Journal of Biology, Agriculture and Healthcare
ISSN 2224-3208 (Paper) ISSN 2225-093X (Online)
Vol.7, No.21, 2017
www.iiste.org

Review on Viral Metagenomics and Its Future Perspective in Zoonotic and Arboviral Disease Surveillance

Begna Bulcha[1] Gebreyohannes Berhane[1] Olana Merera[2]
1.Addis Ababa University college of Veterinary Medicine, Bishoftu, Oromia,Ethiopia
2.Samara University College of Veterinary Medicine, Samara, Afar, Ethiopia

SUMMARY
Viral metagenomics or full-length genome sequencing of enriched viral particle preparations has been frequently applied for viral discovery resulting in the genetic characterization of numerous human and animal viruses. The non- specific nature of viral metagenomics endows it with great potential as a unversal virus detection assay. Historical methods includes, electron microscopy, cell culture, inoculation in suckling mice and serology. Many viruses cannot be cultivated, excluding the use of cell line isolation and serologic techniques, and can only be characterized by molecular methods. The molecular-based techniques provide sensitive and rapid means of virus detection and identification. Most of these tests are designed to be virus-specific that makes them unsuitable for detection of unexpected and/or completely new viruses, as well as novel viral variants that poses great challenge due their re-emrgince property. The recently developed approaches of viral metagenomics provide an effective novel way to screen samples and detect viruses without previous knowlege of the infectious agent, thereby enabling a better diagnosis and disease control. The basic steps involved in viral metagenomics include preparation of viral nucleic acid that is free from host and contaminating nuceic acids, sequence-independent amplification of viral nucleic acid,sequencing and use of bioinformatics tools for analysis of sequnce generated. Viral metagenomics aims to provide the genetic composition of the complete viral population of a sample in an unbiased and culture-independent manner. Viral metagenomics has been successfully used to investigate viral populations in different environments such as seawater, gastreointestinal tracts, and respiratory samples and have demonstrated that there is a high diversity among viruses. Many potential emerging viruses of concern might already be infecting humans, domestic animals or wlidlife but awailt their detection by disease surveilla-nce which can be possible through application of viral metagenomics. This review aims to de-scribes the different possible steps of a viral metagenomics and its future application in viral zonootic and arboviral disease surveillance.
Keywords: Arboviral disease, Bioinformatics, Emerging infectious disease, Sequencing, Surveillance, Viral metagenomics.

1. INTRODUCTION

Viral infections are a major global health concern, and new infectious diseases continue to emerge from year to year (Dong *et al*., 2008). Due to intensive globalization, climatic changes, and viral evolution, among other factors, contributes for the emergence of viruses and new viral diseases in the last decades. Animal pathogens, in particular, viruses are considered to be a significant source of emerging human infections (Cleaveland *et al*., 2001). The identification and optimal characterization of novel viruses affecting both domestic and wild animal population is central to protecting both human and animal health. In this situation, it is crucial to apply powerful methods for broad-range detection and identification of the emerging viruses. According to a recent statistical estimate, there are at least 320,000 mammalian viruses that are yet discovered (Anthony *et al*., 2013).

Viruses can be identified by a wide range of techniques, which are mainly based on comparisons with known viruses. Historical methods of virus identification include electron microscopy, cell culture, inoculation in suckling mice and serology. The major limitation of the historic viral identification is that, it doesn't identify viruses that cannot be cultivated which are only characterized by molecular methods. In combination with classical methods, the molecular-based techniques provide sensitive and rapid means of virus detection and identification. However, most of these diagnostic tests are designed to be virus-specific or aimed at a limited group of infectious agents. This makes them unsuitable for the detection of unexpected and/or completely new viruses, as well as novel viral variants (Belak *et al.,* 2009).

Viral metagenomics is the recent and advanced science used to sequence all viral genomes in a given sample without previous knowledge about their nature. This is done by experiments in which all nucleic acids in a certain sample are sequenced. Next-Generation Sequencing (NGS) techniques make it possible this advanced science to the discovery of completely new viral species (Luisa, 2011). For bacteria, historically, the diversity of a sample used to be expressed by performing phylogenetic analyses based on 16S ribosomal RNA **(Handelsman, 2004**). However, since viruses lack such a universally conserved motif, viral metagenomics refers to the attempt to recover full and partial genomes of all viruses present in the sample (Nieuwenhuijse and Koopmans, 2017*)*.

Viral metagenomics analysis protocols generally start with procedures to remove host and bacterial cells followed by purification of viral nucleic acids, then sequence independent amplification of viral nucleic acid, sequencing and finally use of bioinformatics tools for analysis of sequences generated. Viral metagenomics

circumvents limitations associated with other virus characterization methods, such as PCR or microarrays, and the power of this method resides in targeting total viral nucleic acids without the need for a prior knowledge of the viral types present in a sample (Edwards and Rohwer, 2005).

Viral metagenomics or sequencing of nucleic acids from enriched viral particle preparations has been frequently applied for viral discovery (Chiu *et al.,* 2013), resulting in the genetic characterization of numerous human and animal viruses (De Vlaminck *et al.,* 2013), particularly its non-specific nature endows it with potential as a universal virus detection assay. This procedure aims to provide the genetic composition of the complete viral populations of a sample in an unbiased and culture-independent manner and it has been successfully used to investigate viral populations in different environments such as seawater (Breitbart *et al.,* 2002), gastrointestinal tracts (Victoria *et al.,* 2009), and respiratory samples (Allander *et al.,* 2007) and have demonstrated that there is a high diversity among viruses, and that there is a vast number of viruses that are yet to be discovered. Surveillance and monitoring of viral pathogens circulating in humans and wildlife and the identification of emerging infectious disease (EID) at an early stage is challenging. Many potential emerging viruses of concern might already be infecting humans or animals but await their detection by disease surveillancewhich can be possible through viral metagenomics (Temmam *et al.,* 2014). Emerging infectious diseases are a tremendous burden on economies and public health, and because many cases arise with no known etiology, there is a high demand for advances in viral diagnostic methods (Jones *et al.,* 2008).

Therefore the objectives of this seminar paper are:

- ✓ To review available literature on viral metagenomics.
- ✓ To indicate the anticipated application of viral metagenomics in viral zoonotic and arboviral disease surveillance.

2. HIGHLIGHT OF VIRAL IDENTIFICATION TECHNIQUES

Classically, virus discovery from clinical samples was based on filtration (to remove host cells and other larger microbes), inoculation of the cell free filtrate in suitable cell cultures followed by purification of the viruses from cultures and their characterization (Lipkin and Firth, 2013). Morphological changes in the cultured cells, collectively known as cytopathic effect, such as formation of syncytia, cell rounding, lysis, detachment, or inclusion bodies, indicate the presence and successful infection of the virus (es) in the cells. Virus isolate(s) are purified from the cultured cells or culture supernatant using density gradient and other high speed centrifugal techniques. This is followed by structural characterization of viral particles, antigens, nucleic acids, through different biophysical and biochemical methods (Dong *et al.,* 2008). Although classical methods are sometimes considered as time-consuming, tedious and need significant experimental basis, but the cell inoculation method still remains an exceptional source of enriched viral particles required for serological, molecular characterization and other purposes (Neill *et al.,* 2014).

Clinical virology field shows a gradual substitution of the traditional virus discovery methods with novel molecular biology technology. With the progresses of molecular biology, polymerase chain reaction (PCR)-based methods became the main techniques for virus discovery and allowed the detection of uncultivable viruses, but these techniques required prior knowledge of closely related viral genomes (Fernando, 2012). A great challenge With respect to molecular tools is that, viruses lack a universally conserved genetic marker to target, and PCR assays directed towards conserved sequences within viral groups can only identify close variants of those groups (Staheli *et al*., 2011). Sequence dependent and sequence independent methods are being used for the molecular detection of viruses. Sequence dependent methods, including PCR using consensus primers and hybridization methods such as microarray, require the knowledge of the nucleic acids for the detection of novel viruses (Fawaz *et al*., 2014). Viral metagenomics provides superior capability to detect known and unknown viruses than the traditional and molecular sequence-dependent and sequence-independent methods (Pozzetto, 2002).

3. VIRAL METAGENOMICS

Metagenomics as the name implies is different from genomics in that it takes into account all the genomes of all the microorganisms present in the study environment whether it be a water body like sea, human body or the intestine of an animal at a community level (**Eisen**, 2007).The 16S rRNA gene of the metagenome is sequenced to ascertain the microbial diversity since this gene is highly conserved across species. However, since viruses lack such a universally conserved motif, viral metagenomics refers to the attempt to recover full and partial genomes of all viruses present in the sample (Simmonds, 2015). Viral metagenomics is a sequence, and culture-independent approach that has proven to be a valuable tool for the investigation not only of diseases of unknown etiology but also of the complete viral flora of different reservoirs and vectors. By providing insights into a wide range of unknown potential pathogens and revealing novel aspects of biodiversity, metagenomics is able to detect and characterize novel pathogens (Tang and Chiu, 2010).

