Animal Science Papers and Reports vol. 35 (2017) no. 3, 205-224
Institute of Genetics and Animal Breeding, Jastrzębiec, Poland

Next generation sequencing in animal science - a review

Aleksandra Dunisławska, Jagoda Łachmańska, Anna Sławińska, Maria Siwek*

Department of Animal Biochemistry and Biotechnology,
University of Science and Technology, Mazowiecka 28, 85-084 Bydgoszcz, Poland

(Accepted March 30, 2017)

Next-generation sequencing (NGS), also known as high-throughput sequencing, is a novel method widely used in animal science and veterinary research. Introduction of NGS was a great breakthrough that has revolutionized the world of molecular biology techniques and animal genetics. An unquestionable advantage of NGS is connected with the almost unlimited insight into genetic information it provides. This review discusses the most important applications and achievements in animal genomics thanks to detailed sequence information. Here we present the history of sequencing and its further development. Applications of the NGS technique in whole genome sequencing, whole exome sequencing, targeted sequencing of DNA fragments and RNA sequencing in animal research are discussed.

KEY WORDS: animal genomes / sequencing / targeted DNA sequencing / RNAseq

Introduction

Research on the genetic basis of phenotypic variation and biological processes is the most rapidly growing area of the modern biology. Development of sequencing methods has been triggered by the need for complete genome analysis including comprehensive gene testing. Sequencing facilitates deciphering of large amounts of genetic information. However, the knowledge of the nucleotide sequence alone is not

*The review was supported by grants: UMO-2014/13/B/NZ9/02123 and UMO-2013/11/B/NZ9/00783 funded by the National Science Centre in Cracow (Poland).
**Corresponding author: siwek@utp.edu.pl

sufficient. It is also necessary to implement bioinformatics to process huge amounts of obtained molecular results. Sequencing is the first step in many experiments. A known sequence allows to locate the gene in the genome. Nucleic acid sequence also enables the observation of molecular mechanisms and their functioning. It facilitates determination of the risk of genetic diseases, therefore provides new opportunities in medicine. Insight into the sequence may provide answers to questions related to the evolution and enhance opportunities to detect new mutations. The aim of the review is to discuss the basis of the Next Generation Sequencing (NGS) and to present current achievements of this technology in animal research.

First generation sequencing

The onset of DNA sequencing dates back to the 1970's. A team of researchers headed by Fred Sanger developed a sequencing method based on polynucleotide chain termination [Sanger *et al.* 1977]. As a result, Sanger was named the father of sequencing and a creator of one of the most popular methods, referred to later as Sanger sequencing. This technique involves the use of partial digestion of radiolabeled fragments after two-dimensional fractionation [Sanger *et al.* 1965]. The greatest advantage of this method was connected with the potential for automation. Chemical degradation of the polynucleotide chain was another method developed at that time [Maxam and Gilbert 1977]. However, it was not widely used due to the requirement for radioactive reagents, complexity of the procedure and difficulty in reading the sequence. Decades later, a new method was developed, known as pyrosequencing [Ronaghi *et al.* 1996; Ronaghi *et al.* 1998]. The main advantage of this method is real-time sequencing, this means that the result is provided by synthesizing the new strand. Products resulting from sequencing are subjected to gel electrophoresis. This method is rapid, about 100 times faster as compared to the chain termination method, and fully automated [Ronaghi 2001]. Pyrosequencing provide the basis for the development of a new sequencing approach, referred to later as Next-Generation Sequencing.

First sequenced genomes

Availability and development of new technologies facilitated first attempts at whole genome sequencing. The first fully sequenced genome of a cellular organism was that of a bacterium, *Mycoplasma genitalium* [Fraser *et al.* 1995]. This genome is about 580,000 bp long. Due to the size of bacterial genomes the method of whole–genome random sequencing was applied. This method involves the sequencing and assembly of unselected DNA fragments from the whole chromosome. This approach eliminates initial mapping and is used for many microbial species, for which genomic maps are not available. This method was used to describe the genome of a bacteria *Haemophilus influenza* [Fleischmann *et al.* 1995]. In 1996 the first eukaryote – yeast *Saccharomyces cerevisiae* [Galibert *et al.* 1996] and in 1998, a nematode

Caenorhabditis elegans were sequenced [Sequencing Consortium C. elegans 1998]. Both eukaryotic organisms are important models in molecular genetics.

