Evolution of Next-Generation Sequencing
The discovery of the DNA structure (Watson, 1953) and the subsequent advent of DNA sequencing (Sanger et al., 1977) paved the way for the study of genomes. The DNA sequences of model organisms and humans were elucidated through various innovative techniques. However, it was the next-generation sequencing (NGS) methodologies which accelerated the field of genomics. In the mid-2000’s, the pyrosequencing-based approach was introduced, and over the years, the cost and time of whole-genome sequencing have come down drastically (Margulies et al., 2005). The NGS has also been instrumental in our understanding of health and disease at the molecular level. In this review, we address the merits and limitations of some of the major platforms of NGS and the applications thereof.

Sanger’s dideoxynucleotide chain termination method
In this classical method, the DNA polymerase is tricked to incorporate a modified nucleotide (dideoxynucleotides) with a 3’ hydrogen atom instead of 3’ hydroxy group which is required for chain elongation. Once these dideoxynucleotides are incorporated, the DNA chain elongation terminates. The dNTP molecules used in the reaction are radio-labelled with $^{32}$P, and four dNTPs with one of the modified nucleotides are added per reaction. These samples are resolved on a polyacrylamide gel and the images are captured on X-ray films. The major drawback with this method is the lack of automation. Introduction of thermostable DNA polymerases (Saiki et al., 1985) “D, fluorescent dye-labelled dideoxynucleotides (terminators) (Prober et al., 1987) and capillary sequencing (Marsh et al., 1997) were the major improvements of this method, and these facilitated the unraveling of reference genomes of model organisms. These modifications formed the foundation for NGS.

Next-Generation Sequencing (NGS)
Unlike Sanger’s sequencing, NGS is robust and does not rely on prior knowledge about flanking sequences or

cloning of the DNA fragments. The platform technologies available for NGS can be classified based on the chemistry involved, into (a) sequencing by synthesis and (b) sequencing by ligation and (c) Single molecule sequencing. The different NGS platforms are summarized in Table 1.

(a) Sequencing by synthesis (SBS)
In sequencing by synthesis method, DNA is fragmented and clonally amplified. This is termed as library preparation. During sequencing, addition of nucleotides by the polymerase generates signals which are read by the instrument. The SBS can be further classified as single-nucleotide addition or cyclic reversible termination.

The single-nucleotide addition approach is followed by 454 pyrosequencing (454 Life Sciences, Roche) and Ion Torrent (Life Technologies) techniques. While pyrosequencing method relies on the detection of bioluminescence signal generated due to the pyrophosphate released upon nucleotide incorporation (Margulies et al., 2005), the Ion Torrent platform detects pH change due to release of H+ ions during dNTP incorporation (Rothberg et al., 2011). Ion Torrent utilizes a semiconductor sensor rather than an optical sensor to detect the H+ ions released.

The cyclic reversible termination method utilizes cleavable fluorescent terminator molecules which block the chain elongation (Guo et al., 2008; Ju et al., 2006). Similar to Sanger’s sequencing, a mixture of all four native nucleotides and 3’-blocked dNTP’s are added each time and the unbound dNTP’s are removed by washing. The fluorescent signals from the surface identify the incorporated dNTP. The blocking group and the fluorophore is removed before the beginning of the new cycle. Illumina and GeneReader platform utilize this method for NGS with certain differences. While Illumina platform incorporates each template DNA with the fluorophore-tagged dNTP, the GeneReader platform uses adequate labelled dNTP’s to attain signal for identification (Smith, 2011).

(b) Sequencing by ligation
The Sequencing by Oligonucleotide Ligation and Detection (SOLiD) and Complete genomics platforms follow the sequencing by ligation approach for NGS. The labelled probes and anchor sequences are hybridized and ligated to the DNA strand. The labelled probes can be of one-base or two-base type. The two-base-encoded probes are utilized by the SOLiD platform where each fluorescent signal symbolizes a dinucleotide (Valouev et al., 2008). Despite its early popularity, the short read length of 75 bp greatly limited its use. The Complete genomics technology of the Beijing Genomics Institute utilizes a probe-anchor ligation approach. Here, DNA templates are enriched in solution by ligating it to anchor, circularization and cleavage. Three rounds of ligation, circularization and cleavage add four adapters. The rolling cycle amplification of the circular DNA generates concatamers called nanoballs which are then immobilized on a pattern flow cell and sequenced (Drmanac et al., 2010).

