Review Article

A review of the application of Next Generation Sequencing (NGS) in wild terrestrial vertebrate research

ABSTRACT

In era of sequencing revolution, scientists seek for knowledge about the ever-expanding field of technology, Next Generation Sequence (NGS) to be applied in their research due to its high reliability and rate of discovery. What is NGS? To obtain a detailed understanding about NGS, it is required to look back the history of sequencing and how the NGS stepped into life science. This review paper gives an overview of NGS projects in wild terrestrial vertebrate including applications such as whole genome sequencing and metagenomics.

Keywords: Next-generation sequencing (NGS); wild terrestrial vertebrate; whole genome sequence; metagenomics

INTRODUCTION

In 1953, the double helix structure of DNA or deoxyribonucleic acid was discovered [1]. Later, the first DNA sequencing was perceived after fifteen years elapse in 1968. In line with the development of chemical method [2], modern DNA sequencing began in 1977 and in the same year the first DNA sequence (phage ϕX174) was completed by Sanger and Coulson [3], which demonstrated the sequences were capable to give profound insights into genetic organization. Sanger sequencing was a tool for deciphering complete genes and also entire genomes [4] until the Human Genome Project drafted in year 2001. Although, the first complete cellular genome sequences from bacteria appeared in 1995 [5,6,7], the drastic impact on Next Generation Sequence (NGS) began only after the completion of Human Genome Project in 2003.

NEXT GENERATION SEQUENCING

In era of sequencing revolution, NGS is the most demanding technology that getting greater popularity day by day. NGS has influenced almost every single field in applied and life science. NGS technology utilizes distinct sequencing biochemistry approaches and it is mainly accentuated by its ability to simultaneously perform millions of sequencing reactions. Among the most widely used applications of NGS are the whole genome de novo sequencing, whole genome re-sequencing and also exome, targeted, whole transcriptome, metagenome and epigenome sequencing [8,9].

Although NGS has a wide range of biological outcomes, the cost per sample analyses often limit the use of this technique. Fortunately, the recent development in high-throughput sequencing techniques (Table 1) has reduced the burden. For example, sequencing cost has massively reduced from $5,292.39/Mb in 2001 to $0.06/Mb by April 2013 [10]. It is estimated that the sequencing costs will further reduce with precipitous dropping per-base cost with advancing techniques.

Basically, NGS sequencing has expended from second-generation to the next two levels, third-generation (3G) and fourth-generation (4G) (Table 1). These techniques allowed the genomics to move from platforms that required Polymerase Chain Reaction (PCR) amplification of the template in prior of sequencing to without a prior amplification step as in third-generation sequencing techniques, and to a more refined level of the fourth-generation. Even though, NGS techniques are quite diverse but conceptually they are similar. The preparation of library includes random shearing of DNA followed by ligation with common adaptors (Table 1).

NGS IN WILD TERRESTRIAL VERTEBRATE RESEARCH

Whole genome de novo sequencing, re-sequencing and targeted sequencing

Wild terrestrial vertebrate are defined as undomesticated and free-ranging animals include reptiles, amphibians, birds, and mammals [11]. Since past few decades, researches on wild terrestrial vertebrate have been ascended. Numerous researchers and non-governmental organization (NGO's) investing millions of money in wild terrestrial vertebrate' projects with aim to protect and conserve them from extinct. A repository of wild terrestrial vertebrate' genome sequences is crucial for phylogeography [12,13], demographic history [14], multilocus population genomics [15,16], adaptation studies [17,18], and conservation efforts [19,20].

After the completion of the Human Genome Project, the first nonhuman primate to have its genome sequenced was chimpanzee (*Pan troglodytes verus*) from West Africa. The Chimpanzee Sequencing and Analysis Consortium (2005) reportedly utilized cloned primary blood lymphocyte DNA to generate sequence reads that were assembled via de novo assembly approach and also by aligning sequence reads with the human genome [21]. Their assembly covered 94% of the chimpanzee genome, with a consensus length of 2.7 gigabases (Gb). A genome-wide comparison of the draft chimpanzee genome with the human genome revealed 13,454 pairs of unambiguous human and chimpanzee orthologue genes, of which 29% the encoded proteins were discovered to be identical. The team also compiled a list of human, mouse, rat, and chimpanzee genes with unambiguous gene orthology. Comparisons of West African and Central African chimpanzee sequence reads were also performed to locate polymorphic positions within and between these individuals, which in turn show that the heterozygosity rate of the Central African chimpanzees to be two times the heterozygosity rate of the West African chimpanzees [21].

