, 20482–20487.

Kapoor, A., Slikas, E., Simmonds, P., Chieochansin, T., Naeem, A., Shaukat, S., Alam, M.M., Sharif, S., Angez, M., Zaidi, S., Delwart, E., 2009. A newly identified bocavirus species in human stool. J. Infect. Dis. 199, 196–200.

Kapoor, A., Kumar, A., Simmonds, P., Bhuva, N., Singh Chauhan, L., Lee, B., Sall, A.A., Jin, Z., Morse, S.S., Shaz, B., Burbelo, P.D., Lipkin, W.I., 2015. Virome analysis of transfusion recipients reveals a novel human virus that shares genomic features with hepaciviruses and pegiviruses. mBio 6, e01466–01415.

Karamitros, T., Harrison, I., Piorkowska, R., Katzourakis, A., Magiorkinis, G., Mbisa, J.L., 2016. De novo assembly of human herpes virus type 1 (HHV-1) genome, mining of non-canonical structures and detection of novel drug-resistance mutations using short- and long-read next generation sequencing technologies. PLoS One 11, e0157600.

Karlsson, O.E., Larsson, J., Hayer, J., Berg, M., Jacobson, M., 2016. The intestinal eukaryotic virome in healthy and diarrhoeic neonatal piglets. PLoS One 11, e0151481.

Kilianski, A., Roth, P.A., Liem, A.T., Hill, J.M., Willis, K.L., Rossmaier, R.D., Marinich, A.V., Maughan, M.N., Karavis, M.A., Kuhn, J.H., Honko, A.N., Rosenzweig, C.N., 2016. Use of unamplified RNA/cDNA-Hybrid nanopore sequencing for rapid detection and characterization of RNA viruses. Emerg. Infect. Dis. 22, 1448–1451.

Kistler, A., Avila, P.C., Rouskin, S., Wang, D., Ward, T., Yagi, S., Schnurr, D., Ganem, D., DeRisi, J.L., Boushey, H.A., 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.

Kleiner, M., Hooper, L.V., Duerkop, B.A., 2015. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. BMC Genomics 16, 7.

Kluge, M., Campos, F.S., Tavares, M., de Amorim, D.B., Valdez, F.P., Giongo, A., Roehe, P.M., Franco, A.C., 2016. Metagenomic survey of viral diversity obtained from feces of Subantarctic and South American fur seals. PLoS One 11, e0151921.

Kohl, C., Brinkmann, A., Dabrowski, P.W., Radonic, A., Nitsche, A., Kurth, A., 2015. Protocol for metagenomic virus detection in clinical specimens. Emerg. Infect. Dis. 21, 48–57.

Koonin, E.V., 2010. The wonder world of microbial viruses. Expert Rev. Anti. Infect. Ther. 10, 96 PubMed PMID: 20954874; PubMed Central PMCID: PMC3293457.

Koonin, E.V., Dolja, V.V., 2013. A virocentric perspective on the evolution of life. Curr. Opin Virol. 3, 546–557.

Koonin, E.V., Wolf, Y.I., 2012. Evolution of microbes and viruses: a paradigm shift in evolutionary biology? Front. Cell. Infect. Microbiol. 2, 119.

Koonin, E.V., Dolja, V.V., Krupovic, M., 2015. Origins and evolution of viruses of eukaryotes: the ultimate modularity. Virology 479–480, 2–25.

Kvisgaard, L.K., Hjulsager, C.K., Fahnoe, U., Breum, S.O., Ait-Ali, T., Larsen, L.E., 2013. A fast and robust method for full genome sequencing of porcine reproductive and respiratory syndrome virus (PRRSV) type 1 and type 2. J. Virol. Methods 193, 697–705.

Kwok, H., Chiang, A.K., 2016. From conventional to next generation sequencing of epstein-barr virus genomes. Viruses 8, 60.

Ladner, J.T., Wiley, M.R., Beitzel, B., Auguste, A.J., Dupuis, A.P., 2nd Lindquist, M.E., Sibley, S.D., Kota, K.P., Fetterer, D., Eastwood, G., Kimmel, D., Prieto, K., Guzman, H., Aliota, M.T., Reyes, D., Brueggemann, E.E., St John, L., Hyeroba, D., Lauck, M., Friedrich, T.C., O'Connor, D.H., Gestole, M.C., Cazares, L.H., Popov, V.L., Castro-Llanos, F., Kochel, T.J., Kenny, T., White, B., Ward, M.D., Loaiza, J.R., Goldberg, T.L., Weaver, S.C., Kramer, L.D., Tesh, R.B., Palacios, G., 2016. A multicomponent animal virus isolated from mosquitoes. Cell Host Microbe 20, 357–367.

