Rev. Sci. Tech. Off. Int. Epiz., 2016, **35** (1), 25-42

Novel technologies applied to the nucleotide sequencing and comparative sequence analysis of the genomes of infectious agents in veterinary medicine

F. Granberg [1, 3]*, Á. Bálint [2] & S. Belák [1, 3]

(1) Swedish University of Agricultural Sciences (SLU), Department of Biomedical Sciences and Veterinary Public Health, Box 7028, 750 07 Uppsala, Sweden
(2) National Food Chain Safety Office, Veterinary Diagnostic Directorate, Tábornok u. 2, H-1143 Budapest, Hungary
(3) World Organisation for Animal Health (OIE) Collaborating Centre for the Biotechnology-based Diagnosis of Infectious Diseases in Veterinary Medicine, Uppsala, Sweden
*Corresponding author: Fredrik.Granberg@slu.se

Summary

Next-generation sequencing (NGS), also referred to as deep, high-throughput or massively parallel sequencing, is a powerful new tool that can be used for the complex diagnosis and intensive monitoring of infectious disease in veterinary medicine. NGS technologies are also being increasingly used to study the aetiology, genomics, evolution and epidemiology of infectious disease, as well as host–pathogen interactions and other aspects of infection biology.

This review briefly summarises recent progress and achievements in this field by first introducing a range of novel techniques and then presenting examples of NGS applications in veterinary infection biology. Various work steps and processes for sampling and sample preparation, sequence analysis and comparative genomics, and improving the accuracy of genomic prediction are discussed, as are bioinformatics requirements. Examples of sequencing-based applications and comparative genomics in veterinary medicine are then provided. This review is based on novel references selected from the literature and on experiences of the World Organisation for Animal Health (OIE) Collaborating Centre for the Biotechnology-based Diagnosis of Infectious Diseases in Veterinary Medicine, Uppsala, Sweden.

Keywords

Bacteria – Comparative genomics – Diagnosis – High-throughput sequencing – Infectious agent – Infectious disease – Metagenomics – Next-generation sequencing – Veterinary medicine – Virus.

Introduction

In recent decades, the classic direct diagnostic approaches of microbiology, i.e. virus isolation or bacterial cultivation followed by identification of the causative biological agents, have been increasingly replaced by a range of novel molecular techniques. These latter methods can detect and identify infectious agents in clinical samples containing viral or bacterial nucleic acids or other components such as proteins. Various molecular diagnostics methods can be used to detect and analyse genomic nucleic acid sequences of infectious agents. These include nucleic acid hybridisation assays, various gel-based and real-time polymerase chain reaction (PCR) systems, isothermal amplification approaches, and solid and liquid phase microarrays. Novel diagnostic and comparative genome sequencing techniques are also available for infectious agents. These methods, either alone or in combination, have resulted in a steady

doi: http://dx.doi.org/10.20506/rst.35.1.2415

improvement in diagnostic capability and increasing research efforts into the infection biology of animal diseases.

Several examples of rapid developments in 'molecular diagnostics' (also termed 'biotechnology-based diagnostics') were previously reviewed by the World Organisation for Animal Health (OIE) Collaborating Centre for Biotechnology-based Diagnosis of Infectious Diseases in Veterinary Medicine in Uppsala, Sweden, with a special focus on the various methods of nucleic acid hybridisation and amplification (1, 2, 3). This review will briefly introduce and discuss novel technologies and concepts relevant to nucleotide sequencing and comparative sequence analysis of the genomes of infectious agents. Examples of several techniques with practical applications in infection biology (including diagnostics) will be given, especially those targeting viral infectious agents. Expert opinion on the diagnostic applicability of comparative nucleotide sequence analysis methods in veterinary medicine will also be provided.

Sequencing technologies

Methods for determining nucleotide sequences are commonly referred to as 'sequencing technologies'. Obtaining genetic sequence information first became possible with the introduction of electrophoretic methods for DNA sequencing in the 1970s (4, 5). Sanger sequencing, based on dideoxy chain-termination technology, quickly became the method of choice for DNA sequencing. Over the years, this method was refined and automated with the introduction of fluorophore labelling and multicapillary electrophoresis. Sanger sequencing has remained useful and is still considered the gold standard for many applications (6). PCR, a technique for the selective amplification of DNA sequences *in vitro*, was later developed (7, 8). Together, these techniques offer unprecedented possibilities for molecular diagnostics and the characterisation of infectious agents, and have thus had a profound impact on the study of infection biology.

Second- or next-generation sequencing

Sanger sequencing represented the 'first generation' of sequencing technologies; it was not surpassed until 2005, when 'second-generation' or next-generation sequencing (NGS) technologies became commercially available. This development was driven by a growing demand for faster and cheaper methods of sequencing the large genomes of mammals (9, 10). Thus, NGS has become synonymous with high-throughput sequencing (HTS). Although a number of different NGS platforms are available (Table I), each with a unique sequencing method, most share the same general strategy: clonal amplification of DNA templates on a solid support matrix, followed by sequencing via a cyclic process in massively parallel sequencing reactions (11, 12, 13).

Using this approach, NGS technologies can determine thousands or millions of sequences concurrently. However, the generated sequences, referred to as 'reads', are typically rather short, ranging in length between 50 and 700 base pairs (bp). Longer sequences, known as 'contigs', are commonly obtained by assembling overlapping reads from sequence data using various bioinformatics tools (see below).

Since the introduction of NGS platforms, continuous developments and improvements have resulted in updated instruments with increased capacity to generate sequence data. Consequently, there has been a corresponding dramatic decrease in the cost per base (14, 15). To suit the needs of most average-sized laboratories, less expensive bench-top instruments with reduced capacity but shorter run times have been developed (16, 17). In this way, NGS-based methodologies for identifying and characterising infectious agents have become much more accessible to veterinary diagnostic and basic research laboratories.

A system based on pyrosequencing technology developed by 454 Life Sciences was the first commercially available NGS platform on the market. With relatively long sequence reads, it quickly became the HTS technology of choice for the discovery and *de novo* genome assembly of novel microorganisms (18, 19, 20). However, despite technological improvements and the launch of a bench-top model, 454 sequencing systems were outcompeted by cheaper, more efficient alternatives, resulting in their discontinuation and the withdrawal of technical support from mid-2016. Illumina (formerly Solexa) is now recognised as the dominant HTS technology (21). Using bridge PCR and sequencing-by-synthesis, the Illumina product line now includes both high-capacity platforms (suitable for larger sequencing centres) and bench-top platforms (aimed at a wider range of laboratories and the clinical diagnostics market). Other current NGS platforms include SOLiD sequencers, based on oligonucleotide ligation and detection technology (11), as well as the latest versions of Ion Torrent sequencers that use semiconductor-based hydrogen ion detection to monitor nucleotide incorporation during complementary strand synthesis (22). These different technologies and their corresponding platforms are summarised in Table I.

Third-generation sequencing

More recently, 'third-generation' sequencing technologies have been introduced. These are characterised by long read lengths and the capability to determine the sequence of single molecules without prior amplification (23). Two

Table I
Current sequencing technologies and their main applications

Generation	Platform	Read length	Accuracy (single read % or bases > Q30)	Output range	Run time	Approx. cost per Gb (in US$)	Applications
1st	**Chain termination**						
	Sanger sequencing	400–900 bp	~99.9%	N/A	20 min to 3 h	2,500	Widely used for sequence-based research, but impractical and expensive for larger sequencing projects
2nd	**Illumina**						
	HiSeq X	2 × 150 bp	~90% > Q30	900–1,800 Gb	3 days	7–25	Designed for maximum throughput and low cost population-scale human WGS
	HiSeq 4000	2 × 150 bp	~75% > Q30	125–1,500 Gb	1–3.5 days	25–55	Designed for maximum throughput and low cost production-scale genomics
	NextSeq 500	2 × 150 bp	~75% > Q30	20–120 Gb	12–30 h	40–55	HTS for average laboratories with enough capacity for mammalian WGS. Also suitable for targeted re-sequencing, exome and transcriptome sequencing
	MiSeq	2 × 300 bp	~90% > Q30	0.3–15 Gb	5–55 h	100–260	WGS of small genomes, targeted sequencing and re-sequencing
	SOLiD						
	5500 xl Wildfire	2 × 50 bp	~99.9%	240 Gb	10 days	20–40	Genome re-sequencing, genotyping and transcriptome sequencing
	5500 xl	2 × 60 bp	~99.9%	95 Gb	6 days	100–120	
	Ion Torrent Technology						
	Ion Proton	100–200 bp	~99%	10–32 Gb	2–4 h	35–100	WGS, exome sequencing, transcriptome sequencing and targeted sequencing
	Ion S5	200–400 bp	~99%	0.6–15 Gb	2.5 to 4 h	Unknown	Optimised for targeted sequencing (specific panels available). Also suitable for WGS of small genomes and for exome and transcriptome sequencing
	Ion Torrent	200–400 bp	~99%	0.1–2 Gb	2 to 7 h	375–3,000	WGS of small genomes, targeted sequencing
3rd	**Oxford Nanopore**						
	MinION	5–50 kb	60–95%	0.03–1 Gb	1–48 h (application dependent)	High variation	Long individual reads. Useful for hybrid assemblies to obtain whole-genome sequences. Other uses include full-length cDNA sequencing for isoform analysis and rapid microbial sequencing
	Pacific Biosciences						
	PacBio RS II	8–20 kb	~85% (single) ~99.9% (consensus)	400 Mb 0.4 Gb	2–3 h per cell (up to 16 cells)	700–1,000	Suitable for WGS of small genomes, targeted amplicon sequencing, characterisation of complex populations and characterisation of DNA epigenetic modifications

bp: base pair
cDNA: complementary DNA
Gb: gigabase
HTS: high-throughput sequencing
kb: kilobase
Mb: megabase
Q30: a Phred quality score of 30
WGS: whole-genome sequencing

major technologies are commercially available: single-molecule real-time (SMRT) sequencing, developed by Pacific Biosciences (24); and nanopore-based real-time sequencing with electronic readout, developed by Oxford Nanopore Technologies (25). However, a number of competing platforms are also being developed for third-generation sequencing. The most notable of these is the biological nanopore system, originally designed by Genia Technologies and recently acquired by Roche.