The next hominid nonhuman primate to be sequenced was the Sumatran orangutan (*Pongo abelii*). It is listed by the International Union for Conservation of Nature (IUCN, 2018) as critically endangered, because its habitat increasingly destroyed and fragmented by human encroachment. A female orangutan specimen housed in Gladys Porter Zoo, Brownsville, Texas was sequenced using the whole-genome shotgun sequencing approach, with an average of 5.5-fold coverage across the 3.08 Gb consensus assembly [22]. The orangutan genome assembly was facilitated by referring to human gene models as well as orangutan complementary DNA (cDNA) [22]. Genomic-wide nucleotide identity comparisons and single

nucleotide polymorphism (SNP) between the Sumatran and Bornean orangutan revealed that the Sumatran orangutan was more diverse than their Bornean counterpart, despite having a smaller population size [23]. Further SNP analyses of the orangutan autosomal and mitochondrial genome was carried out whereby they observed that the majority of the orangutan genome underwent negative selection throughout their evolutionary history [24].

Among the old world monkeys, special attention is paid to the rhesus macaque (*Macaca mulatta*) and the cynomolgus macaque (*Macaca fascicularis*) due to their indispensable use as nonhuman primate model organisms. Their similarities to humans with regards to their biology, behaviour, physiology, and genetics make them choice selections for biomedical research, as well as drug response studies [25,26]. Extensive whole genome sequencing endeavors have been undertaken to sequence not only their respective genomes, but genomes of macaques originating from various geographical locations [25,26,27,28]. Another old world monkey, the proboscis monkey (*Nasalis larvatus*), though not a model organism, has had its genome sequenced and assembled by the Proboscis Monkey Functional Genome Consortium from The Department of Zoology, Universiti Malaysia Sarawak in 2014. The team utilized a combination of the 454 Sequencing and Illumina Hiseq sequencing platforms with 290-fold genome coverage across 2.67 Gb. Assembly of the proboscis monkey genome (GenBank assembly accession: GCA_000772465.1) was performed via a reference-guided assembly with the rhesus macaque genome as a reference.

The Amur tiger (*Panthera tigrisa ltaica*) genome was sequenced using the Hiseq 2000 platform with a 83.5-fold coverage across the de novo assembled 2.4 Gb scaffolds [29]. An alignment of the Amur tiger genome sequence with domestic cat genome sequence showed that the tiger genome was 95.6% similar to that of the domestic cat genome. Given the similarity, further comparisons between the tiger and domestic cat genomes revealed 103 orthologous gene families shared uniquely between the tiger and domestic cat. Concurrently with the Amur tiger genome assembly, also sequenced the genomes of four other big cats, including the Bengal tiger (*Panthera tigris tigris*), the African lion (*Panthera leo*), the white African lion (*Panthera leo krugeri*), and snow leopard (*Panthera uncia*) [29]. Comparative genomic analysis between the big cat and domestic cat, revealed a high genomic synteny as well as similar repeat compositions in the genomes, indicating stable genome conservation and similar genome architecture among the big cats and domestic cats. Both the Amur tiger and Bengal tiger are listed as endangered by the IUCN, while the Malayan tiger (*Panthera tigris jacksoni*) as critically endangered [29]. Presently, the Malayan tiger faces imminent extinction due to habitat fragmentation and commercial poaching [30], and has yet to have its genome sequenced. With there being an estimated 500 Malayan tigers left in the wild, an assessment of the Malayan tiger genome and their population genomics is critical for conservation efforts of wild and captive Malayan tigers [31].