Viral metagenomics methods have evolved significantly since they were first developed in 1998 (Handelsman *et al.,*1998). The process of sample preparation has since been streamlined and the sequencing speed increase with

the advent of high throughput sequencing technologies. The replacement of cloning with high throughput methods has revolutionized viral metagenomics. Compared to other viral discovery methods, it is less biased. Potentially, any viruses in the samples, culturable or unculturable, known or novel can be readily detected with the viral metagnomic approach (Mokili *et al.,* 2012). An excellent and very important example for the practical applicability of viral metagenomics is the recent detection of Schmallenberg virus, a novel orthobunyavirus in Europe, with large epizootiological importance (Hoffmann et *al.,* 2012).

The basic steps involved in viral metagenomics include preparation of viral nucleic acid that is free from host and contaminant nucleic acids, sequence independent amplification of viral nucleic acid, sequencing and finally use of bioinformatics tools for analysis of sequences produced (Alavandi and Poornima, 2012).

3.1. Sample preparation for viral metagenomics

New generation sequencing (NGS) metagenomics has emerged as the most promising tool for the detection and discovery of novel infectious agents in clinical sample (Barzon *et al.,* 2011). However, being unbiased method of sequencing, NGS metagenomics is greatly affected by very low virus-to-host genome ratios in clinical samples (He *et al.,* 2013).

As obligate intracellular organisms, viral preparations are usually heavily contaminated by host nucleic acid and it is wise to remove this is as practical in order to ensure as many of the resulting sequence readings of viral rather than host origin. Hence, enrichment of pathogen genetic material or depletion of host genetic materials is essential to maximize sensitivity for discovery of novel viruses in clinical samples (Whon *et al.,* 2012).

3.1.1. Isolation of viral particles

Several approaches have been developed for viral particle concentration includes various size selection filtrations, gradients, differential ultracentrifugation, and chemical and enzymatic pretreatments (Tatiana *et al.,* 2012). These steps serve to concentrate viral particles for more efficient nucleic acid extraction and also help to remove contamination by no-viral cells, maximizing the amount of viral sequence obtained. A fast, simple, and reliable high-yielding method for viral particle recovery is tissue homogenization and cell disruption by freezing and thawing followed by filtering the samples through 0.22µm and 0.45µm pore-size discs **(**Figure 1). Most viruses are smaller than bacteria and filtration using a 0.22µm filter is a common procedure for removing bacteria when searching for viruses (Thurber *et al.,* 2009). However, in recent years, a number of viruses have been discovered that are as large as bacteria (Van Etten *et al.,* 2010) and these would be lost using this procedure. The viral particles can also be separated from other components by density centrifugation using sucrose or cesium chloride gradients (Thurber *et al.,* 2009).

3.1.2. Removal of non-viral nucleic acid

After homogenization of tissues, cells are disrupted by three freeze-thaw phases while leaving the nucleus intact. Nuclei are then pelleted by centrifugation and supernatants are treated by a cocktail of nucleases (RNase, DNase, and Benzonase) to remove cellular nucleic acids and non-particle protected viral nucleic acids (Daly *et al.,* 2011). The presence of the viral capsid surrounding the nucleic acid can be exploited by the use of nucleases. Nuclease treatment is an enrichment method for viral nucleic acid which can be applied directly to the filtered material or could be used as complement to ultracentrifugation. The mechanism for this effect is related to the relative sensitivity of free-host DNA to DNAse treatment and to the relative insensitivity of virus capsid-protected DNA to DNAse treatment. The capsid protects the viral genome while degrading all the external nucleic acids (Allander *et al.,* 2001).The application of these technique in a clinical setting will require that any virus enrichment methods are simple to perform, fast, robust, effective, standardized and do not require significant capital expenditure. The material extracted after nuclease treatment are normally much smaller that amplification is normally required before sequencing (Hall *et al.,* 2014).

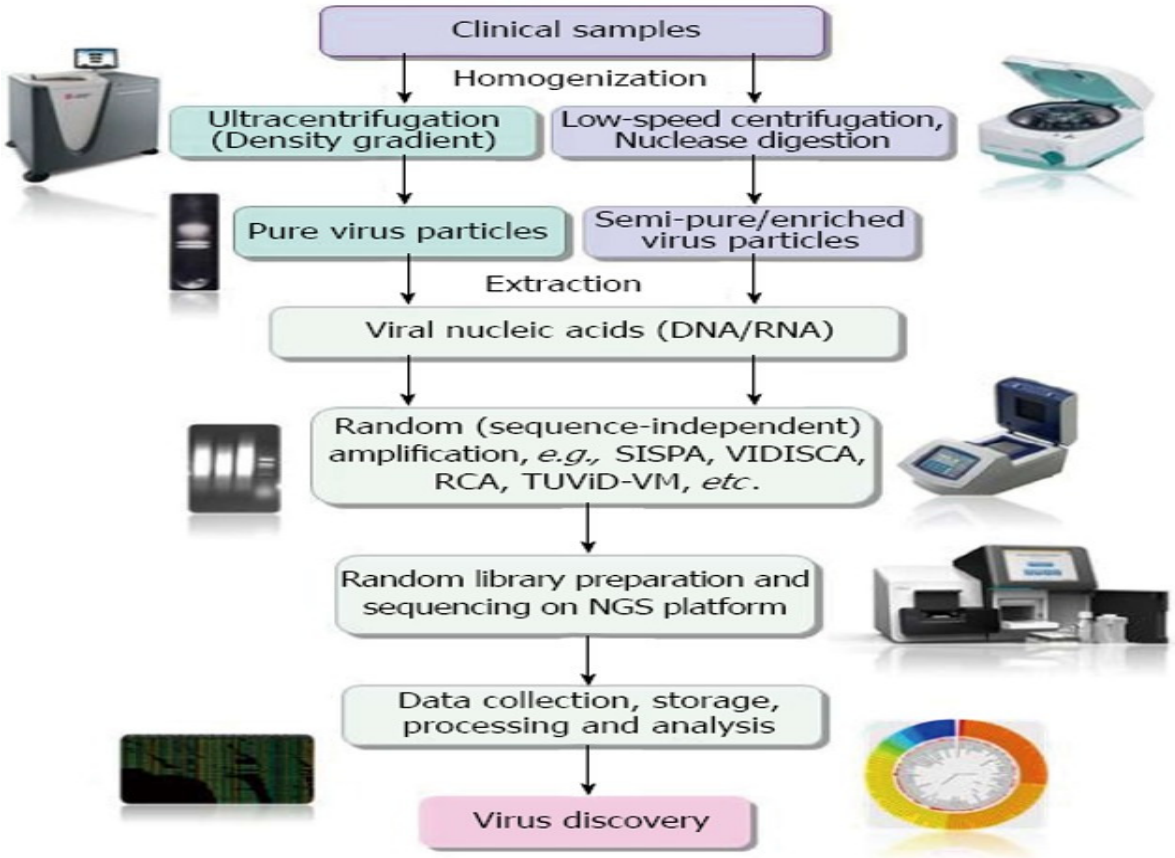

Figure 1: A schematic representation of the different steps involved in NGS based virus metagenomics and discovery.
Source: (Datta *et al.,* 2015).

3.2. Amplification of nucleic acids

The amount of total nucleic acids isolated from viral particles is often too low for sequencing, and may require amplifying viral nucleic acid depending on the sequencing technology used. The viral genomes present in the nucleic acid sample are simultaneously amplified independent of their sequences (Ambrose and Clewley, 2006).

Consequently, after purification, viral nucleic acid needs to be amplified to generate sufficient amounts of DNA for most sequencing platforms. RNA viruses have to be reverse-transcribed before amplification. For viral metagenomics and virus discovery, viral genomes need to be amplified without prior viral sequence knowledge. Currently,different sequence-independent methods have been developedfor amplification purposes including; degenerate PCR, sequence independent single primer amplification (SISPA),degenerate oligonucleotide primed (DOP)-PCR, random PCR, and rolling circle amplification (RCA) (Bexfield and Kellam, 2011). Random PCR, SISPA, and RCA methods are more widely used in viral metagenomics (Alavandi and Poornima, 2012).

3.2.1. Sequence-independent single primer amplification (SISPA)

SISPA involves the partial cleavage of DNA by the endonuclease enzyme, followed by a directional ligation of an asymmetric adaptor to both ends of the DNA molecule. Djikeng *et al*. (2008), develops Reverse-priming (RP)-SISPA adapted from SISPA to generate whole genome shotgun libraries of virus communities. In RP-SISPA, which is a combination of SISPA and random PCR, the cDNA is synthesized from RNA with a mixture containing two primers. These SISPA and RP-SISPA amplification methods are widely used to characterize viruses from tissue samples, clinical biopsies and for viral metagenome analyses (Rosseel *et al.,* 2012).