Lecoq, H., 2001. [Discovery of the first virus, the tobacco mosaic virus: 1892 or 1898?]. Comptes rendus de l'Academie des sciences. Serie III, Sciences de la vie 324, 929–933.

Leland, D.S., Ginocchio, C.C., 2007. Role of cell culture for virus detection in the age of technology. Clin. Microbiol. Rev. 20, 49–78.

Li, L., Pesavento, P.A., Leutenegger, C.M., Estrada, M., Coffey, L.L., Naccache, S.N., Samayoa, E., Chiu, C., Qiu, J., Wang, C., Deng, X., Delwart, E., 2013. A novel bocavirus in canine liver. Virol. J. 10, 54.

Li, M., Zheng, Y., Zhao, G., Fu, S., Wang, D., Wang, Z., Liang, G., 2014. Tibet Orbivirus, a novel Orbivirus species isolated from Anopheles maculatus mosquitoes in Tibet, China. PLoS One 9, e88738.

Li, D., Li, Z., Zhou, Z., Li, Z., Qu, X., Xu, P., Zhou, P., Bo, X., Ni, M., 2016. Direct next-generation sequencing of virus-human mixed samples without pretreatment is favorable to recover virus genome. Biol. Direct 11, 3.

Linnen, J., Wages Jr., J., Zhang-Keck, Z.Y., Fry, K.E., Krawczynski, K.Z., Alter, H., Koonin, E., Gallagher, M., Alter, M., Hadziyannis, S., Karayiannis, P., Fung, K., Nakatsuji, Y., Shih, J.W., Young, L., Piatak Jr., M., Hoover, C., Fernandez, J., Chen, S., Zou, J.C., Morris, T., Hyams, K.C., Ismay, S., Lifson, J.D., Hess, G., Foung, S.K., Thomas, H., Bradley, D., Margolis, H., Kim, J.P., 1996. Molecular cloning and disease association of hepatitis G virus: a transfusion-transmissible agent. Science 271, 505–508.

Lisitsyn, N., Lisitsyn, N., Wigler, M., 1993. Cloning the differences between two complex genomes. Science 259, 946–951.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L., Law, M., 2012. Comparison of next-generation sequencing systems. J. Biomed. Biotechnol. 2012, 251364.

Lorusso, A., Marcacci, M., Ancora, M., Mangone, I., Leone, A., Marini, V., Camma, C., Savini, G., 2014. Complete genome sequence of bluetongue virus serotype 1 circulating in Italy, obtained through a fast next-generation sequencing protocol. Genome Announcement 2.

Mager, D.L., Stoye, J.P., 2015. Mammalian endogenous retroviruses. Microbiol. Spectr. 3 (MDNA3-0009-2014).

Marascio, N., Pavia, G., Strazzulla, A., Dierckx, T., Cuypers, L., Vrancken, B., Barreca, G.S., Mirante, T., Malanga, D., Oliveira, D.M., Vandamme, A.M., Torti, C., Liberto, M.C., Foca, A., The Sinergie-Umg Study, G., 2016. Detection of natural resistance-associated substitutions by ion semiconductor technology in HCV1b positive, direct-acting antiviral agents-naive patients. Int. J. Mol. Sci. 17.

Matranga, C.B., Andersen, K.G., Winnicki, S., Busby, M., Gladden, A.D., Tewhey, R., Stremlau, M., Berlin, A., Gire, S.K., England, E., Moses, L.M., Mikkelsen, T.S., Odia, I.,

Ehiane, P.E., Folarin, O., Goba, A., Kahn, S.H., Grant, D.S., Honko, A., Hensley, L., Happi, C., Garry, R.F., Malboeuf, C.M., Birren, B.W., Gnirke, A., Levin, J.Z., Sabeti, P.C., 2014. Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples. Genome Biol. 15, 519.

Matsui, S.M., Kim, J.P., Greenberg, H.B., Su, W., Sun, Q., Johnson, P.C., DuPont, H.L., Oshiro, L.S., Reyes, G.R., 1991. The isolation and characterization of a Norwalk virus-specific cDNA. J. Clin. Invest. 87, 1456–1461.