Part of the Genome 10k Project, the Avian Genome Consortium [32,33] sequenced the genome of the rhinoceros hornbill (*Buceros rhinoceros silvestri*) together with the genomes of 47 other avian species. The hornbill genome was sequenced using the Illumina Hiseq 2000 sequencing platform, with 35-fold coverage across 1.08 Gb. Assembly was carried out de novo, annotations of protein coding gene was based on chicken, zebra fish, and human gene sets. Comparative genomics analyses among the 48 avian genomes revealed that the avian genomes to be reduced in size due to reduced introns, fragmented microchromosomes, reduced repeat transposon activity, shorter protein coding genes, and large segmental deletions. The avian protein coding genes are on average 50% and 27% shorter than the mammalian and reptilian protein coding genes respectively. In addition, the avian genomes have a reduced number of genes, about 70% of the number of genes found in the human genome. Further phylogenomics analyses utilizing an alignment of around 41.8

million bp nucleotide data sets consisting mainly of orthologous exons from 8251 syntenic protein coding genes, introns from 2516 of these genes, and 3769 ultra conserved elements to infer evolutionary relationships between the 48 avian genomes revealed contradictions in avian phylogenies inferred from morphological characters, DNA-DNA hybridization, and mitochondrial genomes [32,33].

Metagenomics

Metagenomics application is popularly used in studies of assemblage of microorganisms in microbial ecology [34]. Plenty number of bioinformatics tools have been developed for metagenomics analysis. In general, bioinformatics tools are command-based programs, which run on Linux or Ubuntu Operating System but there are few programs developed in Window OS for user-friendly such as MEGAN4 and MG-RAST. Each of the metagenomics tools has their specific functions (Table 2). Numerous of metagenomics studies have been reported to study environmental samples [35] such as hair, feacal, soil and water samples. However, very limited metagenomics studies/research have been studied in wild terrestrial vertebrate.

Metagenomics analysis using a 454 GS Junior Instrument (Roche) enabled the detection of the presence of novel viruses or virus variants of theilovirus, phleboviruses, amdovirus, kobuvirus and picobirnaviruses in 10 different species of wild small carnivores including the American mink (*Neovison vison*), European mink (*Mustela lutreola*), European polecat (*Mustela putorius*), European pine marten (*Martes martes*), stone marten (*Martes foina*), Eurasian otter (*Lutra lutra*) and Eurasian badger (*Meles meles*) from the family of Mustelidae; common genet (*Genetta genetta*) of the family of Viverridae; red fox (*Vulpes vulpes*) of the family of Canidae and European wild cat (*Felis silvestris*) of the family of Felidae) living in the northern part of Spain [36]. Metagenomics approach was used to understand the effective surveillance on wildlife-associate zoonoses in China, especially in bats [37]. The genomes of bats were sequenced using Solexa sequencing technology and nearly 1.2 trillion useful reads were generated. The raw reads were assembled into 4872 contigs whereby 36 viral families were annotated, which consist of 19 vertebrate virus families, 6 plant virus families, 4 insect virus families and 4 phages [37].

Bushpigs (*Potamochoerus larvatus*) have been identified as the potential natural reservoirs for African swine fever virus, however is less known about what other viruses might be carried by bushpigs [38]. Moreover, there is a chance for interaction and sharing of pathogens with domestic relatives due to their habitat and movement at the boundary between the national parks and the farmland [39]. Thus a viral metagenomic study was carried out to determine whether bushpigs are carriers of known and/or unknown porcine viruses using sera samples [39]. The presence of PPV4 and novel TTSuV-1 and 2 variants were identified in the samples. The genetic relationships of these viruses and their distribution in both domestic pigs and in wildlife can be defined by further sequence analysis [39].

The red foxes (*Vulpes vulpes*) which distributed across Northern Hemisphere ranging from urban areas and farmlands to remote forests belongs to most widespread member of the order Carnivora, where these animals are known as carriers for number of pathogens that are harmful to humans, including *Echinococcus multilocularis* and, in certain parts of the world, rabies virus [40]. Using metagenomic approach, the sequences obtained from fecal samples were detected with similarity to the sequences of Parvovirus and Hepevirus together with other viruses like picobirnavirus and astrovirus, however, the majority of the sequences had relatively low homology to known viruses [40].

Wild boars from an animal park in Hungary were subject to viral metagenomics analysis and complete genome sequencing [41]. The study identified Astrovirus sequence contigs in 50% (5/10) of fecal samples by metagenomic analysis. Based on the complete astrovirus genome sequence, this study showed wild boar may be a reservoir for astroviruses that infecting pigs and vice versa, and the PAstV-4 and WB lineages of astroviruses may have a single common origin because of their genetic similarities [41].