(c) Single-molecule sequencing
Two platforms produce real-time reads with single-molecule. The Pacific Biosciences platform employs phospho-linked fluorescent nucleotides which upon incorporation produce signals monitored through a zero-mode waveguide detector (Eid et al., 2009). The Oxford Nanopore Technology introduced the fourth-generation DNA sequencing technology. Unlike other platforms, the nanopore sequencer does not hybridize or incorporate labelled dNTP. The native single-stranded DNA is passed through a protein pore in the presence of electric current. The DNA translocation into the pore causes a block in the voltage which is measured. The shift in the voltage is characteristic of particular DNA sequence in the pore. About ~70 bp per second can be sequenced using this system. Latest addition to the nanopore technology is the CsgG bacterial amyloid secretion pore based sequencer with higher sequencing accuracy (Brown and Clarke, 2016), achieving DNA translocation rate up to 250 bases per second (Carter and Hussain, 2017). Both Nanopore and Pacific Biosciences platform technologies can read 10 to 100 kb of single-stranded DNA, but error rates are also higher with both the platforms. However, in nanopore sequencing technology the addition of specific hairpin adapter to join the two DNA increase the accuracy of the reads. Even though error rates are higher than other platforms, the nanopore sequencer is a portable and low-cost equipment, besides other advantages such as rapidity of library preparation and real-time data generation.
Table 1: Comparison of different NGS platforms

<table>
<thead>
<tr>
<th>Platforms</th>
<th>Read Length</th>
<th>Throughput</th>
<th>Run time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sequence by synthesis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illumina</td>
<td>25 to 300 bp</td>
<td>540 Mb to 900 Gb</td>
<td>4 hours to 11 days</td>
</tr>
<tr>
<td>454- Life Sciences</td>
<td>400 to 700 bp average reads</td>
<td>35 to 700 Mb</td>
<td>10 to 23 hours</td>
</tr>
<tr>
<td><strong>Sequence by ligation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOLiD</td>
<td>50 to 75 bp</td>
<td>80 to 160 Gb</td>
<td>6 to 10 days</td>
</tr>
<tr>
<td>Complete genomics technology</td>
<td>50 to 100 bp</td>
<td>160 to 320 Gb</td>
<td>24 hours</td>
</tr>
<tr>
<td><strong>Single-molecule real-time long reads</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific Biosciences</td>
<td>8 to 12 Kb</td>
<td>500 Mb to 7 Gb</td>
<td>0.5 to 6 hours</td>
</tr>
<tr>
<td>Oxford Nanopore</td>
<td>Up to 200 Kb</td>
<td>Up to 1.5 Gb</td>
<td>50 minutes to 48 hours</td>
</tr>
</tbody>
</table>

Table adapted from (Goodwin et al., 2016).

Application of NGS in Veterinary Science

The high-throughput sequencing data generated by the various platform technologies are unbiased as it reports the sequence of all nucleotides present in the original sample. It can be applied in multiple ways in veterinary science which are mentioned below.

Conservation of animals

India is a diverse country with a wide variety of domestic and wild animals. The diversity in the germplasm may be one of the reason for disease resistance and other desired characters observed in indigenous breeds. Whole genome sequencing of domestic and wild animal species can help record their genetic information and conserve the native breeds for the future. Sequence data with other gene manipulation techniques like CRISPR-Cas9 may help in improving production of indigenous breeds without losing genetic diversity. Unlike PCR based methods (An et al., 2007; Marin et al., 2009), NGS is a sequence-independent detection technique and can be used for forensic identification of endangered animals.

Characterization of gut microbiota of bovine and other ruminants

Our knowledge about the diversity of the microorganisms present in the rumen is limited because most of them cannot be isolated and cultured under laboratory conditions. Rumen microflora consists of obligate anaerobic fungi, bacteria and protozoa. Till date, six genera (Neocallimastix, Piromyces, Caecomyces, Orpinomyces, Anaeromycoses, and Cyllamyces) of obligate anaerobic fungi have been isolated from rumen (Ho, 1995; Ozkose, 2001), with only two reports on Cyllamyces species (Ozkose, 2001; Sridhar, 2007). Different NGS platforms were utilized to assess the rumen microbial community in cows. While most of the studies were related to the rumen bacterial diversity (Edrington et al., 2012; Jami et al., 2013; Zhou et al., 2017), very few of them studied the composition of rumen fungal community (Dill-McFarland et al., 2017). Identification and characterization of rumen microbiota will help in developing better nutritional systems.