This strategy and modifications, such as DNase-SISPA and VIDISCA (van der Hoek *et al.,* 2004), have successfully been used to identify a number of viruses, such as new Anello- and Parvoviruses in humans (Jones *et al.,* 2005) and Bungowannah virus in pigs with porcine myocarditis syndrome (Kirkland *et al.,* 2007).

The advantages of sequence-independent amplification are simplicity and relative speed and the ability to identify and sequence hundreds of viruses simultaneously thereby allowing detection of new or previously unrecognized viruses that are highly divergent from already described ones **(Bodewes *et al.,* 2014).**

3.2.2. Random PCR

Another method of amplification technology is random PCR (rPCR), which requires neither the digestion of the DNA/cDNA nor the ligation of adaptors (Froussard, 1992). Random PCR for viral DNA and RNA library

constructions uses two different primers: a first primer with a defined sequence at its 5' end, followed by a degenerate hexamer or heptamer sequence at the 3' end to randomly prime DNA synthesis, and a second primer complementary to the 5' defined region of the first primer (Allander *et al.,* 2005). Random PCR is an established method for analyzing viromes, finding novel viruses and detecting the presence of known viruses (Wylie *et al.,* 2012). This strategy has been widely used for the investigation of viral metagenomes, and a number of novel viruses have been discovered using this strategy in combination with sequencing. The detection of Astrovirus in the brains of both humans and minks (Blomstrom *et al*., 2010), Bocavirus in human, and Bornaviruses in parrots with proventricular dilatation disease (Honkavuori *et al.,* 2008) are a few examples of viruses that have been discovered with rPCR in combination with sequencing.

3.2.3. Rolling circle amplification

One of the efficient amplification strategy is the use of random primers in combination with a displacement polymerase. The RCA method is an isothermal multiple displacement amplification (MDA) that uses phi29 DNA polymerase. This polymerase possesses several features, such as strand displacement activity, proof-reading activity and generation of very long synthesis products, which make it most suitable for the efficient amplification of circular DNA molecules from complex biological samples (Johne *et al.,* 2009).As multiply-primed RCA enables the detection of circular DNA viral genomes without the need of specific primers, the technique have become a powerful tool for the detection of unknown viruses. It employs random hexamer primers that bind to multiple sites on the virus DNA genome and is based on the strong strand displacement activity of the phi29 DNA polymerase. Viral DNA is exponentially amplified to generate micrograms of DNA (Erlandsson *et al.,* 2011). But phi29 DNA polymerase cannot amplify RNA or short fragments such as cDNA. To overcome this, the method of Whole Transcriptome Amplification (WTA) has been combined with MDA. It includes a ligation step before the amplification, resulting in cDNA that are linked and then amplified by phi29 DNA polymerase (Cheval *et al.,* 2011).

This strategy has been successful in studying different viral metagenomes, and several novel viruses haven been discovered by this method, such as a fibropapilloma virus in sea turtles, Anellovirus in harbor seals (Ng *et al.,* 2011), Bocavirus in pigs (Blomstrom *et al*., 2009), and Papillomaviruses in humans (Rector *et al.,* 2004).

3.3. Nucleic acid sequencing

3.3.1. Sanger sequencing method

To identify the viral nucleic acid in a sample, sequencing is often utilized (Breitbart and Rohwer, 2005). One approach is to construct viral shotgun libraries and sequence these by a standard sequencing technology such as Sanger sequencing (Sanger *et al.,* 1977). This approach creates high-quality sequence data and can nowadays produce sequence reads of up to almost 1000 nucleotide. But as this approach is highly laborious in comparison to its yield, the use of new high-throughput sequencing technologies is often replacing Sanger sequencing for metagenomic studies (Jason *et al.,* 2015). For viral metagenomics, this enables the detection of viruses with low copy numbers. Even though a variety of methods are used to reduce the host and other contaminating nucleic acids, a vast amount still remains and has the potential of masking the viral nucleic acids. Therefore, high-throughput sequencing is often required in the viral metagenomics screening although Sanger sequencing is still often used in the follow-up studies due to its capacity to produce longer sequencing reads (Kircher and Kelso, 2010).

3.3.2. High-throughput sequencing

Metagenomic sequencing technologies differ in library preparation methods and the length of reading frames produced. The next generation high-throughput pyrosequencing by 454 Life Sciences provide greater amounts of sequence data compared to Sanger's sequencing method, enabling detection of even low number of viruses present in the sample (Metzker, 2010) with no need of cloning. Even higher throughput technologies such as Solexa/Illumina and SOLiD systems are now available, which can provide data of as much as 3–6 giga base pair per run. However, the average read length obtained with these systems is about 50–100 base pair (Pareek *et al.,* 2011). The Illumina/Solexa and Roche 454 next generation sequencing platform are the most used platforms for viral metagenomics (Luo *et al.,* 2012).

The Illumina/Solexa method is based on sequencing by synthesis chemistry using fragments of the sample DNA ligated to oligonucleotide adapters. The adapters on a solid support act as primers for DNA polymerase to incorporate reversible terminator nucleotides each labeled with a different fluorescent dye (Mokili *et al.,* 2012). The current market leader, Illumina, manufactures instruments capable of generating billions of 150 base pair paired end reads per run, with read lengths of up to 300 base pair. The Illumina short read platform is widely used for analyses of viral genomes and metagenomes, and, given sufficient sequencing coverage, enables sensitive characterization of low frequency variation within viral populations (**Li *et al.,* 2014**).Novel bat viruses of the genera mamastro, boca, circo, ifla and orthohepadna in bats from Myanmar (He *et al.,* 2013), three novel group 1 coronaviruses from three North American bats (Donaldson *et al.,* 2010), are examples of viruses discovered using Illumina/ Solexa sequencing platform.

The 454 platform is based on pyrosequencing and while it produces the longest reads, 400 nucleotide,

compared to the others, its throughput is less (0.4 to 0.6 gigabase/run). For seque-ncing, DNA is fragmented and ligated to biotinylated specific linkers. The complex DNA/linkers fragment is attached to streptavidin-coated beads that anchor the DNA inside a droplet of water and PCR reagents in oil emulsion. Each fragment is first amplified to produce the template for sequencing reaction. Sequencing is carried out by annealing primers to the linker portion of the template complex, followed by the incorporation of nucleotides by DNA polymerase, which facilitates the extension of the complementary DNA. The pyrophosphate released by this process is measurable by the production of light (Meyer *et al*., 2008). The Roche 454 system measures the pyrophosphate released as the result of nucleotide incorporation during DNA synthesis mediated by DNA polymerase. The amount of light released is proportional to the intensity of the light signal captured by a charge-coupled device camera, which then converts light signals into digital data (Meyer,2007).

3.4. Bioinformatics

Bioinformatics analyzes of viral metagenomes attempts to answer three questions: how many viruses are there (diversity), 'what are they (taxonomy), and what are they doing (function)? (Willner, 2010). Bioinformatic tools compare the viral sequences generated with known sequences maintained in an annotated database such as National Center for Biotechnology Information(NCBI) using the programmes such as Basic Local Alignment Search Tool (BLASTn) and to assign taxonomy (by comparing with known viral sequences and indicate divergent ones) and function to metagenomic sequences (Blomstrom, 2011).

One of the challenges in viral metagenomics can be found in the analysis of the vast amount of sequencing data produced. Unlike re-sequencing of viral genomes with high-throughput sequencing by which it is possible to map the reads into an existing genome, the datasets from metagenomics studies are complicated by the fact that they contain a mixture of different species (Blomstrom, 2011). Also, the genomes in the datasets are usually incomplete with some cases wherein there are only a few numbers of short fragments belonging to each genome. Furthermore, some reads display a high divergence compared to sequences that are deposited in databases. For these reasons, a number of *de novo* assembly algorithms are being evaluated for their suitability to this specific task, and more are sure to be developed (Bao *et al.,* 2011).

Recently several programs and platforms have been developed that can help with both the analysis of the data as well as with the visualization of the sequencing results. Such programs includes, MEtaGenome Analyzer (MEGAN) (Huson *et al.,* 2009), PathSeq (Kostic *et al.,* 2011), Community Cyberinfrastructure for Advanced Microbial Ecology Research and Analysis (CAMERA) (Sun *et al.,* 2011), and Galaxy (Goecks *et al.,* 2010).

Table 1: Summary of a number of viral metagenomics studies and viral discoveries related to veterinary science.