Metagenomic analysis of the viral flora in feaces of pine marten and European badger was carried out in the Netherlands [42]. In this project, researchers have used next-generation sequencing a 454 GS Junior instrument (Roche) technology to gain insight. Based on the metagenomics results, the total seven samples indicated the presence of bacteriophages from the order Caudovirales and family Microviridae. The result for pine marten, eukaryotic viruses with homology to kobuvirus from the Picornaviridae family, bocavirus from the Parvoviridae family, torque teno virus from the Anelloviridae family, and Sclerotiniasclerotiorumhypovirulence-associated DNA virus 1 (SSHADV-1) from the Geminiviridae-like family were identified [42]. Meanwhile, eukaryotic viruses with homology to Bombyxmoricypovirus from the Reoviridae family, columbid circovirus from the Circoviridae family, canine distemper virus from the Paramyxoviridae family, SSHADV-1 from the Geminiviridae-like family, and torque teno virus from the Anelloviridae family were detected in European badgers [42]. In addition, sequences with homology to viruses from the families Paramyxo- and Picornaviridae were also detected.

Zoonotic infections by rodents to human are very common due to their frequent contact [43]. A study was attempted to sequence for viral diversity in feces of 105 wild rodents (mouse, vole, and rat) collected in California and Virginiaand discovered a declining rate of sequences related to the mammalian viruses families Circoviridae, Picobirnaviridae, Picornaviridae, Astroviridae, Parvoviridae, Papillomaviridae, Adenoviridae, and Coronaviridae [44]. Several potentially new viral families related to the Circoviridae and Picornaviridae were characterized [43]. First murine astrovirus genome, papillomavirus genome, adenovirus and adenovirus-associated virus were also sequenced. This study also identified a large fraction of insect viruses namely the Densoviridae, Iridoviridae, Polydnaviridae, Dicistroviriade, Bromoviridae, and Virgaviridaeand plant viruses such asNanoviridae, Geminiviridae, Phycodnaviridae, Secoviridae, Partitiviridae, Tymoviridae, Alphaflexiviridae, and Tombusviridae families from which they rodents obtained through their diet [43].

With the advancement in metagenomics or NGS technologies, recent studies have demonstrated the existence of enormous virus diversity among wild terrestrial vertebrate including those uncharacterized viruses through conventional methods as per discussed above. As far as the animal conservation and welfare in concern, the expansion of knowledge of the virus diversity present in wild terrestrial vertebrate, as well as the potential transmission of these viruses to domestic animals or humans are essential.

CONCLUSION

In a world after the first human genome was successfully drafted [44], the labour and cost of sequencing a genome has reduced remarkably with the introduction of next-generation sequencing (NGS) platforms such as the Illumina's Hiseq, Roche's 454 pyrosequencing, ABI's SOLiD Platform, and various other up and coming platforms (Table 1). NGS is advantageous over Sanger sequencing in that a larger amount of data can be obtained in a much shorter period of time and at a fraction of the cost. These benefits enable genomes of non-model organisms to be sequenced. Asides from biomedical and pharmacological studies, phylogenomics and comparative genome studies will easily benefit from the large

amount of data that are easily obtained from utilizing the NGS platforms, either from the genome proper or from whole mitochondrial genome. As of writing, there are at least 3000 eukaryote genomes at various levels of assemblies listed in NCBI. Wildlife animal are extinct day by day due illegal hunting, deforestation, defaunation and other factors. For future generation, the knowledge of NGS is very important in a way understanding the genome of an animal, genetic composition and diseases that could effect the organism or vice versa. Many researches should be studied in wildlife animal in order to preserve our nature and future.