Another important step in DNA sequencing was announced in 2001. The work on the human genome begun in 1990 was led by the US Department of Energy and the US National Institutes of Health (NIH). The Human Genome Project received 3 billion dollars for 15 years of work. The full sequence of the human genome was given by Venter *et al.* [2001]. It was 90% of the human genome, the highest coverage possible at that time. In 2003 additional information was presented covering 99% of the human genome sequence [Marroquin 2007]. The success of this project speeded up sequencing of other genomes and brought the attention of scientists to the possibility of using genomics in animal breeding.

In the years 2003-2009 genomes of companion animals such as the cat [Pontius *et al.* 2007] and the dog [Lindblad-Toh *et al*. 2005] as well as farm animals including the chicken [Hillier *et al.* 2004], sheep [Archibald *et al.* 2010], cattle [Zimin *et al.* 2009], pig [Groenen *et al.* 2012] and the horse [Orlando *et al.* 2013] were sequenced. The analysis of the genetic material of livestock and domestic animals brings more than deciphering the genetic code. Vast numbers of DNA polymorphisms were used to develop several panels of SNP markers. These panels contributed to the accurate mapping of genetic damage [Berry *et al.* 2011], characterisation of quantitative trait loci [Siwek *et al.* 2015], genetic analysis of population structure [McKay *et al.* 2008; Edea *et al.* 2013] and evolutionary history of species [Alföldi and Lindblad-Toh 2013]. However, the most important aspect was connected with the implementation of molecular methods into animal breeding, aimed at improving breeding value. Genomic selection significantly supports traditional animal breeding, complementing traditional techniques based on phenotypic selection [Dekkers 2012].

Second and third generation sequencing

Thanks to the development of sequencing techniques and computer analysis the use of luminescent measurement of pyrophosphate synthesis has become common [Heather and Chain 2016]. Second generation sequencing is based on sequencing of amplified DNA molecules. Two examples of such methods include Solexa and Pyrosequencing [Ahmadian *et al.* 2000; Heather and Chain 2016]. Pyrosequencing is based on flash chemiluminescence emitted during the incorporation of a new nucleotide into the DNA sequence [Ahmadian *et al.* 2000]. Second generation sequencing incorporates Polymerase Chain Reaction (PCR), generation of a genomic library, immobilisation of sequenced DNA fragments on a solid substrate and fluorescent analysis of incorporated nucleotides as described in Pareek *et al.* [2011]. Third-generation sequencing provides the sequence of a single DNA molecule and may include such methods as tSMS (True Single Molecular Sequencing) and SMRT (Single Molecular Real-Time), in which the sequencing matrix does not have to be preceded by pre-amplification [Schadt *et al.* 2010]. This generation of sequencing does not require the generation of libraries, therefore any

problems resulting from amplification prior to sequencing are eliminated. Among the most important advantages of third generation sequencing when compared to second generation sequencing include higher throughput, a shorter analysis time, analysis of longer fragments within a shorter time, better accuracy facilitating the detection of rare variance, lower costs of whole genome sequencing, or the possibility of sequencing unique individual DNA molecules [Pareek *et al.* 2011]. A common characteristic for both generations of sequencing is the appearance of sequencing in real time and the use of sequencing by synthesis. However, in the case of the third generation approach, sequencing by synthesis is not necessary, or this type of sequencing is replaced by e.g. strand degradation. The similarity between second and third generation sequencing is in the simultaneous analysis of multiple copies. In second generation sequencing DNA copies are obtained by amplification, while in the third generation approach copies are obtained during the isolation of genetic material [Heather and Chain 2016]. Third generation sequencing provides direct RNA sequencing, which was not possible with second generation sequencing, where the RNA sample has to be transcribed into cDNA and amplified by PCR [Wang *et al.* 2009].

Next generation sequencing

NGS is currently the main tool in molecular genetic research. NGS has introduced new possibilities in the creation of high-throughput sequence data of large genomic regions. The entire human genome may be sequenced in one day. It provides an indispensable source of information for DNA and RNA analyses. The main principle of NGS resembles the Sanger sequencing method and is based on signals that are emitted by individual nucleotides during re-synthesis of matrix DNA. A key aspect in this process is the proper preparation of the library. The main advantage of NGS is its high throughput, huge amounts of produced data, high precision, extensive applicability and high quality reading [Reis-Filho 2009]. The next-generation sequencing method takes place in three stages: (1) isolation and creation of the DNA library, (2) amplification of the template, and (3) massive parallel sequencing (Mardis 2008b). There are several commercially available platforms for sequencing, e.g. Illumina, Roche454 or AB SOLiD. A feature common to all of them is the need to prepare a library of single stranded DNA. Sequencing steps vary depending on the platform.