Application	Sample type	Virus	Method	Reference(s)
Etiology- PMC **Etiology-proventricular Dilatation disease in parrot**	**Sera** **Brain**	**Bungowannah virus** **Bornavirus**	**SISPA** **Microarray** **rPCR and 454 sequencing**	**Kirkland *et al.*, 2007** **Kistler *et al.*, 2008** **Honkavuori *et al.*, 2008**
Etiology-fibropapillomatosisin sea turtle	**Fibropapilomas**	**Sea turtle Tornovirus 1**	**Phi29 and shotgun sequencing**	**Ng *et al.*, 2009**
Etiology-shaking mink syndrome	**Brain**	**Astrovirus**	**rPCR and 454 sequencing**	**Blostrom *et al.*, 2010**
Complex disease-PMWS in pigs	**Lymph nodes**	**torque teno virus, porcine Bocavirus**	**Phi29,454 sequencing**	**Blostrom *et al.*, 2009)**
Viral flora- honey bee colony collapse	**Bee**	**Chronic bee paralysis virus, sac brood virus, Israeli acute virus**	**rPCR,454 sequencing**	**Cox-Foster *et al.*, 2007**
Viral flora-bats	**Fecal** **Fecal and oral**	**Parvovirus, coronavirus, Circovirus** **Circovirus, flavivirus**	**rPCR,454 sequencing** **SISPA,454 seq.**	**Li *et al.*, 2010** **Donaldson *et al.*,2010**
Viral flora-monkeys	**Fecal** **Sera**	**chimpanzee stool-associated circular virus** **simian hemorrhagic fever virus**	**rPCR and Sanger sequencing** **rPCR, 454 sequencing**	**Blinkova *et al.*, 2010** **Lauck *et al.*, 2011**

4. FUTURE PERSPECTIVE IN METAGENOMICS BASED SURVEILLANCE PROGRAM

The future perspectives in virology appear that, the metagenomic approach will generate a plethora of genetic information from unknown and potentially infectious agents, some of which could be associated with human diseases. The discovery of viruses will start to precede the characterization of the diseases they cause, well before the pathogenicity of these agents is defined (Mokili *et al*., 2012).

Viruses make up over two-thirds of all new human pathogens, a highly significant overrepresentation given that most current human pathogen species are bacteria, fungi or helminthes (Woolhouse and Gaunt, 2007). There are 219 viral species (belonging to 23 families) that are known to infect humans, among which more than two-thirds are of zoonotic origin (Woolhouse *et al.,* 2012).

A high percentage of novel EIDs in humans are zoonoses caused by viruses of animal origin, and as humans expand their geographical range and come closer into contact with wildlife reservoirs, the likelihood of new diseases in domestic animals and humans increases (Jones *et al.,* 2008). Providing high-level epidemiological monitoring of viral diseases is undeniably a global public health ambition, and despite rapid progress in the development of diagnostic methods in recent years, improvements are needed for better cost, size (Frey *et al.,* 2014).

The recent development of viral metagenomics, *i.e.*, the characterization of the complete viral diversity isolated from an organism or an environment using high-throughput sequencing technologies, is promising for the surveillance of emerging viral zoonotic diseases and can be accomplished by analyzing the viromes of selected animals and arthropods that are closely in contact with humans (Temmam *et al.,* 2014).

4.1. Viral metagenomics in zoonotic and arboviral disease surveillance programs

Zoonotic and arboviral disease surveillance programs have recently integrated entomology and veterinary medicine. To prevent emerging infectious diseases in humans, surveillance programs now focus on the early detection of new or re-emerging infectious agents in hematophagous arthropods and wild or domestic animals, before viral adaptation to human hosts (Figure 2).Viral metagenomics are well-adapted tools for these surveillance programs because they allow the detection of all viral genomes in a given sample without previous knowledge of their nature (Temmam *et al*., 2014).

Because they are easy to sample, arthropods may be used as targets for emerging arbovirus disease surveillance. Recent metagenomics analyses focused on mosquito arthropods have demonstrated the richness of the mosquito virome, including viruses that reflect the nectar or blood meals (Junlen and Drosten, 2013). Because arboviruses are transmitted to vertebrate hosts via the saliva of arthropods, a simple way to determine if emerging viral pathogens may be transmitted to humans is to selectively analyze the virome of the salivary glands of the arthropod, even though dissection is difficult for extremely small arthropods. However, metagenomic studies targeting the entire body of the bloodsucking arthropod not only allow for the description of the viral flora within the arthropod, which highlight the emerging infectious agents or insect-specific viruses as tools for vector population control, but they also allow for the study of interactions between viral and bacterial communities that may result in viral interference (e.g, *Wolbachia* endosymbiont and Dengue virus interactions (Bian *et al.,* 2010).

Wild fauna are other appropriate target animals for emerging zoonoses surveillance. Because of the many restrictions on studying endangered wild animals (such as bats), non-invasive sampling procedures may be used such as collecting urine or feces. Moreover, humans are more frequently in contact with feces or urine of wild animals in their shared environment, rather than with tissues or blood, with the exception of the consumption of bushmeat. As a consequence, most metagenomic studies conducted on wildlife have involved the feces or urine of wild animals (Smith and Wang, 2013).

Recent studies searching for the reservoir of Middle-East Respiratory Syndrome-Coronavirus (MERS-CoV) have shown the potential role of camels in the transmission of MERS-CoV to humans. Camels are not the usual targets of zoonotic surveillance programs, but these recent examples highlight the interest of focusing future viral metagenomic studies on other animal species interacting with humans if one considers their ability to transmit human infectious agents by crossing the species barriers between animals and humans (Briese *et al.,* 2014).

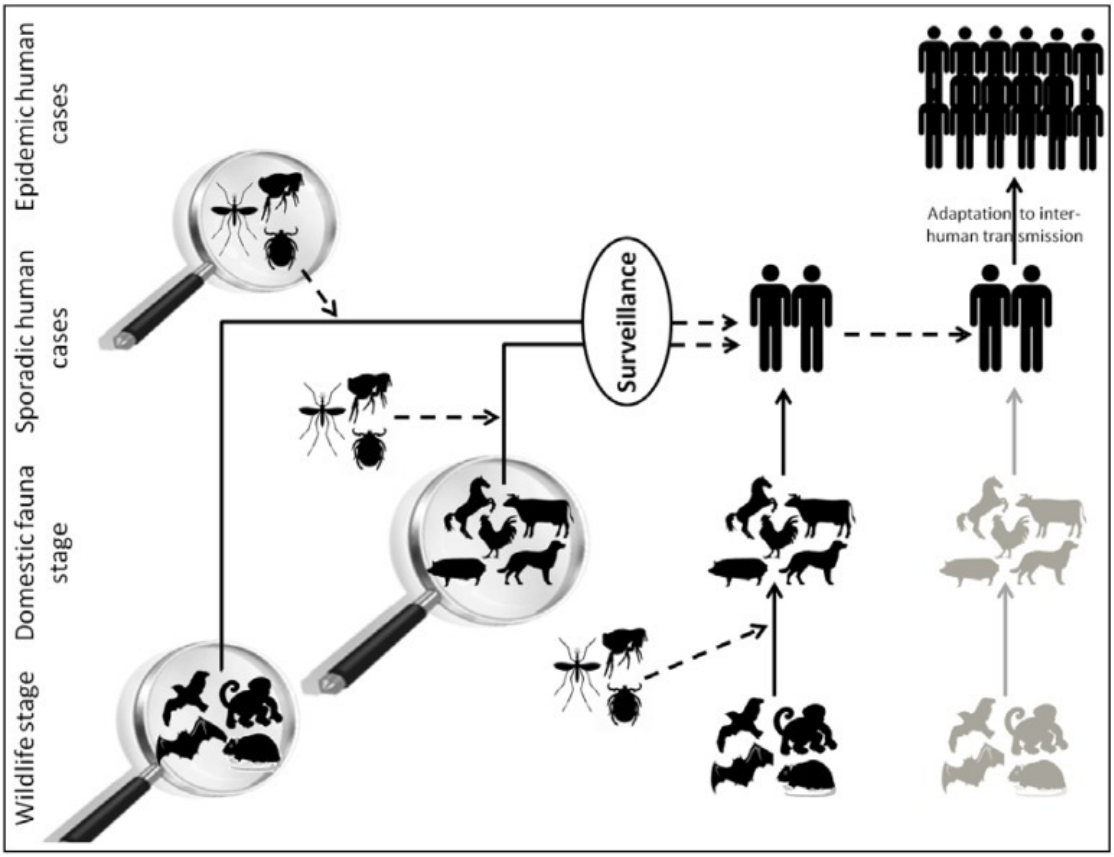

Figure 2: The origins of zoonotic human infections
Source: (Temmam *et al.,* 2014).

4.2. Viral metagenomics on blood-feeding arthropods as a tool for human and animal disease surveillance

Many potential emerging viruses of concern might already be infecting humans or wildlife but await their detection by disease surveillance. The problem is, in remote and under developed regions of the world, often no attention is paid towards possible infectious disease cases until a threshold of serious cases and deaths appears in a cluster and certain epidemic properties are reached (Carroll *et al.,* 2015). Metagenomic sequencing can be used as a promising solution for surveillance purposes as it detects all viruses in a single protocol, delivers additional genomic information for outbreak tracing, and detects novel unknown viruses. Surveillance and monitoring of viral pathogens circulating in humans and wildlife, together with the identification of EIDs, are critical for the prediction of future disease outbreaks and epidemics at an early stage. In this sense viral metagenomics is well suited in the detection and response to viral pathogen outbreaks (Rosario and Breitbart,2011).