Table 1: High-throughput sequencing methods

303

METHOD	ADVANTAGES	DISADVANTAGES	REMARK
Second-generation sequencing techniques			
454 sequencing	Generate long read lengths and relatively fast run times of the instrument	Poor interpretation of homopolymers leading to errors	First introduced NGS technique [45,46]
Illumina (Solexa) Genome Analyzer	Wide use analyser and short read length method a	Irregular incorporation of incorrect dNTPs by polymerases	Low multiplexing ability [47]
ABI SOLiD system	Reduction in error rates relative to Illumina NGS system	Have long run times and need for 2-20 µg DNA	Driven by DNA ligase than polymerase [48]
HiSeq 2000 (Illumina, CA, USA)	Requires less sample < 1 µg	75 (35-100) bp read lengths. More false positives	Addition of fluorescent-labeled nucleotides [49]
Polonator G.007	Decode the base by single-base probe nonamers	In coverage, positive selection rate adequate false- SNP	Ligation based sequencer [50]
Ion Torrent Sequencing	First platform to eliminate cost and complexity with 4-color optical detection used by other NGS platforms	High accuracy and short run time	Non-optical DNA sequencing [51]
SLAF-seq	De novo SNP discovery with reduced cost and high accuracy	Needs complex instrument	Double barcode system ensures simultaneous genotyping of large populations [52]
Third-generation sequencing techniques			
PacBio RS (Pacific Biosciences , CA, USA)	No amplification of template DNA required, real- time monitoring of nucleotide incorporation	High error rates and low reads	Generates long-read lengths 800-1000 bp [53]
HeliscopeT M Sequencer	Nonbiased DNA sequence	High NTP incorporation error rates	Single molecule sequencing [54]
Fourth-generation sequencing techniques			
Oxford Nanopore	Fastest sequencer whole- genome scan within 15 min	Not much data available, high cost perMb	Expanding technique [55]

304 **Table 2: Program and tools for Metagenomics application**

No	Software name	Operating system	Functions/application
1	MG-RAST (an automated analysis platform for metagenomes)	Windows	• Provides quantitative insights into microbial population information based on query sequences data [56]. • User-friendly.
2	IMG/M or known as "The Integrated Microbial Genomes (IMG) system	Linux/Ubuntu	• Command-based
3	CAMERA	Windows	• This tool specifically developed to study microbial ecology and a centralized database for marine microbes [57]
4	CARMA	Windows	• Specifically developed for characterizing the taxonomic composition and genetic diversity of short-read metagenomes [58]
5	MOTHUR	Linux/Ubuntu	• Complete package to answer microbial ecology questions because composed of the development tools such as ARB, DOTUR, SONS, LIBSHUFF, UniFrac, Statistical package, phylogenetic tool. • This tool analyses 222,000 sequences less than 2 hours in personal computer compared to other existing tools where can analyses 102 to 104 sequences only. Besides, this software is very flexible and easy to maintain [59]

REFERENCES

1. Watson JD, Crick FH. Molecular structure of nucleic acids. Nature. 1953;171(4356):737-738.
2. Maxam AM, Gilbert W. A new method for sequencing DNA. Proc Natl Acad Sci. 1977;74(2):560-564.
3. Sanger F, Coulson A, Friedmann T, Air G, Barrell B, Brown N et al. The nucleotide sequence of bacteriophage φX174. J Mol Biol. 1978;125(2):225-246.
4. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci. 1977;74(12);5463-5467.
5. Sulston J, Du Z, Thomas K, Wilson R, Hillier L, Staden R et al. The *C. elegans* genome sequencing project: a beginning. Nature. 1992;356(6364):37-41.
6. Adams MD, Kerlavage AR, Fleischmann RD, Fuldner RA, Bult CJ, Lee NH et al. Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. Nature. 1995;377(6547 Suppl):3-174.
7. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science, 1995;269(5223):496-512.
8. Mardis ER, Wilson RK. (2009). Cancer genome sequencing: a review. Hum Mol Genet. 2009;18(R2):R163-R168.
9. Bras J, Guerreiro R, Hardy J. Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. Nat Rev Neurosci. 2012;13(7):453-464.
10. Spurdle AB, Lakhani SR, Da Silva LM, Balleine RL, Goldgar DE. Bayes analysis provides evidence of pathogenicity for the BRCA1 c. 135-1G> T (IVS3-1) and