McGinn, S., Bauer, D., Brefort, T., Dong, L., El-Sagheer, A., Elsharawy, A., Evans, G., Falk-Sorqvist, E., Forster, M., Fredriksson, S., Freeman, P., Freitag, C., Fritzsche, J., Gibson, S., Gullberg, M., Gut, M., Heath, S., Heath-Brun, I., Heron, A.J., Hohlbein, J., Ke, R., Lancaster, O., Le Reste, L., Maglia, G., Marie, R., Mauger, F., Mertes, F., Mignardi, M., Moens, L., Oostmeijer, J., Out, R., Pedersen, J.N., Persson, F., Picaud, V., Rotem, D., Schracke, N., Sengenes, J., Stahler, P.F., Stade, B., Stoddart, D., Teng, X., Veal, C.D., Zahra, N., Bayley, H., Beier, M., Brown, T., Dekker, C., Ekstrom, B., Flyvbjerg, H., Franke, A., Guenther, S., Kapanidis, A.N., Kaye, J., Kristensen, A., Lehrach, H., Mangion, J., Sauer, S., Schyns, E., Tost, J., van Helvoort, J.M., van der Zaag, P.J., Tegenfeldt, J.O., Brookes, A.J., Mir, K., Nilsson, M., Willcocks, J.P., Gut, I.G., 2016. New technologies for DNA analysis – a review of the READNA Project. New Biotechnol. 33, 311–330.

Merriman, B., Ion Torrent, R., Team, D., Rothberg, J.M., 2012. Progress in ion torrent semiconductor chip based sequencing. Electrophoresis 33, 3397–3417.

Mokili, J.L., Rohwer, F., Dutilh, B.E., 2012. Metagenomics and future perspectives in virus discovery. Curr. Opin. Virol. 2, 63–77.

Muerhoff, A.S., Leary, T.P., Desai, S.M., Mushahwar, I.K., 1997. Amplification and sub-traction methods and their application to the discovery of novel human viruses. J. Med. Virol. 53, 96–103.

Muir, P., Kammerer, U., Korn, K., Mulders, M.N., Poyry, T., Weissbrich, B., Kandolf, R., Cleator, G.M., van Loon, A.M., 1998. Molecular typing of enteroviruses: current status and future requirements: the European Union Concerted Action on Virus Meningitis and Encephalitis. Clin. Microbiol. Rev. 11, 202–227.

Ndze, V.N., Esona, M.D., Achidi, E.A., Gonsu, K.H., Doro, R., Marton, S., Farkas, S., Ngeng, M.B., Ngu, A.F., Obama-Abena, M.T., Banyai, K., 2014. Full genome char-acterization of human Rotavirus A strains isolated in Cameroon, 2010–2011: diverse combinations of the G and P genes and lack of reassortment of the backbone genes. Infect. Genet. Evol. 28, 537–560.

Nichol, S.T., Spiropoulou, C.F., Morzunov, S., Rollin, P.E., Ksiazek, T.G., Feldmann, H., Sanchez, A., Childs, J., Zaki, S., Peters, C.J., 1993. Genetic identification of a han-tavirus associated with an outbreak of acute respiratory illness. Science 262, 914–917.

Nishizawa, T., Okamoto, H., Konishi, K., Yoshizawa, H., Miyakawa, Y., Mayumi, M., 1997. A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. Biochem. Biophys. Res. Commun. 241, 92–97.

Nougairede, A., Bichaud, L., Thiberville, S.D., Ninove, L., Zandotti, C., de Lamballerie, X., Brouqui, P., Charrel, R.N., 2013. Isolation of Toscana virus from the cerebrospinal fluid of a man with meningitis in Marseille, France, 2010. Vector Borne Zoonotic Dis. 13, 685–688.

Nyaga, M.M., Stucker, K.M., Esona, M.D., Jere, K.C., Mwinyi, B., Shonhai, A., Tsolenyanu, E., Mulindwa, A., Chibumbya, J.N., Adolfine, H., Halpin, R.A., Roy, S., Stockwell, T.B., Berejena, C., Seheri, M.L., Mwenda, J.M., Steele, A.D., Wentworth, D.E., Mphahlele, M.J., 2014. Whole-genome analyses of DS-1-like human G2P[4] and G8P [4] rotavirus strains from eastern, western and southern africa. Virus Genes 49, 196–207.