The first available third-generation platform was the PacBio RS II from Pacific Biosciences, which has proven to be suitable for whole-genome sequencing (WGS) of microbial genomes and the targeted sequencing of genomic regions of interest (26, 27, 28). Recently, this company also launched the scalable Sequel system that can generate enough sequence data to enable the whole-genome *de novo* assembly of eukaryotic organisms. Both platforms can produce circular consensus sequencing reads with an average length of 10–15 kilobases (kb) and acceptable accuracy (29). In contrast to the large, bulky sequencers produced by Pacific Biosciences, Oxford Nanopore Technologies initially launched a miniaturised DNA sequencing device, the MinION, in a portable pen-drive format. The MinION was released in an early access programme to initiate a community-driven improvement effort. This led to the generation of read lengths typically ranging from 5 to 50 kb and error rates of between ~5% and 40% (30, 31). The PromethION, a larger desktop version scalable up to 48 flow cells, is scheduled for release, and an even larger modular system, the GridION, is being developed.

Although the first implementation of these technologies resulted in a lower throughput capacity compared with most second-generation sequencing platforms, they have several distinct advantages that make them attractive instruments for clinical laboratories and microbiology research. For example, the capability for real-time data analysis during the sequencing run is an important feature in time-critical applications such as investigating infectious outbreaks, as already demonstrated for viral pathogens (32). Long read lengths have greatly simplified *de novo* genome assembly of novel organisms (33), and the exceptionally long reads produced by the MinION have proven especially useful for hybrid assembly, i.e. when combining the results of two or more sequencing technologies (30). In addition, the use of these platforms is not necessarily limited to DNA sequencing: SMRT sequencing can detect epigenetic modifications such as DNA methylation, which is an important virulence factor for pathogenic bacteria (34); and nanopore-based electronic systems are potentially useful for the direct detection of protein variants and post-translational modifications (such as phosphorylation) linked to pathogenicity (35, 36, 37).

Sampling and sample preparation

The type and quality of specimens can greatly influence the outcome of molecular diagnostic methods. While the nature and stage of disease are primary considerations for sampling, other pre-analytical variables such as the processing, transportation and storage of collected specimens are also very important. These processes have been extensively covered and discussed in the OIE *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals* and *Manual of Diagnostic Tests for Aquatic Animals* (38, 39). Furthermore, the Genomic Standards Consortium (www.gensc.org) strongly recommends that contextual metadata should also be collected at the time of sampling to aid downstream data analysis and interpretation (40).

To increase assay sensitivity and improve the recovery of whole-genome sequences, a variety of sample preparation protocols have been devised for infectious agent enrichment prior to sequencing. Common steps include homogenisation, filtration, ultracentrifugation and nuclease treatment, as well as nucleic acid extraction and pre-amplification (14). Recently, strategies for in-solution capture enrichment of nucleic acids from known pathogen groups have also become popular. It is now feasible to selectively enrich for all known vertebrate viruses and their variants (41, 42).

The target material retained after removing host-derived components and contaminants is often present in such small quantities that amplification is required before sequencing. This is especially the case when clinical samples and/or uncultivable microorganisms are investigated by direct sequencing. Pre-amplification can be performed as a separate step or during sequence library construction (i.e. the generation of platform-specific DNA templates) (43). However, it is important to realise that amplification methods can introduce bias (44, 45, 46). This is usually less important in the detection of infectious agents, when only a positive or negative result is required; however, amplification bias may adversely affect the recovery of whole-genome sequences.

In addition to pre-amplification, second-generation platforms also require the clonal amplification of sequence libraries to generate clusters of identical DNA fragments, which adds another potential source of error (47). Although the lack of clonal amplification is considered one of the main benefits of third-generation sequencing, these methodologies require much more starting material (microgram rather than nanogram amounts of nucleic acids) for the efficient construction of sequence libraries. Overcoming this problem would make them much more useful for investigating infectious agents.

Sequence analysis and comparative genomics

Bioinformatics is the use of computational methods for retrieving, analysing and storing biological data; it is an integral part of all sequencing applications. After the acquisition of sequencing data, a quality control step is typically performed to ensure that the downstream analysis is meaningful. This step is usually followed by screening to remove irrelevant sequence information derived from contaminants, host material or tag sequences (48, 49). Sequence data can be analysed in various ways depending on the purpose of the investigation. The two most common are *i*) direct mapping against a closely related reference genome and *ii*) the *de novo* assembly of reads into longer contigs.

Although numerous algorithms have been developed for the direct mapping of sequence reads, all make different compromises between accuracy and speed (50). The choice of algorithm and software should therefore be based on the specific application and the needs of the user (51). Similarly, a variety of algorithms are available for *de novo* sequence data assembly. However, a common problem is the production of incomplete assemblies resulting from the use of short reads for complex repeat regions (52). An effective solution to this problem is hybrid assembly that combines long but error-prone third-generation sequencing reads and high-quality HTS reads (53).

The steps described above are usually performed within bioinformatics pipelines, which function as workflow management systems designed to automatically direct the flow of sequence data between different software tools. These pipelines are typically customised for different applications, such as pathogen identification (54), metagenomic studies of microbiomes (55), whole-genome assembly (56, 57) and virus variant detection (58). For all applications of genomic sequence data, comparison against known sequences is essential. Comparative genomic analyses facilitate taxonomic classification, functional annotation of genomes and the investigation of evolutionary processes at the molecular level (59). However, for highly divergent or previously unknown infectious agents, it may be necessary to search for remote homologues of protein sequences using methods based on protein sequence profiles, such as profile hidden Markov models (60). Furthermore, to enable phenotypic traits, such as pathogenicity and virulence, to be assessed in comparative studies, it is important to use databases that include relevant phenotypic metadata for the deposited sequences (61).

When planning sequence analysis based on HTS data and comparative genomics, it is necessary to remember that these are resource-intensive tasks. Dedicated high-end computer servers are typically required for the long-running data processing pipelines required to convert raw data into biologically meaningful information packaged in a useful output format. A high-capacity storage solution is also needed for the large data volumes generated by the sequencing platforms and during analysis, and for local storage of recent copies of relevant databases. It is also important not to underestimate the manual interaction needed to manage output files, adjust processing parameters and interpret the data (62). These are all well-known issues associated with this sequencing technology, and great efforts are being made to simplify workflows, automate analysis steps and implement intuitive graphical interfaces to make the technology more accessible to users with limited or no bioinformatics training (63, 64). Even so, the storage of sequenced data remains a huge problem, because the amount of sequence data seems to increase much faster than storage capacity; therefore, data storage might not be indefinitely sustainable (65, 66).

Applications of NGS and comparative genomics in veterinary medicine

In a previous review article published in the OIE *Scientific and Technical Review*, the authors summarised several recent applications and achievements of HTS in veterinary infection biology and diagnostics, including the recent achievements of the OIE Collaborating Centre, Uppsala, Sweden (14). Various HTS techniques were described, along with their applications in veterinary diagnostic microbiology of infectious diseases in different categories of host animals. Examples in swine included detecting and identifying various infectious agents involved in complex diseases, such as post-weaning multisystemic wasting syndrome; detecting infection agents in mixed infections in enteric disease complexes; characterising porcine microbiota (bacterial and viral populations); and host–pathogen interactions. Examples of HTS-based investigations of infections and diseases in cattle and sheep included characterising and determining the phylogeny of new bluetongue virus variants; determining host viral population diversity; and detecting unknown and emerging new pathogens, most notably the Schmallenberg virus. Applications in diseases of poultry and other avian species were also discussed, including the characterisation and phylogenetic analysis of new variants of various viruses; the detection of unknown and 'unexpected' pathogens; investigations of poultry microbiota; host–pathogen interactions; and several aspects of vaccine development and quality control. In addition to pathogens restricted to a single host species (in the

examples described above), multi-host pathogens were also discussed. The review mainly focused on rabies and related viruses, and examples of HTS applications included WGS of Ikoma lyssavirus (14).

The information and discussions already published in the previous OIE review will not be repeated here; however, since it contains many important results, achievements and a long list of selected HTS references, it provides the conceptual background to the current review. As the present review mainly focuses on more recent achievements, these two reviews are complementary and it is recommended that they be read together.

Improved detection and characterisation of infectious agents and early diagnosis

The high biodiversity of infectious agents (such as orthomyxoviruses and paramyxoviruses) and the geographical movement of hosts (e.g. migrating birds or animals transported over continents for sport or touristic purposes) facilitate the distribution of various infectious agents and contribute to the development of new disease reservoirs. These factors also assist the spread of emerging infectious diseases, such as transboundary animal diseases. Since most classic diagnostic methods detect a limited group of infectious agents or have other limitations (e.g. many viruses do not replicate in cultured cells), it is possible that traditional diagnostic approaches will be ineffective in detecting and identifying some of the newly emerging variants of infectious agents.

Recently, HTS-based culture-independent assays have been developed for the non-specific (i.e. broad-target) detection of a wide range of infectious agents. Since these assays are designed to investigate the complete viral and/or microbial genetic composition of a sample, they are referred to as metagenomic detection assays. Metagenomic approaches have led to the identification of a broad variety of 'new' viruses, e.g. novel bocaviruses, torque teno viruses, astroviruses, rotaviruses and kobuviruses that cause porcine disease; new virus variants affecting honeybee populations; and a range of other infectious agents in various host species (3, 67, 68). Kapgate *et al*. described the use of NGS to identify novel avian viruses and investigate their diversity. This led to the identification of a range of new avian viruses, including picobirnavirus, picornavirus, orthoreovirus and avian gamma coronavirus. Viral–host interactions and disease-associated factors have also been reported in newly identified avian viruses (69).