An effective strategy in virus surveillance would need to survey simultaneously a wide range of viral types in a large number of human and wildlife individuals in order to detect viruses before spreading. In order to identify new EIDs before they emerge or re-emerge, wildlife animals that are likely to carry viruses with zoonotic potential, e.g., bats, rodents, birds and primates, are sampled frequently (Temmam *et al.,* 2014).

However, collecting swabs or blood from sufficient numbers of wildlife individuals and the subsequent identification of viruses is challenging. The solution for overcoming this challenge might be presented by the vector itself. Blood feeding arthropods feed on blood from a wide range of hosts including humans, animals and birds. In doing this, they act as syringes sampling numerous vertebrates and collecting the viral diversity over space, time and species (Figure 3) (Molaei *et al*., 2006). Recent development in NGS metagenomics introdu-c-es a term known as Xenosurveillance, which refers to the identification of viral pathogens from total nucleic acids extracted from mosquito blood meals, either by next-generation sequencing metagenomics or conventional PCR assays (Grubaugh *et al*., 2015).

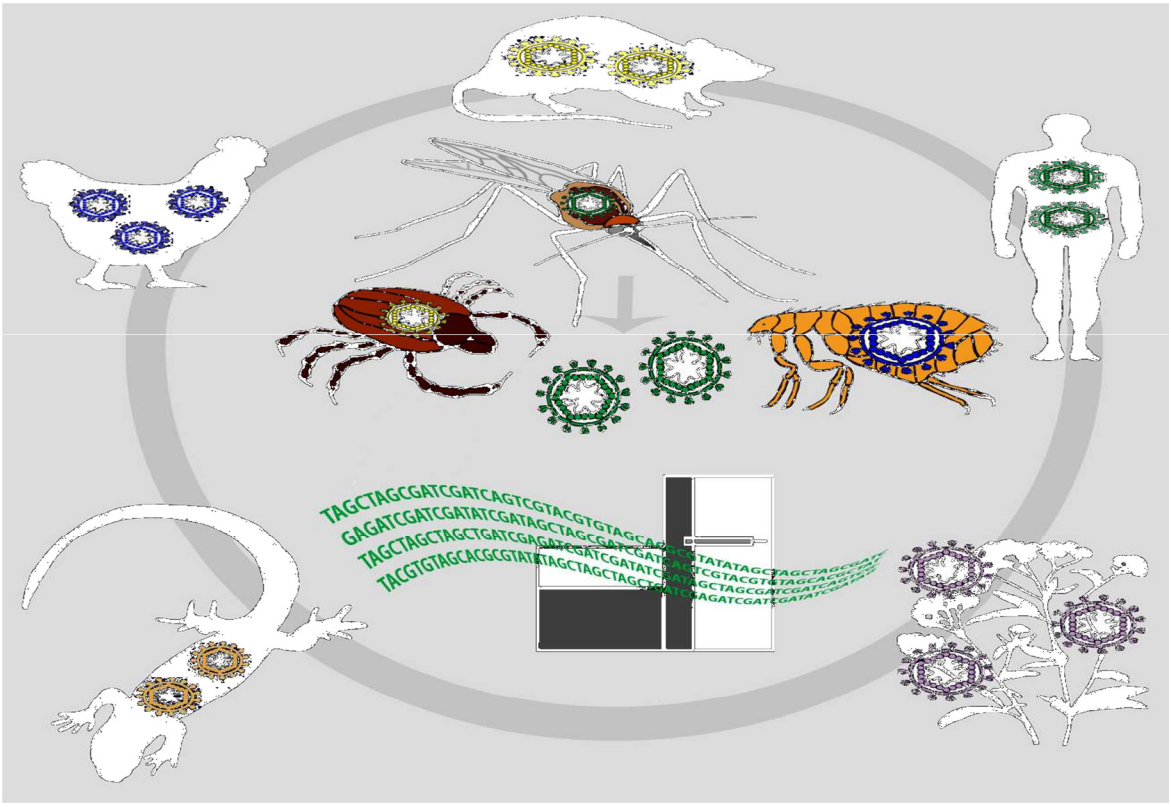

Figure 3: Vector-enabled metagenomics as a tool for virus surveillance.
Source: (Brinkmann *et al.,* 2016).

4.2.1. Mosquitoes as ''Flying Syringes'' for virus surveillance

Recently, done studies on mosquitoes by using metagenomics, demonstrates that mosquitoes could be used as biological syringes to accurately quantify viremias in animals (Kading *et al.,* 2015). In fact, two recent field studies have demonstrated that vertebrate viral pathogens that are not vector-borne could be detected in the bloodmeals of *Culicidae* mosquitoes (Barbazan *et al.,* 2015).

These findings suggest that hematophagous insects, specifically mosquitoes, could make sample acquisition and pathogen surveillance more tractable in remote tropical locales. Novel NGS technologies have led to the discovery of numerous mosquito-borne viruses replicate within the mosquito, but can also be transmitted biologically to vertebrates and infect vertebrate cells (Manzin *et al.,* 2013). The majority of mosquito-borne viruses belong to the families *Togaviridae*, *Flaviviridae* and *Bunyaviridae*, comprising highly pathogenic viruses such as Chikungunya virus, Dengue virus, Yellow fever virus, Japanese encephalitis virus, West Nile virus and Rift Valley fever virus (Zacks and Paessler, 2010).

Most studies on mosquito viruses focus on mosquito-infecting viruses or mosquito-borne viruses, neglecting the presence of vertebrate viruses originating from the blood of the mosquitoes' host that have accumulated in the mosquitoes' intestine during blood feeding. Example of such viruses in the metagenomic survey on viral abundance in mosquitoes (*Culex tritaeniorhynchus*, *Anopheles sinensis*, *Armigeres subalbatus* and *Culex fatigans*) shows 3.6% of all viruses found in the metagenomics survey were neither mosquito infecting nor mosquito-borne, but assumed to be of vertebrate origin (Shi *et al.,* 2015). These viruses were closely related to torque teno sus virus 1 (family *Anelloviridae*, genus Iotatorquevirus) which is widely distributed in pigs (Liu *et al.,* 2013). In addition, sequences belonging to the genus Parvovirus were identified that were closely related to porcine parvovirus. Since mosquitoes are not known to be vectors of torque teno sus virus 1 and porcine parvovirus, it is likely that the mosquito had ingested viremic blood during blood feeding on diseased pigs (Brinkmann *et al.,* 2016).

The combination of using mosquitoes as "flying syringes" and NGS for virus surveillance was recently introduced using the term vector enabled metagenomics. Metagenomics sequencing of mosquito samples from different sites in California showed that a broad range of already known and highly diverse DNA viruses, including anelloviruses, herpesviruses, poxviruses and Papillomaviruses (Ng *et al.,* 2011). These viruses infect a wide range of hosts including humans, mammals and birds and are not assumed to be transmissible by mosquito-es. Viruses of possible human origin were human papillomavirus 23, human herpes virus 1 and human Papillomavirus type 112. It is possible that both Papillomaviruses and herpes virus have been transferred from the human skin to the mosquito during feeding. Anelloviruses, infection of humans,vertbrates and marine mammals, also identified in blood meal of mosquitoes (Biagini *et al*.,2007).

5. CHALLENGES IN USING VIRAL METAGENOMICS

Metagenomics is a promising tool for the detection of new viral species that could potentially be a threat for human and animal health. However, it yet suffers several pitfalls when considering new/highly divergent viral genomes. A critical challenge in viral metagenomes assembly is the lack of a ubiquitous marker, analogous to bacterial 16S rRNA, to identify viral particles and estimate their diversity within ecological niches. Additionally, viral phylogeny based on sequences is impaired by extensive horizontal gene transfer and genome modularity within taxa, which is further complicated by the large numbers of viral particles within environmental samples. This makes it very difficult to find homologous sequences in reference database (Jorge *et al*., 2014).

One key aspect specific to whole metagenomic study strategies is requirement for ever greater amounts of input genomic material for comprehensive metagenomic studies (Petrosino *et al.*,2014). This is an important limitation when the starting material is limited, as in paleogeno-mics (Tringe and Rubin, 2005).

Another crucial issue is the process of DNA extraction. By definition, in a microbial community there are many different species and phylogenetic groups and as a consequence the DNA is encapsulated in cells with different properties. The techniques that are used to lyse cells might also affect the composition of environmental DNA libraries, as the harsh lysis methods that are necessary to extract DNA from every organism will cause degradation of the DNA from some organisms (Tringe and Rubin, 2005). Another challenging issues concerns, the gene reference catalogue. The main difficulty is that the gene reference catalogue needs to be representative enough for each studied sample. Finally, one of the biggest challenges that can be faced is the lack of an adapted statistical framework. The immense dimensionality of the data with millions of variables along with the very particular sparse nature of such data (due to the absence of species and thus genes among samples), make the use of classical statistics unsuitable (Rebecca *et al*., 2016)

Lastly, those developing countries are waiting for developed countries to make advances in science and technology that they later import at great cost has recently been challenged (Alain *et al*., 2011).