Palacios, G., Oberste, M.S., 2005. Enteroviruses as agents of emerging infectious diseases. J. Neurovirol. 11, 424–433.

Palacios, G., Quan, P.L., Jabado, O.J., Conlan, S., Hirschberg, D.L., Liu, Y., Zhai, J., Renwick, N., Hui, J., Hegyi, H., Grolla, A., Strong, J.E., Towner, J.S., Geisbert, T.W., Jahrling, P.B., Buchen-Osmond, C., Ellerbrok, H., Sanchez-Seco, M.P., Lussier, Y., Formenty, P., Nichol, M.S., Feldmann, H., Briese, T., Lipkin, W.I., 2007. Panmicrobial oligonucleotide array for diagnosis of infectious diseases. Emerg. Infect. Dis. 13, 73–81.

Palacios, G., Druce, J., Du, L., Tran, T., Birch, C., Briese, T., Conlan, S., Quan, P.L., Hui, J., Marshall, J., Simons, J.F., Egholm, M., Paddock, C.D., Shieh, W.J., Goldsmith, C.S., Zaki, S.R., Catton, M., Lipkin, W.I., 2008. A new arenavirus in a cluster of fatal transplant-associated diseases. N. Engl. J. Med. 358, 991–998.

Phan, T.G., Li, L., O'Ryan, M.G., Cortes, H., Mamani, N., Bonkoungou, I.J., Wang, C., Leutenegger, C.M., Delwart, E., 2012. A third gyrovirus species in human faeces. J. Gen. Virol. 93, 1356–1361.

Phan, T.G., Vo, N.P., Aronen, M., Jartti, L., Jartti, T., Delwart, E., 2013. Novel human gammapapillomavirus species in a nasal swab. Genome Announcement 1, e0002213.

Phan, T.G., Dreno, B., da Costa, A.C., Li, L., Orlandi, P., Deng, X., Kapusinszky, B., Siqueira, J., Knol, A.C., Halary, F., Dantal, J., Alexander, K.A., Pesavento, P.A., Delwart, E., 2016. A new protoparvovirus in human fecal samples and cutaneous T cell lymphomas (mycosis fungoides). Virology 496, 299–305.

Qin, X.C., Shi, M., Tian, J.H., Lin, X.D., Gao, D.Y., He, J.R., Wang, J.B., Li, C.X., Kang, Y.J., Yu, B., Zhou, D.J., Xu, J., Plyusnin, A., Holmes, E.C., Zhang, Y.Z., 2014. A tick-borne segmented RNA virus contains genome segments derived from unsegmented viral ancestors. Proc. Natl. Acad. Sci. U. S. A. 111, 6744–6749.

Quan, P.L., Wagner, T.A., Briese, T., Torgerson, T.R., Hornig, M., Tashmukhamedova, A., Firth, C., Palacios, G., Baisre-De-Leon, A., Paddock, C.D., Hutchison, S.K., Egholm, M., Zaki, S.R., Goldman, J.E., Ochs, H.D., Lipkin, W.I., 2010. Astrovirus encephalitis in boy with X-linked agammaglobulinemia. Emerg. Infect. Dis. 16, 918–925.

Quan, P.L., Firth, C., Conte, J.M., Williams, S.H., Zambrana-Torrelio, C.M., Anthony, S.J., Ellison, J.A., Gilbert, A.T., Kuzmin, I.V., Niezgoda, M., Osinubi, M.O., Recuenco, S., Markotter, W., Breiman, R.F., Kalemba, L., Malekani, J., Lindblade, K.A., Rostal, M.K., Ojeda-Flores, R., Suzan, G., Davis, L.B., Blau, D.M., Ogunkoya, A.B., Alvarez

Castillo, D.A., Moran, D., Ngam, S., Akaibe, D., Agwanda, B., Briese, T., Epstein, J.H., Daszak, P., Rupprecht, C.E., Holmes, E.C., Lipkin, W.I., 2013. Bats are a major natural reservoir for hepaciviruses and pegiviruses. Proc. Natl. Acad. Sci. U. S. A. 110, 8194–8199.

Reuter, G., Boros, A., Delwart, E., Pankovics, P., 2013. Novel seadornavirus (family re-oviridae) related to banna virus in europe. Arch. Virol. 158, 2163–2167.