Modern sequencing technologies have greatly simplified WGS and thus the study of genomic recombination, which has an important role in bacterial and viral evolution. A practical example is the evolution of feline infectious peritonitis virus, a coronavirus that causes fatal debilitating disease in the Felidae family. Its pathogenicity is believed to result from recombination events that led to concurrent genetic alterations in different parts of the genome. NGS techniques have been used to identify these virulence markers, and their roles in disease pathogenesis have been assessed by reverse genetic methods (70). This information may form the basis of coronavirus vaccine development, which has been an extremely difficult task because of the complicated infection biology of coronaviruses (71). WGS is also a useful method of investigating segmented genomes. For example, NGS of the bluetongue virus responsible for recent disease outbreaks in Europe revealed that the central European virus evolved via the reassortment of multiple genome segments derived from several different bluetongue virus variants. This new genetic information may help improve administrative measures and provide more reliable disease control by vaccination (72).

Surveillance of disease outbreaks and transmission routes

Comparative nucleotide sequence analysis is a useful tool for epidemiological studies and the effective control of infectious diseases. Thus, HTS approaches can function as important tools for studying disease outbreaks by identifying and following transmission routes, thereby facilitating the identification of outbreak origins (i.e. providing improved traceability of infectious agents). For example, NGS-based comparative whole-genome analysis formed part of the diagnostic process used to investigate the sporadic outbreaks of avian influenza virus (AIV) H5N8 in Europe during 2014–2015. Phylogenetic analysis revealed a close resemblance to strains from South Korea and Japan, suggesting a transmission pattern that involved migratory wild birds from Asia, possibly via overlapping flyways and common breeding sites in Siberia (73). This type of HTS work is regularly performed by the collaborative efforts of international consortia and organisations, as well as by individual laboratories and centres of excellence. For example, the Vietnam Initiative on Zoonotic Infections was established to investigate the origins and emergence of zoonotic infections in a region of the Far East. In this countrywide project, several international institutions are collaborating with Vietnamese organisations to combine HTS, clinical data analysis, epidemiology and social science to address relevant One Health questions. The principal objective is to develop an integrated approach to the surveillance of pathogens circulating in both human and animal populations and to assess how frequently these pathogens are exchanged. This infrastructure will enable the systematic investigation of pathogen ecology and evolution, enhance our understanding of viral cross-species transmission events, and identify the risk factors and drivers for zoonotic disease emergence (74).

Another example is Epi-SEQ, which is an EMIDA ERA-NET (Coordination of European Research on Emerging and Major Infectious Diseases of Livestock – European Research Area Network) transnational research project. In this project, NGS-based technologies are applied to address two ubiquitous challenges. The first challenge is how to obtain and analyse sequence data from large collections of viruses from specific outbreaks to create improved models of viral transmission and epidemiology. Prior to the use of NGS, high-resolution (complete genome) sequence data required individual samples to be processed separately, which limited the numbers of samples that could realistically be analysed. Multiplex NGS approaches allow large numbers of samples to be processed and they are generating data of unprecedented resolution and detail. The epidemiology of acutely acting viruses, e.g. foot and mouth disease virus (FMDV), AIV, Newcastle disease virus and classical swine fever virus, are being studied in detail using existing samples in European reference laboratories collected from field outbreaks. These include the complete sample archive of FMDV from the 2001 outbreak in the United Kingdom and of AIV H7N1 viruses from Italian outbreaks between 1999 and 2001.

The second challenge is the use of deep resequencing of viral sequences within single samples to assess viral evolutionary dynamics. Viruses have relatively large population sizes and high replication rates. Their replication machinery introduces errors when the genome is replicated (for some RNA viruses, roughly 0.0001 mutations per nucleotide copied). Consequently, these viruses exhibit high genetic diversity, evolve very rapidly and can quickly adapt to different environmental pressures. The replication errors therefore allow viruses to evade host immune defences. They may also reduce the protective effects of vaccines and antiviral drug treatments, alter viral pathogenicity and host range, and compromise the reliability of diagnostic tests. (For further details, see www.episeq.eu.)

Use of HTS in the field during outbreaks – 'on-site' applications

In the coming years, sequencing technologies will be applied not only in central laboratories but also directly in the field during outbreaks and other manifestations of infectious diseases to significantly shorten the time to diagnosis. In addition, the biosafety level can be higher (no transport-related risks), costs can be lower (no transport-related costs) and diagnoses may be more reliable (no transport-related sample damage or target degradation) for 'on-site' applications than for sequencing in central laboratories.

Reliable on-site diagnosis helps health authorities to rapidly identify the infectious agents and apply specific control measures. A prompt, early diagnosis and rapid implementation of specific control measures can stop outbreaks or prevent their uncontrolled spread. Therefore, on-site HTS has medical and economic benefits, and can improve socio-economic conditions and 'quality of life'.

A practical and functional example of on-site sequencing-based technologies was provided by the European Mobile Laboratory Project (see www.emlab.eu). The project, 'Establishment of Mobile Laboratories for Pathogens up to Risk Group 4 in Combination with CBRN [Chemical, Biological, Radiological and Nuclear] Capacity Building in Sub-Saharan Africa', was funded by the European Commission to build the capacity to bring state-of-the-art technology and diagnostics and highly trained scientists to sites of infectious disease epidemics in Europe and sub-Saharan Africa. This project set up three mobile laboratory units with operational readiness to respond to infectious disease outbreaks on site. A pool of potential responders, i.e. scientists and technicians experienced in infectious disease diagnostics, molecular biology and research, was also established for field missions. The mobile laboratories performed on-site HTS during outbreaks to help in the temporal and spatial analysis of the 2014–2015 Ebola virus (EBOV) outbreaks in West Africa. Deep sequencing of 179 patient samples by the European Mobile Laboratory, the first diagnostics unit to be deployed in the outbreak epicentre in Guinea, revealed the epidemiological and evolutionary history of the outbreak from March 2014 to January 2015. Analysis of the EBOV genome evolution also benefited from a similar effort to sequence patient samples from Sierra Leone. In 2015, Carroll *et al*. confirmed that EBOV probably moved from Guinea into Sierra Leone in April or early May of 2014 and that viruses of the Guinea/Sierra Leone lineage mixed together during June to July 2014. Viral sequences from August, September and October of 2014 indicate that this lineage evolved independently within Guinea. The authors reported that data analysis combined with epidemiological information can be used to retrospectively test the effectiveness of control measures, thus providing unprecedented insight into the evolution of ongoing viral haemorrhagic fever outbreaks (75).

Although this example relates to EBOV and mainly focuses on human medicine (not the main subject of this review), the authors believe that it is important to present the latest developments of on-site HTS. Mobile laboratory facilities can similarly diagnose disease and analyse infectious agents (including zoonotic pathogens) in humans and animals. Thus, it is very likely that mobile laboratory units (or simpler sequencing equipment) will be used during disease outbreaks in animal populations in the near future. This is likely to greatly improve both human and veterinary medicine, as hypothesised above.

Microbiome analysis and between-species comparisons

High-throughput sequencing provides excellent tools to investigate the complex microbiomes of various host

species. Analysis of microbiome composition provides valuable new information on various aspects of infection biology, such as co-infection with several infectious agents, the simultaneous presence of viral and bacterial sub-populations within a single organ, and the evolution and genomic stability of pathogens. This information will aid the development of novel diagnostic approaches and improve control measures. For example, HTS has been successfully used to study the collateral effects of antibiotic use on beneficial gut microbial communities (76). Differences in the microbiome composition of various organ systems between healthy individuals and those with disease have also been demonstrated using HTS (reviewed by Hoffmann *et al*. [77]). Furthermore, HTS-based investigations of the viromes of various mammals indicate that viruses can be viewed as part of the microbiome, interacting with and influencing the behaviour of commensal bacteria and other microbial agents (78).

Characterisation and assessment of genetic diversity within and between viral populations

The latest sequencing technologies have opened up new possibilities for characterising the mutant spectra of viral quasispecies. Simultaneous advancements in ultra-deep HTS technologies and sophisticated bioinformatics tools have enabled low-frequency variants to be detected. Andino and Domingo recently reviewed the use of HTS to further our understanding of virus population dynamics, notably the 'occurrence of intra-mutant spectrum interactions, and the implications of fitness landscapes for virus adaptation and de-adaptation' (79). The impact of coexisting viral sub-populations on disease development has also been assessed, for example in the highly heterogeneous porcine reproductive and respiratory syndrome virus (PRRSV). According to Lu *et al*., HTS-based whole-genome comparison of viral variants has identified functionally important regions of PRRSV and enabled its infection biology to be studied at unprecedented molecular resolution (80).

Investigating the diversity of bacterial communities

The novel sequencing technologies have recently revolutionised various fields of microbiology: bacteriology shows a similar strong progress to that of virology. Since publication of the first two complete bacterial genome sequences in 1995, the science of bacteriology has dramatically changed. Bacterial genome sequencing is now considered a standard procedure; information from tens of thousands of bacterial genomes has altered our view of the bacterial world. A recent review by Land *et al*. illustrates the extensive diversity of bacterial communities using the genome sequences now available from 50 different bacterial phyla. It also summarises the many practical applications of novel sequencing technologies in bacteriology, such as genome-scale metabolic modelling, biosurveillance, bioforensics and infectious disease epidemiology (81).

Complexity of infection biology, including responses to antimicrobial treatments and vaccination

High-throughput sequencing approaches provide excellent tools for studying the complexity of infection biology, including host–pathogen interactions, the genomic features involved in infectious agent transmission and host immune responses. For example, the role of pathogen genomics in assessing disease transmission was reviewed by Sintchenko and Holmes (82). The authors discussed the new insights into disease spread and transmission obtained from WGS, particularly when combined with genomic-scale phylogenetic analyses. The mechanisms of pathogen transmission were described, including their examination by various sequencing approaches.

Truong *et al*. described how the use of HTS in immunological research revealed differing immune responses in the intestinal mucosa of two inbred lines with necrotic enteritis (83). In another application, the influence of host-targeted antiviral drugs (which inhibit viral replication) on viral evolution and diversity has been investigated by various research groups, including Plummer *et al*. in the case of dengue virus infection biology (84).