BRCA2 c. 7977-1G> C (IVS17-1) variants displaying in vitro splicing results of equivocal clinical significance. Hum mutat. 2010;31(2):E1141-E1145.
11. Yarrow G. Wildlife and wildlife management. Fact Sheet. 2009; 36.
12. Morin PA, Archer FI, Foote AD, Vilstrup J, Allen EE, Wade P et al. Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. Genome Res. 2010;20(7):908-916.
13. McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. Applications of next-generation sequencing to phylogeography and phylogenetics. MolPhylogenetEvol. 2013;66(2):526-538.
14. Zhao S, Zheng P, Dong S, Zhan X, Wu Q, Guo X et al. Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. Nat Genet. 2013;45(1):67-71.
15. McCormack JE, Maley JM, Hird SM, Derryberry EP, Graves GR, Brumfield RT. Next-generation sequencing reveals phylogeographic structure and a species tree for recent bird divergences. MolPhylogenet and Evol. 2012;62(1):397-406.
16. Ellegren H. Genome sequencing and population genomics in non-model organisms. Trends EcolEvol. 2014;29(1):51-63.
17. Zhan X, Pan S, Wang J, Dixon A, He J, Muller MG et al. Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. Nat Genet. 2013;45(5):563-566.
18. Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan D et al. Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. Genome Res. 2014;24(8):1308-1315.
19. Steiner CC, Putnam AS, Hoeck PE, Ryder OA. Conservation genomics of threatened animal species. Annu Rev AnimBiosci. 2013;1(1):261-281.
20. Der Sarkissian C, Ermini L, Schubert M, Yang MA, Librado P, Fumagalli M et al. Evolutionary genomics and conservation of the endangered Przewalski's horse. Curr Biol. 2015;25(19):2577-2583.
21. Sequencing TC, Consortium A. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature. 2005;437(7055): 69-87.
22. Locke DP, Hillier LW, Warren WC, Worley KC, Nazareth LV, Muzny DM et al. Comparative and demographic analysis of orang-utan genomes. Nature. 2011;469(7331):529-533.
23. Meijaard E, Wich S. Putting orang-utan population trends into perspective. Curr Biol. 2007;17(14):R540.
24. Ma X, Kelley JL, Eilertson K, Musharoff S, Degenhardt JD, Martins ALet al. (2013). Population genomic analysis reveals a rich speciation and demographic history of orang-utans (*Pongo pygmaeus* and *Pongo abelii*). PLoS One. 2013;8(10):e77175.
25. Ebeling M, Küng E, See A, Broger C, Steiner G, Berrera M et al. Genome-based analysis of the nonhuman primate *Macacafascicularis* as a model for drug safety assessment. Genome Res. 2011;21(10):1746-1756.
26. Yan G, Zhang G, Fang X, Zhang Y, Li C, Ling F et al. Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. Nat Biotechnol. 2011;29(11):1019-1023.
27. Higashino A, Sakate R, Kameoka Y, Takahashi I, Hirata M, Tanuma R et al. Whole-genome sequencing and analysis of the Malaysian cynomolgus macaque (*Macacafascicularis*) genome. Genome Biol. 2012;13(7):R58.
28. Zimin AV, Cornish AS, Maudhoo MD, Gibbs RM, Zhang X, Pandey S et al. A new rhesus macaque assembly and annotation for next-generation sequencing analyses. Biol Direct. 2014;9:20.
29. Cho YS, Hu L, Hou H, Lee H, Xu J, Kwon S et al. The tiger genome and comparative analysis with lion and snow leopard genomes. Nat Commun. 2013;4:2433.