Table 2: Important bioinformatics challenges associated with application of next-generation sequencers in diagnosis of viruse by viral metagenomics and the proposed action to overcome the challenges.

Bioinformatics challenges associated with application of NGS in viral diagnostics	**Proposed action to overcome the challenges**
Generation of huge volumes of data by NGS platforms-"data deluge"	Advancement in storage and computation facilities, availability of computer with greater storage and highly powerful processors, cluster/grid computing and cloud computing. Computation facilities needs to be updated with emergence of newer platforms delivering larger datasets
Challenges in uploading data for submission to databases and supercomputing servers for analysis	Requirement of uninterrupted and extremely fast networks
Challenges in storage, public archival and ease of access	Creation of specialized data archive such as the Sequence Read Archive by National Instute of Health (NIH) and ENA (European nucleotide Archive) by Europian Bioinformatics Institute (EBI). Sharing of data within the two major databases (NIH, EBI and) for public accessibility
Challenges in analysis and visualization of large volumes of data, beyond the scope of computation facilities available in molecular biology laboratories	Creation of metagenomic or NGS data analysis pipelines and integrated tool kits, such as those available at NIH-NCBI, Genome Browser, and availability of cloud computing based servers such as Galaxy
Challenges in alignment, *de novo* assembly, gene prediction and phylogenetic analyses NGS datasets, especially short read datasets	Availability of alignment algorithms/programs such as Bowtie, Cloudburst, Zoom, Burrows-Wheeler Aligner Short Read Mapping Package(SHRiMP), Maximum Oligonucleotide Mapping (MOM), SeqMap, Metagene, Velvet, FragGeneScan, BLAST,Avadis, Eagle View
Interpretation of huge amount of data generated in metagenomic analyses by NGS platforms	Proper interpretation of analyzed data is of utmost importance to identify newer pathogens as well as their clinical significance

Source: (Datta *et al*., 2015).

6. CONCLUSION AND RECOMMENDATION

Viruses are abundant biological entities on earth and the emergence of viral pathogens has become a serious threat

to animals and humans worldwide Over 60% of these emerging pathogens are zoonotic in origin. The knowledge on the diversity of viruses in healthy and disease situations becomes important for understanding their role on the health of human and animal species. Viral metagenomics has proven to be useful for understanding viral diversity, describing novel viruses in new diseases and recognized as an important tool for discovering novel viruses in human and veterinary medicine. The techniques of viral metagenomics open novel possibilities for the direct comparative analysis of the genetic compositions of various clinical samples and for the detection of new, emerging viruses. Achievements obtained by viral metagenomics show significant advantages over traditional methods of identifying a viral pathogen, including no need of sequence information for that pathogen, identifying multiple pathogens in a single assay and eliminating the need for time-consuming culturing or antibody laboratory tests. The future of the field is promising, with emerging technologies showing potential to eliminate certain challenges and this is of paramount in emerging and re-emerging viral disease surveillance.

Based on the above conclusion, the following recommendations are pinpointed:

- ✓ Monitoring and surveillance of the viral diversity of wildlife and humans in today's rapidly changing ecosystems can be the key to predicting EIDs before they spread.
- ✓ The present high demand for advances in viral diagnostic methods, should be fulfilled through revolutionized viral metageniomics.
- ✓ Viral meta-genomics is an area which has opened the black box of viruses, so the technology should be devolved among developing country like Ethiopia and be taught to students and young scientists.
- ✓ There should be further developments in virus-specific nucleic acid extraction methods, bioinformatics data processing applications, and unifying data visualization tools that are needed to gain insightful surveillance knowledge from suspect samples.
- ✓ As the detection and characterization of novel viruses are of paramount importance in the forecasting of future outbreaks of viral diseases in humans and animals, the advancement of viral metagenomics should be encouraged.
- ✓ There should be high motivation, creativity and expertise increment to handle some bottleneck of viral metagenomics.

7. REFERENCES

Alain, K., Calka, N. and Angiuoli, S. (2011): DNA extractions from deep subseafloor sedime-nts: novel cryogenic-mill-based procedure and comparison to existing protocols. *J. Microbiol. Methods*, 87: 355-362.

Alavandi, S. and Poornima, M. (2012): Viral Metagenomics, A tool for virus discovery and diversity in aquaculture. *Indian J. Virol.*, 23:88-98.

Allander, T., Andreasson, K., Gupta, S., Bjerkner, A., Bogdanovic,G., Persson, M., Dalianis, T., Ramqvist, T. and Andersson, B.(2007): Identification of a third human polyomav-irus. *J. Virol.*, 81: 4130-4136.

Allander, T., Emerson, S., Engle, R., Purcell, R. and Bukh, J. (2001): A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proc. Nat. Acad Sci.*, 98: 11609-11614.

Allander, T., Tammi, M., Eriksson, M., Bjerkner, A., Tiveljung, L. and Andersson, B. (2005): Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proc. Nat. Acad Sci*., 102: 12891-12896.

Ambrose, H. E. and Clewley, J. P. (2006): Virus discovery by sequence-independent genome amplification. *Rev. Med. Virol.,* 16: 365-383.

Anthony, S., Epstein, J., Murray, K., Navarrete-Macias, I., Zambrana-Torrelio, C., Solovyov, A., Ojeda-Flores, R., Arrigo, N., Islam, A. and Ali Khan, S. (2013): A strategy to estimate unknown viral diversity in mammals. *Micro Bio.*, 4: 508-513

Bao, S., Jiang, R., Kwan, W., Wang, B., Ma, X. and Song, Y. (2011): Evaluation of next-generation sequencing software in mapping and assembly. *J. Hum. Genet.*, 56:406-414.

Barbazan, P., Thitithanyanont, A., Misse, D., Dubot, A. and Bosc, P. (2015) Detection of H5N1 avian influ-enza virus from mosquitoes collected in an infected poultry farm in Thailand. *Vector Borne Zoonotic Dis.*, 8: 105-109.

Barzon, L., Lavezzo, E., Militello, V., Toppo, S. and Palù, G. (2011): Applications of next-generation sequencing technologies to diagnostic virology. *Int. J. Mol. Sci.*, 12: 7861-7884

Belak, S., Thoren, P., LeBlanc, N. and Viljoen, G. (2009): Advances in viral disease diagnostic and molecular epidemiological technologies. *Expert. Rev. Mol. Diagn.*, 9: 367-381.

Bexfield, N. and Kellam, P. (2011): Metagenomics and the molecular identification of novel viruses. *Vet. J.*, 190: 191-198.

Biagini, P., Uch, R., Belhouchet, M., Attoui, H., Cantaloube, J., Brisbarre, N. and de Micco, P. (2007): Circular genomes related to anelloviruses identified in human and animal samples by using a combined rolling-circle amplification/sequence-independent single primer amplification approach. *J. Gen. Virol.*, 88: 2696-2701.

Bian, G., Xu, Y., Lu, P., Xie, Y. and Xi, Z. (2010): The endosymbiotic bacterium *Wolbachia* induces resistance to