Vaccine industry

The NGS data could be of help at different levels in vaccine industry. In case of bacterial vaccines, the whole genome sequencing can establish the identity of the bacterial strain used in the vaccine. Bacterial strains can also be verified for batch-to-batch consistency as well as to validate the absence of extraneous agents. When compared to bacterial vaccines, NGS has a wider application in viral vaccines. First, the vaccine material is supposed to be devoid of anything other than the antigenic material. To avoid the unintentional introduction of adventitious agents in the vaccine manufacturing process, detailed characterization of seed stocks, final bulk and formulated vaccine is necessary (Kumar et al., 2012). Specifically, these materials must be free from adventitious agents such as bacteria, fungi, mycoplasma/spiroplasma, mycobacteria, rickettsia, protozoa, parasites, transmissible spongiform encephalopathy (TSE) agents and viruses. Despite following stringent quality protocols, contamination of vaccines with extraneous agents can happen. Notable among these are (a) simian virus-40 contamination of the poliovirus and adenovirus vaccines (Geissler et al., 1985; Hilleman, 1998; Lewis, 1998; Sangar et al., 1999), (b) bacteriophage (pV-1) contamination in poliovirus vaccine and measles, mumps and rubella (MMR) vaccine (Haselkorn et al., 1978), (c) reverse transcriptase activity in MMR vaccine (Pyra et al., 1994; WHO, 1998), (d) bovine viral diarrhea virus in mumps, mumps and measles, and MMR vaccines
We regularly isolate and type bluetongue virus (BTV) from field samples through conventional cell culture and RT-PCR, respectively. Many times, we encounter difficulty in typing BTV isolates due to infections with mixed serotypes and emergence of a new variant which cannot be typed using the primers designed to identify the serotype by RT-PCR. There are 27 BTV serotypes identified until now and geographical variants called topotypes are present within a serotype. The genomic variation due to diverse serotypes and topotypes adds to the complexity of typing (Reddy et al., 2016). We routinely use NGS for typing of BTV isolates which could not be typed due to lack of reagents. Now, NGS is becoming popular to identify new serotypes, topotypes, variants and reassortants (Boyle et al., 2012; Studer et al., 2002), (e) avian retroviruses in yellow fever vaccine (Hussain et al., 2003), (f) avian leucosis virus in measles and mumps vaccine (Tsang et al., 1999), (d) porcine circovirus 1 contamination of rotavirus vaccine (Victoria et al., 2010). Several veterinary vaccines have been reported to be contaminated with adventitious agents such as (a) feline endogenous retrovirus in feline and canine attenuated vaccines (Yoshikawa et al., 2014), (b) reticuloendotheliosis virus in Marek’s disease vaccine (Takagi et al., 1996), (c) bovine viral diarrhoea virus in swine fever vaccines (Wensvoort and Terpstra, 1988), and (d) bluetongue virus in canine vaccines (Akita et al., 1994; Evermann et al., 1994; Wilbur et al., 1994), lumpy skin disease and sheep pox vaccines (Bumbarov et al., 2016).

Apart from identifying the adventitious agents, NGS can also establish the identity of the vaccine stains. NGS is a highly sensitive method for pathogen identification typing. We have used Ion torrent platform to sequence RNA and DNA isolated from pharyngeal mucosa during a disease outbreak investigation in commercial poultry. Apart from genome sequence of the causative agents, we could identify the genomes of some of the live vaccines used in poultry, several bacteriophages, and viruses of fish and maize. The latter two genomes probably originated from the poultry feed (unpublished data), and demonstrate the power of this technology in diagnostics. This makes NGS a more suitable test for identification of disease causing agents.

NGS is also helpful in identification of mixed infections especially of those that reduce production or are immunosuppressive with no clear clinical symptoms. For example, pyrosequencing of DNA from post-weaning multisystemic wasting syndrome (PWMS) affected pig lymph node samples revealed the presence of torque teno virus and a novel boca-like parvovirus apart from the causative agent porcine circovirus 2 (Blomstrom et al., 2010; Blomstrom et al., 2009). These complex conditions necessitate the regular use of NGS to establish concrete relation between pathogens and disease. Further, NGS is a good technology to identify, characterize and understand the epidemiology of non-cultivable infectious agents such as Jaagsiekte, a retrovirus of sheep.