Reuter, J.A., Spacek, D.V., Snyder, M.P., 2015. High-throughput sequencing technologies. Mol. Cell 58, 586–597.

Rhoads, A., Au, K.F., 2015. PacBio sequencing and its applications. Genomics Proteom. Bioinf. 13, 278–289.

Rosseel, T., Ozhelvaci, O., Freimanis, G., Van Borm, S., 2015. Evaluation of convenient pretreatment protocols for RNA virus metagenomics in serum and tissue samples. J. Virol. Methods 222, 72–80.

Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P., Penaranda, S., Bankamp, B., Maher, K., Chen, M.H., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J.L., Chen, Q., Wang, D., Erdman, D.D., Peret, T.C., Burns, C., Ksiazek, T.G., Rollin, P.E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Gunther, S., Osterhaus, A.D., Drosten, C., Pallansch, M.A., Anderson, L.J., Bellini, W.J., 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300, 1394–1399.

Sauvage, V., Foulongne, V., Cheval, J., Ar Gouilh, M., Pariente, K., Dereure, O., Manuguerra, J.C., Richardson, J., Lecuit, M., Burguiere, A., Caro, V., Eloit, M., 2011. Human polyomavirus related to African green monkey lymphotropic polyomavirus. Emerg. Infect. Dis. 17, 1364–1370.

Schleiss, M.R., McAllister, S., Armien, A.G., Hernandez-Alvarado, N., Fernandez-Alarcon, C., Zabeli, J.C., Ramaraj, T., Crow, J.A., McVoy, M.A., 2014. Molecular and biological characterization of a new isolate of guinea pig cytomegalovirus. Viruses 6, 448–475.

Shi, M., Lin, X.D., Vasilakis, N., Tian, J.H., Li, C.X., Chen, L.J., Eastwood, G., Diao, X.N., Chen, M.H., Chen, X., Qin, X.C., Widen, S.G., Wood, T.G., Tesh, R.B., Xu, J., Holmes, E.C., Zhang, Y.Z., 2015. Divergent Viruses Discovered in Arthropods and Vertebrates Revise the Evolutionary History of the Flaviviridae and Related Viruses. J. Virol. 15–02036 PubMed PMID: 26491167; PubMed Central PMCID: PMC4702705.

Simons, J.N., Leary, T.P., Dawson, G.J., Pilot-Matias, T.J., Muerhoff, A.S., Schlauder, G.G., Desai, S.M., Mushahwar, I.K., 1995a. Isolation of novel virus-like sequences associated with human hepatitis. Nat. Med. 1, 564–569.

Simons, J.N., Pilot-Matias, T.J., Leary, T.P., Dawson, G.J., Desai, S.M., Schlauder, G.G., Muerhoff, A.S., Erker, J.C., Buijk, S.L., Chalmers, M.L., et al., 1995b. Identification of two flavivirus-like genomes in the GB hepatitis agent. Proc. Natl. Acad. Sci. U. S. A. 92, 3401–3405.

Stenglein, M.D., Sanders, C., Kistler, A.L., Ruby, J.G., Franco, J.Y., Reavill, D.R., Dunker, F., Derisi, J.L., 2012. Identification, characterization, and in vitro culture of highly divergent arenaviruses from boa constrictors and annulated tree boas: candidate etiological agents for snake inclusion body disease. mBio 3, e00180–00112.

Steyer, A., Gutierrez-Aguire, I., Kolenc, M., Koren, S., Kutnjak, D., Pokorn, M., Poljsak-Prijatelj, M., Racki, N., Ravnikar, M., Sagadin, M., Fratnik Steyer, A., Toplak, N., 2013. High similarity of novel orthoreovirus detected in a child hospitalized with acute gastroenteritis to mammalian orthoreoviruses found in bats in Europe. J. Clin. Microbiol. 51, 3818–3825.

Thurber, R.V., Haynes, M., Breitbart, M., Wegley, L., Rohwer, F., 2009. Laboratory pro-cedures to generate viral metagenomes. Nat. Protoc. 4, 470–483.

Tokarz, R., Williams, S.H., Sameroff, S., Sanchez Leon, M., Jain, K., Lipkin, W.I., 2014. Virome analysis of Amblyomma americanum, Dermacentor variabilis, and Ixodes scapularis ticks reveals novel highly divergent vertebrate and invertebrate viruses. J. Virol. 88, 11480–11492.