dengue virus in *Aedes aegypti*. *PLoS Pathog*., *6*:e833.
Blinkova, O., Victoria, J., Ndjango, J., Peeters, M., Travis, D., Lonsdorf, E.and Wilson, M. (2010): Novel circular DNA viruses in stool samples of wildliving chimpanzees. *J. Gen. Virol.,* 91: 74-86.
Blomstrom, A. (2011): Viral metagenomics as an emerging and powerful tool in veterinary medicine, *Veterinary Quarterly*, 31: 107-114
Blomstrom, A., Belak, S., Fossum, C., McKillen, J., Allan, G., Wallgren, P. and Berg M. (2009): Detection of a novel porcine boca-like virus in the background of porcine Circovirus type 2 induced Postweaning multisystemic wasting syndrome. *Virol. Res.*, 146: 125-129.
Blomstrom, A., Widen, F., Hammer, A., Belak, S. and Berg, M. (2010): Detection of a novel astrovirus in brain tissue of mink suffering from shaking mink syndrome by use of viral metagenomics. *J. Clin. Microbiol*,, 48: 4392-4396.
Bodewes, R., Lempp, C., Schurch, A., Habierski, A., Hahn, K. and Lamers, M. (2014): Novel divergent nidovirus in a python with pneumonia. *J. Gen. Virol*. 95:2480-2485.
Breitbart, M. and Rohwer, F. (2005): Methods for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. *Biotechniques*, 39: 729-736.
Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J., Segall, A., Mead, D., Azam, F. and Rohwer, F. (2002): Genomic analysis of uncultured marine viral communities. *Proc. Nat. Acad. Sci.*, 99: 14250-14255.
Briese, T., Mishra, N., Jain, K., Zalmout, I., Jabado, O.J., Karesh, W., Daszak, P. Mohammed, O., Alagaili, A. and Lipkin, W. (2014): Middle east respiratory syndrome coronavirus quasis-pecies that include homologues of human isolates revealed through whole-genome analysis and virus cultured from dromedary camels in Saudi Arabia. *Micro Bio.*, *5*: e114614.
Brinkmann, A., Nitsche, A. and Kohl, C. (2016): Viral Metagenomics on Blood-Feeding Arthropods as a Tool for Human Disease Surveillance. *Int. J. Mol. Sci.*, 17: 1743-1781.
Carroll, M., Matthews, D., Hiscox, J., Elmore, M., Pollakis, G., Rambaut, A., Hewson, R., Garcia-Dorival, I., Bore, J. and Koundouno, R. (2015): Temporal and spatial analysis of the Ebolavirus outbreak in West Africa. *Nature*, 524: 97-101.
Cheval, J., Sauvage, V., Frangeul, L., Dacheux, L., Guigon, G., Dumey, N., Pariente, K. and Rousseaux, C. (2011): Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples. *J. Clin. Microbiol.,* 49: 32-68.
Chiu, C, Lipkin, W. and Firth, C. (2013): Viral pathogen discovery. *Current opinion in microbiology,* 16:468-78.
Cleaveland, S., Laurenson, M. and Taylor, L. (2001): Diseases of humans and their domestic mammals: Pathogen characteristics, host range and the risk of emergence. *Biol. Sci.*, 356: 991-999.
Cox-Foster, D., Conlan, S., Holmes, E., Palacios, G., Evans, J., Moran, N., Quan, P., Briese, T., Hornig, M. and Geiser, D. (2007): A metagenomic survey of microbes in honey bee colony collapse disorder. *Science*, 318: 283-287.
Daly, G., Bexfield, N., Heaney, J., Stubbs, S., Mayer, A., Palser, A., Kellam, P., Drou, N., Caccamo, M. and Tiley, L. (2011): A viral discovery methodology for clinical biopsy samples utilizing massively parallel next generation sequencing. *PLoS One*, 6:e28879
Datta, S., Budhauliya, R., Das, B., Chatterjee, S., VijayVeer, V. (2015): Next-generation sequencing in clinical virology: Discovery of new viruses.*World J. Virol.*, 4: 265-276
De Vlaminck, I., Khush, K., Strehl, C., Kohli, B., Luikart, H., Neff, N., Okamoto, J., Snyder, T., Cornfield, D., Nicolls, M., Weill, D., Bernstein, D., Valantine, H. and Quake, S. (2013): Temporal response of the human virome to immunosuppression and antiviral therapy. *Mol. Cell*, 155:1178-87.
Delwart, EL. (2007): Viral metagenomics. *Rev. Med. Virol.*, 17: 115-131.
Djikeng, A., Halpin, R., Kuzmickas, R., Depasse, J., Feldblyum, J., Sengamalay, N., Afonso, C., Zhang, X., Anderson, N., Ghedin, E. and Spiro, D. (2008): Viral genome sequen-cing by random priming methods. *BMC Genomics*, 95: 467-471.
Donaldson, E., Haskew, A., Gates, J., Huynh, J, Moore, C., Frieman, M. (2010): Metagenom-ic analysis of the viromes of three North American bat species: viral diversity among different bat species that share a common habitat. *J. Virol*., 84: 13004-13018.
Dong, J., Olano, J., McBride, J. and Walker, D. (2008): Emerging pathogens: challenges and successes of molecular diagnostics. *J. Mol. Diagn.,* 10: 185-197.
Edwards, R. and., Rohwer, F. (2005): Viral metagenomics. *Nat. Rev.Microbiol.*, 3: 504-5 10.
Eisen, J. A. (2007):Environmental Shotgun Sequencing: Its Potential and Challenges for Studying the Hidden World of Microbes. *PLoS Biology*, 5: e82.
Erlandsson, L., Rosenstierne, M., McLoughlin, K., Jaing, C. and Fomsgaard, A. (2011): The Microbial Detection Array Combined with Random Phi29-Amplification Used as a Diagnostic Tool for Virus. *PLoS One*, 6: e22631.
Fawaz, M., Ashok, A., Mishra, A. and Kambla, N. (2014): Viral Metagenomics: *A Rev.Art.*, 2: 81-89.

Fernando, C. (2012): Application of Molecular Diagnostic Techniques for Viral Testing.*Virol. J.*, 6: 104-114.
Frey, K., Herrera-Galeano, J. and Redden, C. (2014): Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood, *BMC Genomics*, 15: 96.
Froussard, P. (1992): A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Res.*, 20: 2900
Goecks, J., Nekrutenko, A. and Taylor, J. (2010): Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *PLoS One,*11:e86.
Grubaugh, N., Sharma, S., Krajacich, B., Fakoli, L., Bolay, F., Diclaro, J., Johnson, W., Ebel, G., Foy, B. and Brackney, D. (2015): Xenosurveillance: A novel mosquito-based approach for examining the human pathogen landscape. *PLoS Negl. Trop. Dis.,* 4: e0003628.
Hall, R., Wang, J., Todd, A., Bissielo, A., Yen, S., Strydom, H.and Peacey, M. (2014): Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *J. Virol. Meth.,* 195: 194-204.
Handelsman, J. (2004): Metagenomics: Application of genomics to uncultured microorganism. *Microbiol. Mol. Biol. Rev.*, 68: 669-685.
Handelsman, J., Rondon, R., Brady, F., Clardy, J., and Goodman, M. (1998): Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chem. Biol*. 5:245-249.
He, B., Li, Z., Yang, F., Zheng, J., Feng, Y., Guo, H., Li, Y., Wang, Y., Su, N. and Zhang, F. (2013): Virome profiling of bats from Myanmar by metagenomic analysis of tissue samples reveals more novel Mammalian viruses. *PLoS One,* 8: e61950.
Hoffmann, B., Scheuch, M., Hoper, D., Jungblut, R., Holsteg, M., Fischer, M., Breithaupt, A. and Mettenleiter, T. (2012): Novel Orthobunyavirus in Cattle, Europe. *Emerg. Infect. Dis.,* 18: 469-472.
Honkavuori, K., Shivaprasad, T., Williams, B., Quan, P., Hornig, M., Street, C., Palacios ,G., Hutchison, S., Franca, M.and Egholm, M. (2008): Novel borna virus in psittacinebirds with proventricular dilatation disease. *Emerg. Infect. Dis.*, 14: 1883-1886.
Huson, D., Richter, D., Mitra, S., Auch, A. and Schuster, S. (2009): Methods for comparative metagenomics. *BMC Bioinfo.*, 10: 1-12.
Jason, A., Damek, S. and Michael, P. (2015): High throughput sequencing technologies.*Mol. Cell,* 58: 586-597.
Johne, R., Hermann, M., Rector, A., Ranst, M. and Stevens, H. (2009): Rolling-circle amplification of viral DNA genomes using phi29 polymerase. *Tre. micro. Art*., 5: 45-71.
Jones, K., Patel N., Levy M., Storeygard, A., Balk D., Gittleman J. and Daszak P. (2008): Global trends in emerging infectious diseases. *Nature*, 451: 990-993.
Jones, M., Kapoor, A., Lukashov, V., Simmonds, P., Hecht, F. and Delwart, E. (2005): New DNA viruses identified in patients with acute viral infection syndrome. *J. Virol.*, 79: 8230-8236.
Jorge, F., Vazquez, C., Rodrigo, Garcia L,V, Brocal, M. and Pignatelli, M.(2014):Comparison of different assembly and annotation tools on analysis of simulated viral metagenomic communities in the gut. *BMC Genomics*, 15: 37
Junlen, S. and Drosten, C. (2013): Virus discovery and recent insights into virus diversity in arthropods. *Curr. Opin. Microbiol.*, 16: 507–513.
Kading, R., Biggerstaff, B., Young G. and Komar, N. (2014): Mosquitoes used to draw blood for arbovirus viremia determinations in small vertebrates. *PLoS One,* 9: e99342.
Kircher, M. and Kelso, J. (2010): High throughput DNA sequencing concepts and limitations. *BioEssays*, 32: 524–536.
Kirkland, P., Frost, M., Finlaison, D., King, K., Ridpath, J. and Gu, X. (2007): Identification of a novel virus in pigs - Bungowannah virus: a possible new species of pestivirus. *Virol. Res.*, 129: 26-34.
Kistler, A., Gancz, A., Clubb, S., Lublin, A., Mechani, S.and Farnoushi, Y. (2008.): Recovery of divergent avian bornaviruses from cases of proventricular dilatation disease: identification of a candidate etiologic agent. *Virol. J.*, 5: 88.
Kostic, A., Ojesina, A., Pedamallu, C., Jung, J., Verhaak, R., Getz, G. and Meyerson, M. (2011): PathSeq: software to identify or discover microbes by deep sequencing of human tissue. *Nat. Biotechnol.*, 29: 393-396.
Lauck, M., Hyeroba, D., Tumukunde, A., Weny, G., Connor, D., Friedrich, T. and Goldberg, T. (2011): Novel, divergent simian hemorrhagic fever viruses in a wild ugandan red colobus monkey discovered using direct pyrosequencing. *PLoS One*, 6: e19056.
Li, J. Z. (2014): Comparison of Illumina and 454 Deep Sequencing in Participants Failing Raltegravir-Based Antiretroviral Therapy. *PLoS One*, 9: e90485.
Li, S., Chan W., Lai C., Tsai K. and Hsu C. (2011): UMARS: Un-M Appable Reads Solution. *BMC Bioinfo.*, 12:1-9.
Lipkin, W. (2010): Microbe hunting. *Microbiol. Mol. Biol. Rev.,* 74: 363-377.
Liu, S., Vijayendran, D., Bonning, B. (2011): Next generation sequencing technologies for insect virus discovery.