Reversion to virulence is one of the major concerns with live attenuated vaccines. Different nucleic acid tests are used to identify the virulent markers in vaccine stains. In general, limited data is generated using nucleic acid based techniques i.e., Mutant Analysis by PCR and Restriction Enzyme Cleavage (MAPREC) test for poliovirus. NGS methods are applied to determine the mutations necessary for attenuation of poliovirus (Victoria et al., 2010). Similar methods can be used to detect virulent markers in vaccine viruses to improve the safety of vaccines. Several attenuated vaccines are employed for veterinary use, and their characterization by NGS can improve vaccine safety.

**Diagnosis of infectious diseases and monitoring of outbreaks**

Identification of infectious disease etiology is generally done by microscopy, microbiological, virological, immunological or molecular methods. Compared to the other methods, molecular methods have higher sensitivity and specificity. However, knowledge about genome sequence of the etiological agent is a prerequisite to use molecular methods such as PCR or RT-PCR. For differential diagnosis by classical methods, several tests need to be done to confirm the etiological agent. In addition, availability of reagents such as special media, cell lines, antibodies, primers, and availability of validated methods are the major constrains in diagnosis of infectious agents. Quick and accurate diagnosis is important during disease outbreaks to devise strategies to control the spread of the disease in question. Unlike classical methods of diagnosis, where availability of agent-specific diagnostics is a prerequisite, NGS is a generic method that can be applied to identify any type of infectious agent i.e., viruses, bacteria, fungus, protozoa.

NGS is a good technology to identify, characterize and understand the epidemiology of non-cultivable infectious agents such as Jaagsiekte, a retrovirus of sheep.
In the Indian context, emergence of nephropathogenic avian infectious bronchitis virus (Bayry et al., 2005), porcine circovirus2 (Karuppannan et al., 2016), re-emerging BTV serotypes (Hemadi et al., 2016; Krishnajyothi et al., 2016; Rao et al., 2016), avian influenza (Bhat et al., 2017; Nagarajan et al., 2017) and new variants of foot and mouth disease virus (Mahapatra et al., 2016; Rudreshappa et al., 2012; Yuvaraj et al., 2013), and porcine respiratory and reproductive syndrome virus (PRRSV) (Rajkhowa et al., 2015) are some examples where NGS would have helped.

Antimicrobial resistance to antibiotics is seen as one of the greatest threats in Human and veterinary practice. Bacteria can genetically evolve by introducing mutations in genes that encode proteins that are targets for antibiotics and by acquiring foreign DNA coding for resistance determinants. Various NGS methods were applied to detect antibiotic resistance markers in Escherichia coli, Bacillus anthracis, Yersinia pestis, Francisella tularensis and Mycobacterium tuberculosis (Daum et al., 2012; Stefan et al., 2016; Veenemans et al., 2014). Using NGS data for detection of the many genes involved in antibiotic resistance/sensitivity of different strains. Species’ general of bacteria present in a clinical sample yields far more information from a single test than other methods like culture, biochemical identification, PCR, microarrays and traditional antibiograms and does not require prior knowledge of the resistance phenotype of the isolate(s). However, for most species of bacteria, insufficient data about antibiotic resistance/sensitivity genes and mechanism involved is known and hence, data need to be generated to use NGS-inferred antibiotic resistance/sensitivity to guide clinical decisions (Ellington et al., 2017). The current cost and speed of NGS is prohibitive to use NGS for testing genotypic antibiograms. However, with the advent of faster and affordable NGS methods like Nanopore, it is possible to identify antibiotic resistance markers in clinical samples that can help in therapeutic decisions in critical cases (ONT, 2017).

Although NGS cannot be used as routine test for disease diagnosis due to cost and time required, it will become an unavoidable alternative at least for identification of the pathogen. With the development of new portable and affordable nanopore technology of NGS, it is possible to use this technology as routine diagnostic test in the near future.

Identification of exotic diseases

Diseases can spread from one country to another through various routes and one of the controllable routes is the import of animals and animal products. To control disease spread by movement of animals and animal products, countries implement import protocols mostly guided by World Organization for Animal Health (OIE). As per the OIE terrestrial animal health code, the live animal imports require the health certificate of the animals for diseases (mostly infectious diseases) from the exporting country. To control transboundary movement of diseases, pre-export testing and post-import quarantine is being practiced by most of the countries. Most of these tests are serology- or nucleic acid-based. In case of RNA viruses, where nucleic acid diversity is more, primers designed for one geographical area may not detect the viruses from other geographical area (Maan et al., 2012a; Mertens et al., 2007; Reddy et al., 2016). Apart from that, the test needs to be done against pathogens. For example, import of live pigs into India requires the pre-export screening of the animals for African swine fever, porcine brucellosis, classical swine fever, porcine respiratory and reproductive syndrome, porcine epidemic diarrhoea virus, porcine delta corona virus, transmissible gastroenteritis, swine vesicular disease, swine influenza A virus, Tescho virus encephalomyelitis, and Aujeszky’s disease (DAHD, 2017). The animals are quarantined in India for 30 days and observed for any symptoms of disease before handing it over to the consignee.