High-throughput sequencing is highly applicable to several aspects of vaccine development and control. For example, an HTS-based approach involving the Illumina MiSeq platform was used to investigate the phylogenetic relationship between the Rovac vaccine and other lapinised vaccines of classical swine fever. A comparative genetic analysis of near-complete genomes identified the Rovac vaccine as a possible ancestor of Russian vaccine strains but not of the C-strain vaccine (85). Thus, the comparative analysis of large segments of viral genomes can advance the science of vaccinology, especially regarding the identity and origin of various virus strains used in vaccines. Another example concerns the use of vaccination to prevent disease caused by rapidly evolving RNA viruses, such as PRRSV. Novel sequencing technologies can now provide fast, reliable tools to confirm or exclude the possibility that vaccine reversion has occurred during severe disease outbreaks (86). This is very useful for the successful design and implementation of eradication programmes.

HTS-based investigations into the composition and quality of pharmaceutical products and vaccines

Although not directly relevant to this review, it is worth mentioning the frequent use of HTS in other related

biotechnological fields, for example the testing and control of pharmaceutical products and vaccines. These applications can have complementary or synergistic effects on pathogen diagnosis and control in veterinary infection biology.

These and related applications have been widely discussed in recent publications. For example, Dumarest *et al*. described the viral diversity in the swine intestinal mucus used to manufacture heparin. Heparin is one of the main pharmaceutical products manufactured from raw animal material. To describe the viral burden associated with this raw material, researchers performed HTS on mucus samples collected from European pigs and destined for heparin manufacture They identified *Circoviridae* and *Parvoviridae* family members as the most prevalent contaminating viruses, together with viruses from the *Adenoviridae, Anelloviridae, Astroviridae, Birnaviridae, Caliciviridae, Picornaviridae*, and *Reoviridae* families. These investigations indicated the effectiveness of HTS and related bioinformatics in improving biosafety and both animal and human health (87).

These examples convincingly show the capacity of HTS-based diagnostics to revolutionise veterinary and human medicine in the near future because of their speed and accuracy in detecting and identifying pathogens and their ability to provide essential information and support for rapid, effective treatment measures.

Discussion

There are two main approaches to the 'direct diagnosis' of infectious diseases: conventional diagnostic methods and molecular diagnostic methods. Conventional diagnosis includes classic methods, such as virus isolation and bacteria cultivation, along with various biological assays for the identification and characterisation of isolates and viral and bacterial strains. Molecular diagnosis techniques encompass various nucleic acid detection and amplification assays, e.g. nucleic acid hybridisation, PCR and isothermal amplification assays, solid and liquid phase microarrays, and recently developed metagenomics methods. The main topic of this review, genome sequencing of infectious agents, falls into the molecular diagnosis category because sequencing methods complement and strengthen molecular diagnostic approaches. For example, comparative sequence analysis of PCR products, including long-template PCR amplicons, provides valuable information for both diagnostic and infection biology studies. Similarly, metagenomic approaches such as the detection and characterisation of 'unknown infectious agents' have been greatly facilitated, enhanced, and enabled by the various sequencing technologies.

The experiences of the diagnostic laboratories indicate that both conventional and molecular diagnostic methods have strengths and weaknesses. For example, a strength of conventional diagnostic methods, such as virus isolation, is that they can simultaneously detect a range of infectious agents. More importantly, successful virus isolation results in a valuable collection of viral isolates and/or strains, which are essential for studying infection biology and developing diagnostic panels, new vaccines and immunotherapeutic products. On the other hand, a weakness is its inability to isolate and detect infectious agents that do not replicate in the available cell culture systems.

Several molecular diagnostic methods (e.g. so-called 'wide-range PCR panels') can also detect a wide range of infectious agents in a single assay. Metagenomics assays have an even broader detection range: they can detect 'all' (or most) infectious agents in samples, irrespective of their character and/or origin. It is important to note that molecular methods also have limitations compared with conventional methods. A major weakness is that, unlike classic virus isolation, most molecular diagnostic methods do not yield isolates or strains. This is a serious limitation for diagnostic research because it is essential to obtain new isolates and strains (for the reasons already discussed). However, molecular clones can sometimes be chemically synthesised from sequence data and transfected into cultured cells to recover viral strains. Nevertheless, for many viruses, we still lack the molecular tools for the *in vitro* generation of viruses.

It is important to emphasise the clear tendency over the last few decades for molecular diagnostic approaches to partially or fully replace conventional methods in most diagnostic laboratories. In many diagnostic laboratories, conventional diagnostic methods (e.g. virus isolation) have been completely lost. Molecular diagnostic approaches are replacing conventional methods because of their lower associated costs (e.g. of equipment, reagent costs and facilities) and the lack of available experts in conventional diagnostic techniques. Although this initially appears a reasonable progression, in the authors' opinion, a major strategic mistake is being made. Based on the experiences of the OIE Collaborating Centre in Uppsala (and of many of the authors' virology and bacteriology colleagues throughout the OIE network), the authors maintain that a combination of both conventional and novel molecular diagnostic techniques are needed to ensure a high level of diagnostic ability. They maintain that virus isolation is absolutely necessary because viral isolates and strains are needed to study viral infection biology, develop novel diagnostic assays and produce new vaccines, among other veterinary medicine tasks. Thus, the authors' recommendations are to:

– maintain conventional methods and make further improvements to these, even if they are costly

– develop various molecular assays, for both specific (narrow-target) and non-specific (broad-target) detection

– combine conventional and molecular approaches to improve existing diagnostic facilities

– standardise and validate the various approaches by following OIE recommendations, and share information globally via the OIE network.

Today, various nucleotide sequencing-based approaches are directly improving multidisciplinary, complex diagnostic panels and are being coupled to conventional and/or novel molecular approaches. Virus isolates or strains obtained by classic methods can be directly investigated by WGS and comparative sequence analyses. This combined analysis is improving diagnostic approaches and providing deeper insight into infection biology in veterinary medicine.

Although the introduction and use of novel sequencing-based technologies is providing huge real and potential benefits in veterinary medicine today, some drawbacks and issues still need to be addressed. Most importantly, few (if any) HTS-based approaches have been validated for routine diagnostic use in veterinary medicine. Instead, they have mainly been used in research applications to identify novel infectious agents and investigate host–pathogen interactions or as adjuncts to provide more complete genetic characterisation (e.g. WGS). Nonetheless, it is important to consider using HTS-based approaches for diagnostic purposes in routine settings. The variety of available sequencing platforms and their proposed applications mean that validation is likely to be required in a platform- and application-specific manner. It is also important to consider not only the sequencing technologies themselves but also all steps in the processes from sample collection to the final results of the bioinformatics analysis. A need for computational resources and structured data storage should also be considered before HTS-based methods are routinely applied in our laboratories (88).

In summary, this review provides current and prospective views on the opportunities and challenges of novel sequencing technologies and comparative sequence analysis in veterinary medicine, with a particular focus on applications with a potential impact on disease diagnosis, control and management (89). After considering examples of the recent progress of novel sequencing technologies in veterinary medicine, it is important to ask whether the massive amount of information arising from high-throughput technologies can improve the diagnosis, control and management of infectious diseases in practice. The examples indicate a positive answer. These techniques are being widely applied today, and this is due to continue in the future. The novel sequencing methodologies, harmonised in the OIE network, provide powerful novel tools for combating infectious diseases according to the principles of 'One World, One Health'.

Acknowledgements

This work was supported by the Epi-SEQ research project funded under the second joint call for transnational research projects by EMIDA ERA-NET (FP7 Project No. 219235) and by the Strong Research Environments BioBridges project (No. 2011–1692) grant from the Swedish Research Council Formas.

Les nouvelles technologies appliquées au séquençage nucléotidique et à l'analyse séquentielle comparative des génomes des agents infectieux en médecine vétérinaire

F. Granberg, Á. Bálint & S. Belák

Résumé

Le séquençage de nouvelle génération (également désigné « séquençage à très haut débit » ou « séquençage massivement parallèle ») est un nouvel outil extrêmement puissant permettant de procéder à des diagnostics sophistiqués et d'assurer un contrôle vétérinaire intensif de maladies infectieuses complexes.

Les technologies du séquençage de nouvelle génération sont également d'un grand secours pour étudier l'étiologie, la génomique, l'évolution et l'épidémiologie des maladies infectieuses ainsi que les interactions hôtes–agents pathogènes et bien d'autres aspects de la biologie des maladies infectieuses.
Pour présenter les progrès et les accomplissements les plus récents dans ce domaine, les auteurs décrivent d'abord une série de techniques innovantes puis quelques exemples d'applications du séquençage de nouvelle génération dans le champ de la biologie des maladies animales infectieuses. Ils exposent un certain nombre d'étapes et de processus opérationnels régissant la sélection et la préparation des échantillons, l'analyse séquentielle et les études de génomique comparative, ainsi que ceux qui permettent d'améliorer la justesse prédictive de la génomique ; les exigences particulières de la bio-informatique sont également évoquées. Cette analyse est complétée par quelques exemples d'applications de l'analyse séquentielle et de la génomique comparative en médecine vétérinaire. Cette synthèse est basée sur une sélection de références bibliographiques récentes ainsi que sur l'expérience acquise par le Centre collaborateur de l'Organisation mondiale de la santé animale (OIE) pour le diagnostic basé sur la biotechnologie des maladies infectieuses en médecine vétérinaire, situé à Uppsala (Suède).

Mots-clés
Agent causal – Bactérie – Diagnostic – Génomique comparative – Maladie infectieuse – Médecine vétérinaire – Métagénomique – Séquençage à très haut débit – Séquençage de nouvelle génération – Virus.