J. Virol., 3: 1849-1869.
Luisa, B. (2011): Applications of Next-Generation Sequencing Technologies to Diagnostic Virology. *Int. J. Mol. Sci.,* 12: 7861-7884
Luo, C., Tsementzi, D., Kyrpides, N., Read, T. and Konstantinidis, K. (2012): Direct Comparisons of Illumina vs. Roche 454 Sequencing Technologies on the Same Microbial Community DNA Sample. *PLoS One,* 7: e30087.
Manzin, A., Martina, B., Gould, E., Bagnarelli, P. and Sambri, V. (2013): Human arthropod-borne viral infections. *BioMed Res. Int*., 60: 85-100.
Metzker, M. (2010): Sequencing technologies - the next generation. *Nat. Rev. Genet.*, 11: 31-46.
Meyer, M., Stenzel, U. and Hofreiter, M. (2008): Parallel tagged sequencing on the 454 platform. *Nat. Protoc.*, 3: 267-278.
Meyer, P. (2007): The history of pyrosequencing. *Meth. Mol. Bio.,* 37: 1-14.
Mokili, J., Rohwer, F. and Dutilh, B. (2012): Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol.*, 2: 63-77.
Molaei, G., Andreadis, T., Armstrong, P., Anderson J., Vossbrinck, C. (2006): Host feeding patterns of *Culex* mosquitoes and West Nile virus transmission, Northeastern United States. *Emerg. Infect. Dis*., 12: 468-474.
Neill, J., Bayles, D. and Ridpath, F. (2014): Simultaneous rapid sequencing of multiple RNA virus genomes. *J. Virol. Methods*, 201: 68-72.
Ng, T., Manire, C., Borrowman, K., Langer, T., Ehrhart, L. and Breitbart, M. (2009): Disco-very of a novel single-stranded DNA virus from a sea turtle fibropapilloma by using viral metagenomics. *J. Virol.*, 83: 2500-2509.
Ng, T., Wheeler, E., Greig, D., Waltzek, T., Gulland, F. and Breitbart, M. (2011): Metagen-omic identification of a novel anellovirus in Pacific harbor seal (*Phoca vitulina richardsii*) lung samples and its detection in samples from multiple years. *J. Gen. Virol.,* 92: 1318-1323.
Nieuwenhuijse, D. and Koopmans, M. (2017): Metagenomic Sequencing for Surveillance of Food and Waterborne Viral Diseases. *Front. Microbiol*., 8: 230.
Pareek, C., Smoczynski, R. and Tretyn, A. (2011): Sequencing technologies and genome sequencing. *J. Appl. Genet.*, 52: 413-435.
Petrosino, J., Highlander, S., Ann, R., Richard, A. and Gibbs, V. (2014): Metagenomic Pyrosequencing and Microbial Identification. *Clin. Chem*., 55: 856-86.
Pozzetto, B. (2002): Methods of diagnosis in virology. Medical Virology; Mammette, A., (Ed). Presses Universities de Lyon: Lyon, France; pp. 187-210.
Rebecca, R., Bede, C., Avraam, T., David, L. R. and Mattia, P. (2016): Challenges in the analysis of viral metagenomes. *Virus Evolution*, 2: 134-157
Rector, A., Bossart, G., Ghim, S., Sundberg, P., Jenson, A. and Van Ranst, M. (2004): Char-acterization of a novel close to-root papillomavirus from a Florida manatee by using multiply primed rolling circle amplification. Papillomavirus type 1. *J. Virol.*, 78:12698-12702.
Rohwer, F. (2002): Genomic analysis of uncultured marine viral communities. *Proc. Nat. Acad Sci.*, 99: 14250-14255.
Rosario, K. and Breitbart, M. (2011): Exploring the viral world through metagenomics. *Curr.Opin.Virol.,* 1: 289-2 97
Rosseel, T., Scheuch, M., Hoper, D., De Regge, N., Caij, A., Vandenbussche, F. and Van Borm, S. (2012): DNase SISPA-next generation sequencing confirms Schmallenberg virus in Belgian field samples and identifies genetic variation in Europe. *PLoS One*, 7: e41967.
Sanger, F., Nicklen, S., Coulson, A. (1977): DNA sequencing with chain terminating inhibi-tors. *Proc. Nat. Acad Sci*., 74: 5463-5467.
Shi, C., Liu, Y., Hu, X., Xiong, J., Zhang, B. and Yuan, Z. (2015): A metagenomic survey of viral abundance and diversity in mosquitoes from Hubei province. *PLoS One*, 10: e0129845
Simmonds, P. (2015): Methods for Virus Classification and the Challenge of Incorporating Metagenomic Sequence Data. *Journal of General Virology*, 6: 1193-206.
Smith, I. and Wang, L. (2013): Bats and their virome: An important source of emerging viruses capable of infecting humans. *Curr.Opin. Virol.*, 3: 84-91.
Staheli, J., Boyce, R., Kovarik, D. and Rose, T. (2011): CODEHOP PCR and CODEHOP PCR primer design. *Methods Mol. Biol*., 67: 57-73.
Sun, Sh., Chen, J., Li, W., Altintas, I., Abel, Lin., Peltier, S. And Stocks, K. (2011): Comm-unity Cyberinfrastructure for Advanced Microbial Ecology Research and Analysis. *Nucleic Acids Research,* 39: 1338-1340.
Tang, P. and Chiu, C. (2010): Metagenomics for the discovery of novel human viruses. *Futur. Microbiol*., 5: 177-189.
Tatiana, D., Agnès, N. and Séverine, C. (2012): Next-generation sequencing of ultra-low copy samples'*Current Topics in virology*, 10: 456-537.

Temmam, S., Davoust, B., Berenger, J., Raoult, D. and Desnues, Ch. (2014): Viral Metagenomics on Animals as a Tool for the Detection of Zoonoses Prior to Human Infection. *Int. J. Mol. Sci.*, 15: 10377-10397.

Thurber, R., Haynes, M., Breitbart, M., Wegley, L. and Rohwer, F. (2009): Laboratory procedures to generate viral metagenomes. *Nat. Protoc.*, 4: 470-483.

Tringe, S. and Rubin E. (2005): Metagenomics: DNA sequencing of environmental samples. *Nat Rev Genet.*, 6: 805-814.

Van der Hoek, L., Pyrc, K., Jebbink, M., Vermeulen-Oost, W., Berkhout, R., Wolthers, K., Wertheim-van Dillen, P. and Kaandorp, J. (2004): Identification of a new human coronavirus. *Nat. Med.*, 10: 368-373.

Van Etten, J., Lane, L. and Dunigan, D. (2010): DNA viruses: the really big ones (giruses). *Annu. Rev. Microbiol.*, 64: 83-99.

Victoria, J., Kapoor, A., Li, L., Blinkova, O., Slikas, B., Wang, C., Naeem, A., Zaidi, S. and Delwart, E. (2009.): Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. *J. Virol.*, 83: 4642-4651.

Whon, T., Kim, M., Roh, S., Shin, N., Lee, H. and Bae, J. (2012): Metagenomic characteriz-ation of airborne viral DNA diversity in the near-surface atmosphere. *J .Virol.,* 86: 8221-8231.

Willner, DL. (2010): Viral metagenomics in host-associated systems. San Diego: San Diego state University. *Bioessays,* 33: 43-51

Woolhouse, M. and Gaunt, E. (2007): Ecological origins of novel human pathogens. *Crit. Rev. Microbiol.,* 33: 231-242.

Woolhouse, M., Scott, F., Hudson, Z., Howey, R. and Chase-Topping, M. (2012): Human viruses: Discovery and emergence. *Phil. Trans. R. Soc. B.*, 367: 2864-2871.

Wylie, K., Mihindukulasuriya, K., Sodergren, E., Weinstock, G. and Storch, G. (2012): Seq-uence analysis of the human virome. *Genome Res.,* 25: 1910-1920.

Zacks, M. and Paessler, S. (2010): Encephalitic alphaviruses. *Vet. Microbiol.*, 140: 281-286.