Despite pre-import and post-quarantine tests, several incidences of entry and establishment of causative agents from other countries are reported from India. These include exotic serotypes and topotypes of BTV (Gollapalli et al., 2012; Maan et al., 2012b; Rao et al., 2016; Rao et al., 2012a; Susmitha et al., 2012), and highly pathogenic strain of PRRSV similar to Chinese isolates, (Rajkhowa et al., 2015). Currently, lumpy skin disease (LSD) is being considered as one of the future threats to Indian livestock industry. LSD, once considered as African disease, has now spread to throughout the middle-eastern countries in the last quarter century (Ali et al., 1990; House et al.,
1990; OIE, 2017; Yeruham et al., 1995) and affected some of the European and central Asian countries (OIE, 2017). Recent reports of LSD positive samples from some districts of Odisha (OIE, 2019) is a cause of concern. There is an immediate need for an effective control and advanced surveillance measure to contain the spread of the disease. LSDV, the causative agent of LSD is closely related to sheep pox and goat pox viruses and it is difficult to differentiate these viruses by PCR-based or serological methods. It is further complicated by the presence of a clinically similar disease called pseudo-lumpy skin disease (PLSD) caused by bovine herpesvirus-2. We have used NGS to identify BHV-2 from skin samples of cattle affected with PLSD and successfully identified the causative agent, and similar methods can be used to screen suspected samples for quick and accurate diagnosis.

Recent reports of African swine fever (ASF) outbreaks from various Asian countries (OIE, 2019) pose a threat to India. The current wave of ASF started from China in August 2018. By 2019, the outbreaks of ASF were recorded from Mongolia (in January), Vietnam (in February), Cambodia (in April), Lao PDR (in June) and Myanmar (in August). India shares its eastern border with Myanmar, there are chances that the African swine fever (ASF) may seep in to the north-eastern states of India. As the pig population is more in the north-eastern states of India, the advent of ASF will have substantial economic impact.

Introduction of new pathogens into a naïve population may cause huge economic losses and it is important to identify and contain the exotic diseases before they get established and spread through the length and breadth of the country. Here, NGS has a clear advantage over other classical methods of pathogen identification due to its speed, sensitivity and generic nature.

CONCLUSION

NGS technology is likely to play an important role in the field of veterinary medicine and animal husbandry. In veterinary services, NGS is expected to play a vital role in disease diagnosis and control. Although most of the current NGS technologies are expensive to establish and require highly skilled molecular biologists and bioinformaticians, advent of technologies like Nanopore, which is fast and portable may become a disruptive technology in veterinary and medical diagnosis, replacing most of the current technologies being used in this field. This portable platform has played a crucial role in the monitoring of outbreaks of Salmonella enterica (Quick et al., 2015) and Ebola hemorrhagic fever (Quick et al., 2016). Unlike platforms such as Illumina, sequence error rates are higher in nanopore-based sequencing, but for the purpose of diagnosis, errors in sequence are tolerable provided the test could precisely identify the causative agent. With the traits of a cost-effective, rapid and portable alternative for DNA sequencing, technologies like nanopore sequencing will be necessary for disease monitoring and diagnostic laboratory sooner or later.

REFERENCES


relationships and possible routes of entry into Australia. J. Virol, 86: 6724-6731.
DAHD 2017. Veterinary certificate for import of live swine for breeding purposes into India In, D.F. Department of Animal Husbandry, ed.

Journal of Animal Research: v.9 n.6, December 2019


Kumar, D., Beach, N.M., Meng, X.J. and Hegde, N.R. 2012. Use of PCR-based assays for the detection of the adventitious agent porcine circovirus type 1 (PCV1) in vaccines, and for confirming the identity of cell substrates and viruses used in vaccine production. J. Virol. Methods, 179: 201-211.


Application of NGS in veterinary science


Journal of Animal Research: v.9 n.6, December 2019