Técnicas novedosas aplicadas a la secuenciación de nucleótidos y el análisis comparado de secuencias genómicas de agentes infecciosos en medicina veterinaria

F. Granberg, Á. Bálint & S. Belák

Resumen
La secuenciación de próxima generación, también denominada secuenciación profunda, de alto rendimiento o masivamente paralela, es una nueva y poderosa herramienta para efectuar diagnósticos complejos y vigilar muy de cerca enfermedades infecciosas complejas en el ámbito de la medicina veterinaria. Estas técnicas también se utilizan cada vez más para estudiar la etiología, genómica, evolución y epidemiología de las enfermedades infecciosas, así como las interacciones entre patógeno y anfitrión y otros aspectos de la biología de las infecciones.

Los autores resumen brevemente una serie de logros y adelantos obtenidos últimamente en este ámbito, presentando en primer lugar un conjunto de técnicas novedosas y ofreciendo después ejemplos de aplicaciones de la secuenciación de próxima generación a la biología de las infecciones veterinarias. También describen varios protocolos y procesos de trabajo para la obtención y preparación de muestras, el análisis de secuencias y las labores de genómica comparada, explican cómo mejorar la exactitud de la predicción genómica y examinan las herramientas bioinformáticas necesarias para ello. A continuación presentan ejemplos de aplicaciones basadas en técnicas de secuenciación y en la genómica comparada en medicina veterinaria. Este artículo está basado en referencias muy recientes, tomadas de publicaciones científicas, y en la experiencia del Centro Colaborador de la Organización Mundial de Sanidad Animal (OIE) para el Diagnóstico de las enfermedades infecciosas de la medicina veterinaria basado en la biotecnología, sito en Upsala (Suecia).

Palabras clave
Agente infeccioso – Bacteria – Diagnóstico – Enfermedad infecciosa – Genómica comparada – Medicina veterinaria – Metagenómica – Secuenciación de alto rendimiento – Secuenciación de próxima generación – Virus.

References

1. Belák S. (2007). – Molecular diagnosis of viral diseases, present trends and future aspects: a view from the OIE Collaborating Centre for the application of polymerase chain reaction methods for diagnosis of viral diseases in veterinary medicine. *Vaccine*, **25** (30), 5444–5452. doi:10.1016/j.vaccine.2006.11.068.

2. Belák S., Thorén P., LeBlanc N. & Viljoen G. (2009). – Advances in viral disease diagnostic and molecular epidemiological technologies. *Expert. Rev. Mol. Diagn.*, **9** (4), 367–381. doi:10.1586/erm.09.19.

3. Belák S. & Liu L. (2013). – Recent advances in veterinary diagnostic virology: report from a Collaborating Centre of the World Organisation for Animal Health (OIE). *In* Advanced techniques in diagnostic microbiology (Y.-W. Tang & C.W. Stratton, eds). Springer, New York, 661–678.

4. Maxam A.M. & Gilbert W. (1977). – A new method for sequencing DNA. *Proc. Natl Acad. Sci. USA*, **74** (2), 560–564.

5. Sanger F., Nicklen S. & Coulson A.R. (1977). – DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74** (12), 5463–5467.

6. Smith L.M., Sanders J.Z., Kaiser R.J., Hughes P., Dodd C., Connell C.R., Heiner C., Kent S.B. & Hood L.E. (1986). – Fluorescence detection in automated DNA sequence analysis. *Nature*, **321** (6071), 674–679. doi:10.1038/321674a0.

7. Mullis K., Faloona F., Scharf S., Saiki R., Horn G. & Erlich H. (1986). – Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.*, **51** (Pt 1), 263–273.

8. Mullis K.B. & Faloona F.A. (1987). – Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Meth. Enzymol.*, **155**, 335–350.

9. Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G., Smith H.O., Yandell M., Evans C.A., Holt R.A., Gocayne J.D., Amanatides P., Ballew R.M., Huson D.H., Wortman J.R., Zhang Q., Kodira C.D., Zheng X.H., Chen L., Skupski M., Subramanian G., Thomas P.D., Zhang J., Gabor Miklos G.L., Nelson C., Broder S., Clark A.G., Nadeau J., McKusick V.A., Zinder N., Levine A.J., Roberts R.J., Simon M., Slayman C., Hunkapiller M., Bolanos R., Delcher A., Dew I., Fasulo D., Flanigan M., Florea L., Halpern A., Hannenhalli S., Kravitz S., Levy S., Mobarry C., Reinert K., Remington K., Abu-Threideh J., Beasley E., Biddick K., Bonazzi V., Brandon R., Cargill M., Chandramouliswaran I., Charlab R., Chaturvedi K., Deng Z., Di Francesco V., Dunn P., Eilbeck K., Evangelista C., Gabrielian A.E., Gan W., Ge W., Gong F., Gu Z., Guan P., Heiman T.J., Higgins M.E., Ji R.R., Ke Z., Ketchum K.A., Lai Z., Lei Y., Li Z., Li J., Liang Y., Lin X., Lu F., Merkulov G.V., Milshina N., Moore H.M., Naik A.K., Narayan V.A., Neelam B., Nusskern D., Rusch D.B., Salzberg S., Shao W., Shue B., Sun J., Wang Z., Wang A., Wang X., Wang J., Wei M., Wides R., Xiao C., Yan C., Yao A., Ye J., Zhan M., Zhang W., Zhang H., Zhao Q., Zheng L., Zhong F., Zhong W., Zhu S., Zhao S., Gilbert D., Baumhueter S., Spier G., Carter C., Cravchik A., Woodage T., Ali F., An H., Awe A., Baldwin D., Baden H., Barnstead M., Barrow I., Beeson K., Busam D., Carver A., Center A., Cheng M.L., Curry L., Danaher S., Davenport L., Desilets R., Dietz S., Dodson K., Doup L., Ferriera S., Garg N., Gluecksmann A., Hart B., Haynes J., Haynes C., Heiner C., Hladun S., Hostin D., Houck J., Howland T., Ibegwam C., Johnson J., Kalush F., Kline L., Koduru S., Love A., Mann F., May D., McCawley S., McIntosh T., McMullen I., Moy M., Moy L., Murphy B., Nelson K., Pfannkoch C., Pratts E., Puri V., Qureshi H., Reardon M., Rodriguez R., Rogers Y.H., Romblad D., Ruhfel B., Scott R., Sitter C., Smallwood M., Stewart E., Strong R., Suh E., Thomas R., Tint N.N., Tse S., Vech C., Wang G., Wetter J., Williams S., Williams M., Windsor S., Winn-Deen E., Wolfe K., Zaveri J., Zaveri K., Abril J.F., Guigo R., Campbell M.J., Sjolander K.V., Karlak B., Kejariwal A., Mi H., Lazareva B., Hatton T., Narechania A., Diemer K., Muruganujan A., Guo N., Sato S., Bafna V., Istrail S., Lippert R., Schwartz R., Walenz B., Yooseph S., Allen D., Basu A., Baxendale J., Blick L., Caminha M., Carnes-Stine J., Caulk P., Chiang Y.H., Coyne M., Dahlke C., Mays A., Dombroski M., Donnelly M., Ely D., Esparham S., Fosler C., Gire H., Glanowski S., Glasser K., Glodek A., Gorokhov M., Graham K., Gropman B., Harris M., Heil J., Henderson S., Hoover J., Jennings D., Jordan C., Jordan J., Kasha J., Kagan L., Kraft C., Levitsky A., Lewis M., Liu X., Lopez J., Ma D., Majoros W., McDaniel J., Murphy S., Newman M., Nguyen T., Nguyen N., Nodell M., Pan S., Peck J., Peterson M., Rowe W., Sanders R., Scott J., Simpson M., Smith T., Sprague A., Stockwell T., Turner R., Venter E., Wang M., Wen M., Wu D., Wu M., Xia A., Zandieh A. & Zhu X. (2001). – The sequence of the human genome. *Science*, **291** (5507), 1304–1351. doi:10.1126/science.1058040.

10. Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczky J., LeVine R., McEwan P., McKernan K., Meldrim J., Mesirov J.P., Miranda C., Morris W., Naylor J., Raymond C., Rosetti M., Santos R., Sheridan A., Sougnez C., Stange-Thomann Y., Stojanovic N., Subramanian A., Wyman D., Rogers J., Sulston J., Ainscough R., Beck S., Bentley D., Burton J., Clee C., Carter N., Coulson A., Deadman R., Deloukas P., Dunham A., Dunham I., Durbin R., French L., Grafham D., Gregory S., Hubbard T., Humphray S., Hunt A., Jones M., Lloyd C., McMurray A., Matthews L., Mercer S., Milne S., Mullikin J.C., Mungall A., Plumb R., Ross M., Shownkeen R., Sims S., Waterston R.H., Wilson R.K., Hillier L.W., McPherson J.D., Marra M.A., Mardis E.R., Fulton L.A., Chinwalla A.T., Pepin K.H., Gish W.R., Chissoe S.L., Wendl M.C., Delehaunty K.D., Miner T.L., Delehaunty A., Kramer J.B., Cook L.L., Fulton R.S., Johnson D.L., Minx P.J., Clifton S.W., Hawkins T., Branscomb E., Predki P., Richardson P., Wenning S., Slezak T., Doggett N., Cheng J.F., Olsen A., Lucas S., Elkin C., Uberbacher E., Frazier M., Gibbs R.A., Muzny D.M., Scherer S.E., Bouck J.B., Sodergren E.J., Worley K.C., Rives C.M., Gorrell J.H., Metzker M.L., Naylor S.L., Kucherlapati R.S., Nelson D.L., Weinstock G.M., Sakaki Y., Fujiyama A., Hattori M., Yada T., Toyoda A., Itoh T., Kawagoe C., Watanabe H., Totoki Y., Taylor T., Weissenbach J., Heilig R., Saurin W., Artiguenave F., Brottier P., Bruls T., Pelletier E., Robert C., Wincker P., Smith D.R., Doucette-Stamm L., Rubenfield M., Weinstock K., Lee H.M., Dubois J., Rosenthal A., Platzer M., Nyakatura G., Taudien S., Rump A., Yang H., Yu J., Wang J., Huang G., Gu J., Hood L., Rowen L., Madan A., Qin S., Davis R.W., Federspiel N.A., Abola A.P., Proctor M.J., Myers R.M., Schmutz J., Dickson M., Grimwood J., Cox D.R., Olson M.V., Kaul R., Shimizu N., Kawasaki K., Minoshima S., Evans G.A., Athanasiou M., Schultz R., Roe B.A., Chen F., Pan H., Ramser J., Lehrach H., Reinhardt R., McCombie W.R., de la Bastide M., Dedhia N., Blocker H., Hornischer K., Nordsiek G., Agarwala R., Aravind L., Bailey J.A., Bateman A., Batzoglou S., Birney E., Bork P., Brown D.G., Burge C.B., Cerutti L., Chen H.C., Church D., Clamp M., Copley R.R., Doerks T., Eddy S.R., Eichler E.E., Furey T.S., Galagan J., Gilbert J.G., Harmon C., Hayashizaki Y., Haussler D., Hermjakob H., Hokamp K., Jang W., Johnson L.S., Jones T.A., Kasif S., Kaspryzk A., Kennedy S., Kent W.J., Kitts P., Koonin E.V., Korf I., Kulp D., Lancet D., Lowe T.M., McLysaght A., Mikkelsen T., Moran J.V., Mulder N., Pollara V.J., Ponting C.P., Schuler G., Schultz J., Slater G., Smit A.F., Stupka E., Szustakowki J., Thierry-Mieg D., Thierry-Mieg J., Wagner L., Wallis J., Wheeler R., Williams A., Wolf Y.I., Wolfe K.H., Yang S.P., Yeh R.F., Collins F., Guyer M.S., Peterson J., Felsenfeld A., Wetterstrand K.A., Patrinos A., Morgan M.J., de Jong P., Catanese J.J., Osoegawa K., Shizuya H., Choi S. & Chen Y.J. (2001). – Initial sequencing and analysis of the human genome. *Nature*, **409** (6822), 860–921. doi:10.1038/35057062.