30. Kawanishi K, Sunquist ME. Conservation status of tigers in a primary rainforest of Peninsular Malaysia. BiolConserv. 2004;120(3):329-344.
31. Seidensticker J. Saving wild tigers: a case study in biodiversity loss and challenges to be met for recovery beyond 2010. IntegrZool, 2010;5(4):285-299.
32. Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 2014;346(6215):1320-1331.
33. Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C et al. Comparative genomics reveals insights into avian genome evolution and adaptation. Science. 2014;346(6215):1311-1320.
34. Poinar HN, Schwarz C, Qi J, Shapiro B, MacPhee RD, Buigues B et al. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. Science. 2006;311(5759):392-394.
35. Riesenfeld CS, Schloss PD, Handelsman J. Metagenomics: genomic analysis of microbial communities. Annu Rev Genet. 2004;38:525-552.
36. Bodewes R, Ruiz-Gonzalez A, Schapendonk C, Van Den Brand J, Osterhaus A, Smits SL. Viral metagenomic analysis of feces of wild small carnivores. Virol J. 2014;11:89.
37. Yang F, Wang Y, Zheng W, He B, Jiang T, Li Y et al. [Metagenomic analysis of bat virome in several Chinese regions]. Sheng wu gong chengxuebao= Chinese J Biotechnol. 2013;29(5):586-600.
38. Jori F, Bastos AD. Role of wild suids in the epidemiology of African swine fever. EcoHealth. 2009;6:296–310.
39. Blomström AL, Ståhl K, Masembe C, Okoth E, Okurut AR, Atmnedi P et al. Viral metagenomic analysis of bushpigs (*Potamochoeruslarvatus*) in Uganda identifies novel variants of Porcine parvovirus 4 and Torque teno sus virus 1 and 2. Virol J. 2012;9:192.
40. Bodewes R, van der Giessen J, Haagmans BL, Osterhaus AD, Smits SL. Identification of multiple novel viruses, including a parvovirus and a hepevirus, in feces of red foxes. J Virol. 2013;87(13):7758-7764.
41. Reuter G, Nemes C, Boros Á, Kapusinszky B, Delwart E, Pankovics P. Astrovirus in wild boars (*Sus scrofa*) in Hungary. Arch Virol. 2012;157(6):1143-1147.
42. van den Brand JM, van Leeuwen M, Schapendonk CM, Simon JH, Haagmans BL, Osterhaus AD, Smits SL. Metagenomic analysis of the viral flora of pine marten and European badger feces. J Virol. 2012;86(4):2360-2365.
43. Phan TG, Kapusinszky B, Wang C, Rose RK, Lipton HL, Delwart EL. The fecal viral flora of wild rodents. PLoSPathog. 2011;7(9):e1002218.
44. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton G et al. The sequence of the human genome. Science. 2001;291(5507):1304-1351.
45. Wang J, Lin M, Crenshaw A, Hutchinson A, Hicks B, Yeager M et al. High-throughput single nucleotide polymorphism genotyping using nanofluidic Dynamic Arrays. BMC genomics. 2009;10(1):561.
46. Zook JM, Chapman B, Wang J, Mittelman D, Hofmann O, Hide W, Salit M. Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. Nat Biotechnol. 2014;32(3):246-251.
47. Peters EJ, McLeod HL. Ability of whole-genome SNP arrays to capture must have pharmacogenomic variants. Pharmacogenomics. 2008;9(11):1573-1577.
48. Liu L, Li Y, Li S, Hu N, He Y, Pong R et al. Comparison of next-generation sequencing systems. J Biomed Biotechnol. 2012;2012:251364.
49. Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM et al. Accurate multiplex polony sequencing of an evolved bacterial genome. Science. 2005;309(5741):1728-1732

50. Timp W, Mirsaidov UM, Wang D, Comer J, Aksimentiev A, Timp G. Nanopore sequencing: Electrical measurements of the code of life. IEEE Trans Nanotechnol. 2010;9(3):281-294.
51. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M et al. An integrated semiconductor device enabling non-optical genome sequencing. Nature. 2011;475(7356):348-352.
52. Sun X, Liu D, Zhang X, Li W, Liu H, Hong W et al. SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. PLoS One. 2013;8(3):e58700.
53. Roberts RJ, Carneiro MO, Schatz MC. The advantages of SMRT sequencing. Genome Biol. 2013;14(6):405.
54. Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I et al. (2008). Single-molecule DNA sequencing of a viral genome. Science. 2008;320(5872):106-109.
55. Thomas CS, Glassman MJ, Olsen BD. Solid-state nanostructured materials from self-assembly of a Globular Protein–Polymer Diblock Copolymer. Acs Nano. 2011;5(7):5697-5707.
56. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M et al. The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC bioinformatics. 2008;9(1):386.
57. Sun S, Chen J, Li W, Altintas I, Lin A, Peltier S et al. Community cyberinfrastructure for advanced microbial ecology research and analysis: the CAMERA resource. Nucleic Acids Res. 2010;39(suppl_1): D546-D551.
58. Gerlach W, Jünemann S, Tille F, Goesmann A, Stoye J. WebCARMA: a web application for the functional and taxonomic classification of unassembled metagenomic reads. BMC Bioinformatics. 2009;10(1):1.
59. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009;75(23):7537-7541.