Tombacz, D., Sharon, D., Olah, P., Csabai, Z., Snyder, M., Boldogkoi, Z., 2014. Strain kaplan of pseudorabies virus genome sequenced by PacBio single-Molecule real-Time sequencing technology. Genome Announcement 2.

Tombacz, D., Csabai, Z., Olah, P., Havelda, Z., Sharon, D., Snyder, M., Boldogkoi, Z., 2015. Characterization of novel transcripts in pseudorabies virus. Viruses 7, 2727–2744.

Van den Hoecke, S., Verhelst, J., Vuylsteke, M., Saelens, X., 2015. Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing. BMC Genomics 16, 79.

Victoria, J.G., Kapoor, A., Dupuis, K., Schnurr, D.P., Delwart, E.L., 2008. Rapid identifi-cation of known and new RNA viruses from animal tissues. PLoS Pathog. 4, e1000163.

Wang, D., Coscoy, L., Zylberberg, M., Avila, P.C., Boushey, H.A., Ganem, D., DeRisi, J.L., 2002. Microarray-based detection and genotyping of viral pathogens. Proc. Natl. Acad. Sci. U. S. A. 99, 15687–15692.

Wittmann, J., Dreiseikelmann, B., Rohde, M., Meier-Kolthoff, J.P., Bunk, B., Rohde, C., 2014. First genome sequences of Achromobacter phages reveal new members of the N4 family. Virol. J. 11, 14.

Wylie, T.N., Wylie, K.M., Herter, B.N., Storch, G.A., 2015. Enhanced virome sequencing using targeted sequence capture. Genome Res. 25, 1910–1920.

Xia, H., Hu, C., Zhang, D., Tang, S., Zhang, Z., Kou, Z., Fan, Z., Bente, D., Zeng, C., Li, T., 2015. Metagenomic profile of the viral communities in Rhipicephalus spp ticks from Yunnan, China. PLoS One 10, e0121609.

Xu, B., Liu, L., Huang, X., Ma, H., Zhang, Y., Du, Y., Wang, P., Tang, X., Wang, H., Kang, K., Zhang, S., Zhao, G., Wu, W., Yang, Y., Chen, H., Mu, F., Chen, W., 2011. Metagenomic analysis of fever, thrombocytopenia and leukopenia syndrome (FTLS) in Henan Province, China: discovery of a new bunyavirus. PLoS Pathog. 7, e1002369.

Yan, L., Zhang, H., Ma, H., Liu, D., Li, W., Kang, Y., Yang, R., Wang, J., He, G., Xie, X., Wang, H., Wei, L., Lu, Z., Shao, Q., Chen, H., 2015. Deep sequencing of hepatitis B virus basal core promoter and precore mutants in HBeAg-positive chronic hepatitis B patients. Sci. Rep. 5, 17950.

Yin, Y., Jiang, L., Fang, D., Jiang, L., Zhou, J., 2013. Differentially expressed genes of human microvascular endothelial cells in response to anti-dengue virus NS1 antibodies by suppression subtractive hybridization. Viral Immunol. 26, 185–191.

Yu, J.M., Li, J.S., Ao, Y.Y., Duan, Z.J., 2013. Detection of novel viruses in porcine fecal samples from China. Virol. J. 10, 39.

van der Heijden, M., de Vries, M., van Steenbeek, F.G., Favier, R.P., Deijs, M., Brinkhof, B., Rothuizen, J., van der Hoek, L., Penning, L.C., 2012. Sequence-independent VIDISCA-454 technique to discover new viruses in canine livers. J. Virol. Methods 185, 152–155.

van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J., Wolthers, K.C., Wertheim-van Dillen, P.M., Kaandorp, J., Spaargaren, J., Berkhout, B., 2004. Identification of a new human coronavirus. Nat. Med. 10, 368–373.

Yu, J.M., Li, L.L., Zhang, C.Y., Lu, S., Ao, Y.Y., Gao, H.C., Xie, Z.P., Xie, G.C., Sun, X.M., Pang, L.L., Xu, J.G., Lipkin, W.I., Duan, Z.J., 2016. A novel hepatovirus identified in wild woodchuck Marmota himalayana. Sci. Rep. 6, 22361.