11. Shendure J., Porreca G.J., Reppas N.B., Lin X., McCutcheon J.P., Rosenbaum A.M., Wang M.D., Zhang K., Mitra R.D. & Church G.M. (2005). – Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, **309** (5741), 1728–1732. doi:10.1126/science.1117389.

12. Margulies M., Egholm M., Altman W.E., Attiya S., Bader J.S., Bemben L.A., Berka J., Braverman M.S., Chen Y.J., Chen Z., Dewell S.B., Du L., Fierro J.M., Gomes X.V., Godwin B.C., He W., Helgesen S., Ho C.H., Irzyk G.P., Jando S.C., Alenquer M.L., Jarvie T.P., Jirage K.B., Kim J.B., Knight J.R., Lanza J.R., Leamon J.H., Lefkowitz S.M., Lei M., Li J., Lohman K.L., Lu H., Makhijani V.B., McDade K.E., McKenna M.P., Myers E.W., Nickerson E., Nobile J.R., Plant R., Puc B.P., Ronan M.T., Roth G.T., Sarkis G.J., Simons J.F., Simpson J.W., Srinivasan M., Tartaro K.R., Tomasz A., Vogt K.A., Volkmer G.A., Wang S.H., Wang Y., Weiner M.P., Yu P., Begley R.F. & Rothberg J.M. (2005). – Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, **437** (7057), 376–380. doi:10.1038/nature03959.

13. Bennett S. (2004). – Solexa Ltd. *Pharmacogenomics*, **5** (4), 433–438. doi:10.1517/14622416.5.4.433.

14. Belák S., Karlsson O.E., Leijon M. & Granberg F. (2013). – High-throughput sequencing in veterinary infection biology and diagnostics. *Rev. Sci. Tech. Off. Int. Epiz.*, **32** (3), 893–915.

15. Wetterstrand K.A. (2015). – DNA sequencing costs: data from the NHGRI Genome Sequencing Program (GSP). Available at: www.genome.gov/sequencingcosts (accessed on 30 October 2015).

16. Loman N.J., Misra R.V., Dallman T.J., Constantinidou C., Gharbia S.E., Wain J. & Pallen M.J. (2012). – Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.*, **30** (5), 434–439. doi:10.1038/nbt.2198.

17. 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.*, Article ID: 251364. doi:10.1155/2012/251364.

18. 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** (10), 991–998. doi:10.1056/NEJMoa073785.

19. Cox-Foster D.L., Conlan S., Holmes E.C., Palacios G., Evans J.D., Moran N.A., Quan P.L., Briese T., Hornig M., Geiser D.M., Martinson V., vanEngelsdorp D., Kalkstein A.L., Drysdale A., Hui J., Zhai J., Cui L., Hutchison S.K., Simons J.F., Egholm M., Pettis J.S. & Lipkin W.I. (2007). – A metagenomic survey of microbes in honey bee colony collapse disorder. *Science*, **318** (5848), 283–287. doi:10.1126/science.1146498.

20. Rothberg J.M. & Leamon J.H. (2008). – The development and impact of 454 sequencing. *Nat. Biotechnol.*, **26** (10), 1117–1124. doi:10.1038/nbt1485.

21. Quail M.A., Smith M., Coupland P., Otto T.D., Harris S.R., Connor T.R., Bertoni A., Swerdlow H.P. & Gu Y. (2012). – A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*, **13**, 341. doi:10.1186/1471-2164-13-341.

22. Rothberg J.M., Hinz W., Rearick T.M., Schultz J., Mileski W., Davey M., Leamon J.H., Johnson K., Milgrew M.J., Edwards M., Hoon J., Simons J.F., Marran D., Myers J.W., Davidson J.F., Branting A., Nobile J.R., Puc B.P., Light D., Clark T.A., Huber M., Branciforte J.T., Stoner I.B., Cawley S.E., Lyons M., Fu Y., Homer N., Sedova M., Miao X., Reed B., Sabina J., Feierstein E., Schorn M., Alanjary M., Dimalanta E., Dressman D., Kasinskas R., Sokolsky T., Fidanza J.A., Namsaraev E., McKernan K.J., Williams A., Roth G.T. & Bustillo J. (2011). – An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, **475** (7356), 348–352. doi:10.1038/nature10242.

23. Schadt E.E., Turner S. & Kasarskis A. (2010). – A window into third-generation sequencing. *Hum. Molec. Genet.*, **19** (R2), R227–R240. doi:10.1093/hmg/ddq416.

24. Eid J., Fehr A., Gray J., Luong K., Lyle J., Otto G., Peluso P., Rank D., Baybayan P., Bettman B., Bibillo A., Bjornson K., Chaudhuri B., Christians F., Cicero R., Clark S., Dalal R., Dewinter A., Dixon J., Foquet M., Gaertner A., Hardenbol P., Heiner C., Hester K., Holden D., Kearns G., Kong X., Kuse R., Lacroix Y., Lin S., Lundquist P., Ma C., Marks P., Maxham M., Murphy D., Park I., Pham T., Phillips M., Roy J., Sebra R., Shen G., Sorenson J., Tomaney A., Travers K., Trulson M., Vieceli J., Wegener J., Wu D., Yang A., Zaccarin D., Zhao P., Zhong F., Korlach J. & Turner S. (2009). – Real-time DNA sequencing from single polymerase molecules. *Science*, **323** (5910), 133–138. doi:10.1126/science.1162986.

25. Kasianowicz J.J., Brandin E., Branton D. & Deamer D.W. (1996). – Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl Acad. Sci. USA*, **93** (24), 13770–13773. doi:10.1073/pnas.93.24.13770.

26. Huddleston J., Ranade S., Malig M., Antonacci F., Chaisson M., Hon L., Sudmant P.H., Graves T.A., Alkan C., Dennis M.Y., Wilson R.K., Turner S.W., Korlach J. & Eichler E.E. (2014). – Reconstructing complex regions of genomes using long-read sequencing technology. *Genome Res.*, **24** (4), 688–696. doi:10.1101/gr.168450.113.

27. Huser D., Gogol-Doring A., Chen W. & Heilbronn R. (2014). – Adeno-associated virus type 2 wild-type and vector-mediated genomic integration profiles of human diploid fibroblasts analyzed by third-generation PacBio DNA sequencing. *J. Virol.*, **88** (19), 11253–11263. doi:10.1128/JVI.01356–14.

28. Koren S., Harhay G.P., Smith T.P., Bono J.L., Harhay D.M., McVey S.D., Radune D., Bergman N.H. & Phillippy A.M. (2013). – Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol.*, **14** (9), R101. doi:10.1186/gb-2013-14-9-r101.

29. Pacific Biosciences (2014). – New chemistry boosts average read length to 10–15 kb for PacBio® RS II. Available at: www.pacb.com/blog/new-chemistry-boosts-average-read/ (accessed on 30 October 2015).

30. Goodwin S., Gurtowski J., Ethe-Sayers S., Deshpande P., Schatz M.C. & McCombie W.R. (2015). – Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. *Genome Res.*, **25** (11), 1750–1756. doi:10.1101/gr.191395.115.

31. Quick J., Quinlan A.R. & Loman N.J. (2014). – A reference bacterial genome dataset generated on the MinION portable single-molecule nanopore sequencer. *GigaScience*, **3**, 22. doi:10.1186/2047-217X-3-22.

32. 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** (1), 99. doi:10.1186/s13073-015-0220-9.

33. Sakai H., Ken N., Ogiso-Tanaka E., Takahashi Y., Iseki K., Muto C., Satou K., Teruya K., Shiroma A., Shimoji M., Hirano T., Itoh T., Kaga A. & Tomooka N. (2015). – The power of single molecule real-time sequencing technology in the *de novo* assembly of a eukaryotic genome. *Sci. Rep.*, **5**, Article No. 16780. doi:10.1038/srep16780.

34. Manso A.S., Chai M.H., Atack J.M., Furi L., De Ste Croix M., Haigh R., Trappetti C., Ogunniyi A.D., Shewell L.K., Boitano M., Clark T.A., Korlach J., Blades M., Mirkes E., Gorban A.N., Paton J.C., Jennings M.P. & Oggioni M.R. (2014). – A random six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nat. Commun.*, **5**, Article No. 5055. doi:10.1038/ncomms6055.

35. Ge R. & Shan W. (2011). – Bacterial phosphoproteomic analysis reveals the correlation between protein phosphorylation and bacterial pathogenicity. *Genomics Proteomics Bioinformatics*, **9** (4–5), 119–127. doi:10.1016/S1672-0229(11)60015-6.

36. Nivala J., Marks D.B. & Akeson M. (2013). – Unfoldase-mediated protein translocation through an α-hemolysin nanopore. *Nat. Biotechnol.*, **31** (3), 247–250. doi:10.1038/nbt.2503.

37. Rosen C.B., Rodriguez-Larrea D. & Bayley H. (2014). – Single-molecule site-specific detection of protein phosphorylation with a nanopore. *Nat. Biotechnol.*, **32** (2), 179–181. doi:10.1038/nbt.2799.

38. World Organisation for Animal Health (OIE) (2015). – Manual of diagnostic tests for aquatic animals, 17th Ed. OIE, Paris. Available at: www.oie.int/international-standard-setting/aquatic-manual/access-online/ (accessed on 25 January 2016).

39. World Organisation for Animal Health (OIE) (2015). – Manual of diagnostic tests and vaccines for terrestrial animals, 23rd Ed. OIE, Paris. Available at: www.oie.int/en/international-standard-setting/terrestrial-code/access-online/ (accessed on 25 January 2016).

40. Field D., Amaral-Zettler L., Cochrane G., Cole J.R., Dawyndt P., Garrity G.M., Gilbert J., Glockner F.O., Hirschman L., Karsch-Mizrachi I., Klenk H.P., Knight R., Kottmann R., Kyrpides N., Meyer F., San Gil I., Sansone S.A., Schriml L.M., Sterk P., Tatusova T., Ussery D.W., White O. & Wooley J. (2011). – The Genomic Standards Consortium. *PLoS Biol.*, **9** (6), e1001088. doi:10.1371/journal.pbi0.1001088.

41. 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** (5), e01491–15. doi:10.1128/mBio.01491-15.

42. Wylie T.N., Wylie K.M., Herter B.N. & Storch G.A. (2015). – Enhanced virome sequencing through solution-based capture enrichment. *Genome Res.,* **25**, 1910–1920. E-pub.: 22 September. doi:10.1101/gr.191049.115.

43. Head S.R., Komori H.K., LaMere S.A., Whisenant T., Van Nieuwerburgh F., Salomon D.R. & Ordoukhanian P. (2014). – Library construction for next-generation sequencing: overviews and challenges. *Biotechniques*, **56** (2), 61–64, 66, 68, passim. doi:10.2144/000114133.

44. Karlsson O.E., Belák S. & Granberg F. (2013). – The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. *Biosecur. Bioterror.*, **11** (Suppl. 1), 227–234. doi:10.1089/bsp.2013.0008.

45. Pinard R., de Winter A., Sarkis G.J., Gerstein M.B., Tartaro K.R., Plant R.N., Egholm M., Rothberg J.M. & Leamon J.H. (2006). – Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. *BMC Genomics*, **7**, 216. doi:10.1186/1471-2164-7-216.

46. Rosseel T., Van Borm S., Vandenbussche F., Hoffmann B., van den Berg T., Beer M. & Höper D. (2013). – The origin of biased sequence depth in sequence-independent nucleic acid amplification and optimization for efficient massive parallel sequencing. *PLoS ONE*, **8** (9), e76144. doi:10.1371/journal.pone.0076144.

47. Van Dijk E.L., Jaszczyszyn Y. & Thermes C. (2014). – Library preparation methods for next-generation sequencing: tone down the bias. *Experim. Cell Res.*, **322** (1), 12–20. doi:10.1016/j.yexcr.2014.01.008.

48. Falgueras J., Lara A.J., Fernández-Pozo N., Cantón F.R., Pérez-Trabado G. & Claros M.G. (2010). – SeqTrim: a high-throughput pipeline for pre-processing any type of sequence read. *BMC Bioinformatics*, **11**, 38. doi:10.1186/1471-2105-11-38.

49. Daly G.M., Leggett R.M., Rowe W., Stubbs S., Wilkinson M., Ramirez-Gonzalez R.H., Caccamo M., Bernal W. & Heeney J.L. (2015). – Host subtraction, filtering and assembly validations for novel viral discovery using next generation sequencing data. *PLoS ONE*, **10** (6), e0129059. doi:10.1371/journal.pone.0129059.

50. Schbath S., Martin V., Zytnicki M., Fayolle J., Loux V. & Gibrat J.F. (2012). – Mapping reads on a genomic sequence: an algorithmic overview and a practical comparative analysis. *J. Comput. Biol.*, **19** (6), 796–813. doi:10.1089/cmb.2012.0022.

51. Hatem A., Bozdag D., Toland A.E. & Catalyurek U.V. (2013). – Benchmarking short sequence mapping tools. *BMC Bioinformatics*, **14**, 184. doi:10.1186/1471-2105-14-184.

52. Simpson J.T. & Pop M. (2015). – The theory and practice of genome sequence assembly. *Annu. Rev. Genomics Hum. Genet.*, **16**, 153–172. doi:10.1146/annurev-genom-090314-050032.

53. Koren S. & Phillippy A.M. (2015). – One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly. *Curr. Opin. Microbiol.*, **23**, 110–120. doi:10.1016/j.mib.2014.11.014.

54. Naccache S.N., Federman S., Veeraraghavan N., Zaharia M., Lee D., Samayoa E., Bouquet J., Greninger A.L., Luk K.C., Enge B., Wadford D.A., Messenger S.L., Genrich G.L., Pellegrino K., Grard G., Leroy E., Schneider B.S., Fair J.N., Martinez M.A., Isa P., Crump J.A., DeRisi J.L., Sittler T., Hackett J. Jr, Miller S. & Chiu C.Y. (2014). – A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. *Genome Res.*, **24** (7), 1180–1192. doi:10.1101/gr.171934.113.

55. Scheuch M., Höper D. & Beer M. (2015). – RIEMS: a software pipeline for sensitive and comprehensive taxonomic classification of reads from metagenomics datasets. *BMC Bioinformatics*, **16**, 69. doi:10.1186/s12859-015-0503-6.

56. Wan Y., Renner D.W., Albert I. & Szpara M.L. (2015). – VirAmp: a galaxy-based viral genome assembly pipeline. *GigaScience*, **4**, 19. doi:10.1186/s13742-015-0060-y.

57. Liao Y.C., Lin H.H., Sabharwal A., Haase E.M. & Scannapieco F.A. (2015). – MyPro: a seamless pipeline for automated prokaryotic genome assembly and annotation. *J. Microbiol. Meth.*, **113**, 72–74. doi:10.1016/j.mimet.2015.04.006.

58. Verbist B.M., Thys K., Reumers J., Wetzels Y., Van der Borght K., Talloen W., Aerssens J., Clement L. & Thas O. (2015). – VirVarSeq: a low-frequency virus variant detection pipeline for Illumina sequencing using adaptive base-calling accuracy filtering. *Bioinformatics*, **31** (1), 94–101. doi:10.1093/bioinformatics/btu587.

59. Ratnere I. & Dubchak I. (2009). – Obtaining comparative genomic data with the VISTA family of computational tools. *Curr. Protoc. Bioinformatics.*, Chapter 10, Unit 10.6. doi:10.1002/0471250953.bi1006s26.

60. Skewes-Cox P., Sharpton T.J., Pollard K.S. & DeRisi J.L. (2014). – Profile hidden Markov models for the detection of viruses within metagenomic sequence data. *PLoS ONE*, **9** (8), e105067. doi:10.1371/journal.pone.0105067.

61. Freschi L., Jeukens J., Kukavica-Ibrulj I., Boyle B., Dupont M.J., Laroche J., Larose S., Maaroufi H., Fothergill J.L., Moore M., Winsor G.L., Aaron S.D., Barbeau J., Bell S.C., Burns J.L., Camara M., Cantin A., Charette S.J., Dewar K., Deziel E., Grimwood K., Hancock R.E., Harrison J.J., Heeb S., Jelsbak L., Jia B., Kenna D.T., Kidd T.J., Klockgether J., Lam J.S., Lamont I.L., Lewenza S., Loman N., Malouin F., Manos J., McArthur A.G., McKeown J., Milot J., Naghra H., Nguyen D., Pereira S.K., Perron G.G., Pirnay J.P., Rainey P.B., Rousseau S., Santos P.M., Stephenson A., Taylor V., Turton J.F., Waglechner N., Williams P., Thrane S.W., Wright G.D., Brinkman F.S., Tucker N.P., Tummler B., Winstanley C. & Levesque R.C. (2015). – Clinical utilization of genomics data produced by the international *Pseudomonas aeruginosa* consortium. *Front. Microbiol.*, **6**, 1036. doi:10.3389/fmicb.2015.01036.

62. Camerlengo T., Ozer H.G., Onti-Srinivasan R., Yan P., Huang T., Parvin J. & Huang K. (2012). – From sequencer to supercomputer: an automatic pipeline for managing and processing next generation sequencing data. *AMIA Jt Summits Transl. Sci. Proc.*, 2012, 1–10.

63. Bedoya-Reina O.C., Ratan A., Burhans R., Kim H.L., Giardine B., Riemer C., Li Q., Olson T.L., Loughran T.P. Jr, Vonholdt B.M., Perry G.H., Schuster S.C. & Miller W. (2013). – Galaxy tools to study genome diversity. *GigaScience*, **2** (1), 17. doi:10.1186/2047-217X-2-17.

64. Giardine B., Riemer C., Hardison R.C., Burhans R., Elnitski L., Shah P., Zhang Y., Blankenberg D., Albert I., Taylor J., Miller W., Kent W.J. & Nekrutenko A. (2005). – Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.*, **15** (10), 1451–1455. doi:10.1101/gr.4086505.

65. Anon. (2012). – The data deluge. Editorial. *Nat. Cell Biol.* **14** (8), 775. doi:10.1038/ncb2558.

66. EMBL-European Bioinformatics Institute (2014). – ELIXIR: Data for life. Available at: www.elixir-europe.org/system/files/ELIXIR_2014_brochure_full.pdf (accessed on 25 November 2014).

67. Belák S., Karlsson O.E., Blomström A.L., Berg M. & Granberg F. (2013). – New viruses in veterinary medicine, detected by metagenomic approaches. *Vet. Microbiol.*, **165** (1–2), 95–101. doi:10.1016/j.vetmic.2013.01.022.

68. Granberg F., Vicente-Rubiano M., Rubio-Guerri C., Karlsson O.E., Kukielka D., Belák S. & Sanchez-Vizcaino J.M. (2013). – Metagenomic detection of viral pathogens in Spanish honeybees: co-infection by aphid lethal paralysis, Israel acute paralysis and Lake Sinai viruses. *PLoS ONE*, **8** (2), e57459. doi:10.1371/journal.pone.0057459.

69. Kapgate S.S., Barbuddhe S.B. & Kumanan K. (2015). – Next generation sequencing technologies: tool to study avian virus diversity. *Acta Virol.*, **59** (1), 3–13.

70. Graham R.L. & Baric R.S. (2010). – Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J. Virol.*, **84** (7), 3134–3146. doi:10.1128/Jvi.01394–09.

71. Thiel V., Thiel H.J. & Tekes G. (2014). – Tackling feline infectious peritonitis via reverse genetics. *Bioengineered*, **5** (6), 396–400. doi:10.4161/bioe.32133.

72. Hornyák A., Malik P., Marton S., Doro R., Cadar D. & Bányai K. (2015). – Emergence of multireassortant bluetongue virus serotype 4 in Hungary. *Infect. Genet. Evol.*, **33**, 6–10. doi:10.1016/j.meegid.2015.03.036.

73. Bouwstra R., Heutink R., Bossers A., Harders F., Koch G. & Elbers A. (2015). – Full-genome sequence of influenza A(H5N8) virus in poultry linked to sequences of strains from Asia, the Netherlands, 2014. *Emerg. Infect. Dis.*, **21** (5), 872–874. doi:10.3201/eid2105.141839.

74. Rabaa M.A., Tue N.T., Phuc T.M., Carrique-Mas J., Saylors K., Cotten M., Bryant J.E., Nghia H.D., Cuong N.V., Pham H.A., Berto A., Phat V.V., Dung T.T., Bao L.H., Hoa N.T., Wertheim H., Nadjm B., Monagin C., van Doorn H.R., Rahman M., Tra M.P., Campbell J.I., Boni M.F., Tam P.T., van der Hoek L., Simmonds P., Rambaut A., Toan T.K., Van Vinh Chau N., Hien T.T., Wolfe N., Farrar J.J., Thwaites G., Kellam P., Woolhouse M.E. & Baker S. (2015). – The Vietnam Initiative on Zoonotic Infections (VIZIONS): a strategic approach to studying emerging zoonotic infectious diseases. *Ecohealth*, **12** (4), 1–10. doi:10.1007/s10393-015-1061-0.

75. Carroll M.W., Matthews D.A., Hiscox J.A., Elmore M.J., Pollakis G., Rambaut A., Hewson R., García-Dorival I., Bore J.A., Koundouno R., Abdellati S., Afrough B., Aiyepada J., Akhilomen P., Asogun D., Atkinson B., Badusche M., Bah A., Bate S., Baumann J., Becker D., Becker-Ziaja B., Bocquin A., Borremans B., Bosworth A., Boettcher J.P., Cannas A., Carletti F., Castilletti C., Clark S., Colavita F., Diederich S., Donatus A., Duraffour S., Ehichioya D., Ellerbrok H., Fernandez-Garcia M.D., Fizet A., Fleischmann E., Gryseels S., Hermelink A., Hinzmann J., Hopf-Guevara U., Ighodalo Y., Jameson L., Kelterbaum A., Kis Z., Kloth S., Kohl C., Korva M., Kraus A., Kuisma E., Kurth A., Liedigk B., Logue C.H., Lüdtke A., Maes P., McCowen J., Mély S., Mertens M., Meschi S., Meyer B., Michel J., Molkenthin P., Muñoz-Fontela C., Muth D., Newman E.N., Ngabo D., Oestereich L., Okosun J., Olokor T., Omiunu R., Omomoh E., Pallasch E., Palyi B., Portmann J., Pottage T., Pratt C., Priesnitz S., Quartu S., Rappe J., Repits J., Richter M., Rudolf M., Sachse A., Schmidt K.M., Schudt G., Strecker T., Thom R., Thomas S., Tobin E., Tolley H., Trautner J., Vermoesen T., Vitoriano I., Wagner M., Wolff S., Yue C., Capobianchi M.R., Kretschmer B., Hall Y., Kenny J.G., Rickett N.Y., Dudas G., Coltart C.E., Kerber R., Steer D., Wright C., Senyah F., Keita S., Drury P., Diallo B., de Clerck H., Van Herp M., Sprecher A., Traore A., Diakite M., Konde M.K., Koivogui L., Magassouba N., Avšič-Županc T., Nitsche A., Strasser M., Ippolito G., Becker S., Stoecker K., Gabriel M., Raoul H., Di Caro A., Wolfel R., Formenty P. & Günther S. (2015). – Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa. *Nature*, **524** (7563), 97–101. doi:10.1038/nature14594.

76. Looft T., Allen H.K., Cantarel B.L., Levine U.Y., Bayles D.O., Alt D.P., Henrissat B. & Stanton T.B. (2014). – Bacteria, phages and pigs: the effects of in-feed antibiotics on the microbiome at different gut locations. *ISME J.*, **8** (8), 1566–1576. doi:10.1038/ismej.2014.12.

77. Hoffmann A.R., Proctor L.M., Surette M.G. & Suchodolski J.S. (2015). – The microbiome: the trillions of microorganisms that maintain health and cause disease in humans and companion animals. *Vet. Pathol.*, **53** (1), 10–21. doi:10.1177/0300985815595517.

78. Cadwell K. (2015). – The virome in host health and disease. *Immunity*, **42** (5), 805–813. doi:10.1016/j.immuni.2015.05.003.

79. Andino R. & Domingo E. (2015). – Viral quasispecies. *Virology*, **479–480**, 46–51. doi:10.1016/j.virol.2015.03.022.

80. Lu Z.H., Archibald A.L. & Ait-Ali T. (2014). – Beyond the whole genome consensus: unravelling of PRRSV phylogenomics using next generation sequencing technologies. *Virus Res.*, **194**, 167–174. doi:10.1016/j.virusres.2014.10.004.

81. Land M., Hauser L., Jun S.R., Nookaew I., Leuze M.R., Ahn T.H., Karpinets T., Lund O., Kora G., Wassenaar T., Poudel S. & Ussery D.W. (2015). – Insights from 20 years of bacterial genome sequencing. *Funct. Integr. Genomics*, **15** (2), 141–161. doi:10.1007/s10142-015-0433-4.

82. Sintchenko V. & Holmes E.C. (2015). – The role of pathogen genomics in assessing disease transmission. *BMJ*, **350**, h1314. doi:10.1136/bmj.h1314.

83. Truong A.D., Hong Y.H. & Lillehoj H.S. (2015). – High-throughput sequencing reveals differing immune responses in the intestinal mucosa of two inbred lines afflicted with necrotic enteritis. *Vet. Immunol. Immunopathol.*, **166** (3–4), 116–124. doi:10.1016/j.vetimm.2015.06.008.

84. Plummer E., Buck M.D., Sanchez M., Greenbaum J.A., Turner J., Grewal R., Klose B., Sampath A., Warfield K.L., Peters B., Ramstedt U. & Shresta S. (2015). – Dengue virus evolution under a host-targeted antiviral. *J. Virol.*, **89** (10), 5592–5601. doi:10.1128/JVI.00028-15.

85. Zhou W.G., Gao S.D., Podgorska K., Stadejek T., Qiu H.J., Yin H., Drew T. & Liu L.H. (2014). – Rovac is the possible ancestor of the Russian lapinized vaccines LK-VNIVViM and CS strains but not the Chinese strain (C-strain) vaccine against classical swine fever. *Vaccine*, **32** (49), 6639–6642. doi:10.1016/j.vaccine.2014.09.058.

86. Bálint Á., Balka G., Horváth P., Kecskeméti S., Dan A., Farsang A., Szeredi L., Bányai K., Bartha D., Olasz F., Belák S. & Zádori Z. (2015). – Full-length genome sequence analysis of a Hungarian porcine reproductive and respiratory syndrome virus isolated from a pig with severe respiratory disease. *Arch. Virol.*, **160** (2), 417–422. doi:10.1007/s00705-014-2265-2.

87. Dumarest M., Muth E., Cheval J., Gratigny M., Hebert C., Gagnieur L. & Eloit M. (2015). – Viral diversity in swine intestinal mucus used for the manufacture of heparin as analyzed by high-throughput sequencing. *Biologicals*, **43** (1), 31–36. doi:10.1016/j.biologicals.2014.10.004.

88. Dunne W.M. Jr, Westblade L.F. & Ford B. (2012). – Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. *Eur. J. Clin. Microbiol. Infect. Dis.*, **31** (8), 1719–1726. doi:10.1007/s10096-012-1641-7.

89. Van Borm S., Belák S., Freimanis G., Fusaro A., Granberg F., Höper D., King D.P., Monne I., Orton R. & Rosseel T. (2015). – Next-generation sequencing in veterinary medicine: how can the massive amount of information arising from high-throughput technologies improve diagnosis, control, and management of infectious diseases? *Meth. Molec. Biol.*, **1247**, 415–436. doi:10.1007/978-1-4939-2004-4_30.