The most important application of NGS is re-sequencing of already known genomes to determine genetic polymorphisms or mutations in selected genes [Zhang *et al*. 2015]. NGS may also be used in the analyses of large-scale DNA methylation as a tool applied in epigenomics. This method replaced the previously used pyrosequencing, methylation specific PCR and Sanger sequencing in the analysis of bisulfite converted DNA [Masser *et al.* 2015]. Application of NGS for DNA methylation sequencing solved problems with low sample throughput, short read length and low quantitative accuracy. NGS incorporated in gene mapping allows the study of DNA-protein interactions as well as the location of the protein coding sequence. The greatest

challenge and limitation of NGS is connected with computational capabilities of computers.

Application of NGS in animal genomics

Microbiological approach

The next generation sequencing technique provides microorganism sequencing without no need for prior culture [Garza and Dutilh 2015]. It thus facilitates identification of new antibiotics or metabolites produced by microorganisms that may be of potential use in human or veterinary medicine [Bryant *et al.* 2012; Challis 2014]. Thanks to the rapid insight into the genome, it is possible to locate genes responsible for the synthesis of new compounds which are inactive under laboratory conditions, e.g. *Streptomyces* which produce secondary metabolites of importance for the pharmaceutical industry [Ikeda *et al.* 2003; Gomez-Escribano *et al.* 2015]. NGS supports the discovery of new bioactive natural products that could potentially reveal new possibilities for its biosynthesis. It also promotes better understanding of biosynthesis of microorganisms isolated from different habitats which may be difficult to culture and manipulate in the laboratory. An example of such an application is *Streptomyces leeuwenhoek*, isolated from the Atacama desert in Chile, which is used in the production of new polyketide antibiotics [Gomez-Escribano *et al.* 2015]. Another important achievement was connected with deciphering the complete sequence of the bacteriophage with antimicrobial properties that inhabits chicken intestines and which could be used to fight bacterial diseases [Diaz-Sanchez *et al.* 2013].

Genomics of endangered species

Another application of NGS is to sequence species threatened with extinction, already extinct organisms or human ancestors. One of the projects is connected with genome sequencing of the woolly mammoth, which provided us with a better understanding of the evolution of elephants and mammoths [Rohland *et al.* 2010]. In 2010 a group of researchers [Li *et al.* 2010] announced that they had sequenced 2.25 GB, which represents 94% of the entire genome of the giant panda. The aim of that study was to find potential genes that determine the specific bamboo diet and allow for the exclusion of diet dependence affecting the gut microbiome [Li *et al.* 2010]. The same reasoning for genome sequencing for traits related with the threat of extinction was behind analysing the genetic material from koalas [Johnson *et al.* 2014]. Genomes of exotic animals became an object of interest for scientists. In 2014 Dastjerdi *et al.* [2014] published a report on the sequencing of the Indian elephant genome. There are three extant species of elephants: two African and one Indian. One of the goals of elephant genome analysis was to identify the sequence of elephant endotheliotropic herpesvirus (EEHV) that causes the deadly haemorrhagic disease in elephants. For this purpose genetic material from the two elephants which died of haemorrhagic disease was sequenced. Analysis confirmed the presence of EEHV

genetic material. Its genome size is 180,800 bp and contains 120 protein-coding genes [Wilkie *et al.* 2014]. The NGS technique made it possible to sequence the viral genome directly from post-mortem material. This direction of research holds promise for the development of studies on a vaccine against the virus, which is a huge threat to young animals and endangered species.

Cho *et al.* [2013] sequenced five endangered species of the *Panthera* genus: the African lion, African white lion, Asian tiger, white Bengal tiger and the snow leopard. A comparative analysis of the data showed a close genetic relationship of all these species. Additionally, sequencing revealed the gene and the SNP that determines white coat colour of the African white lion (TYR260G>A mutation) and genetic changes in *EGLN1* responsible for adaptation to high altitude (Met39>Lys39) [Cho *et al.* 2013]. A study of specific species adaptation using DNA sequencing was performed on the painted turtle. This analysis revealed a genetic mechanism which is responsible for the extraordinary ability of this species to survive without oxygen for as long as several months. Research of turtle DNA also focused on genetic traits associated with tooth loss, immune function, longevity and gender differentiation [Shaffer *et al.* 2013]. NGS also solved the secret of evolutionary changes caused by the animal transition from the terrestrial to aquatic environment. Whole genome sequencing was performed based on material collected from the minke whale, fin whale, finless porpoise and the bottlenose dolphin. It demonstrated the expansion of families of genes involved in the production of protein responsible for stress and anaerobic metabolism. The sequences revealed a number of changes related to the physiological and morphological adaptation to aquatic lifestyle, particularly resistance to stress caused by a lack of oxygen and a high level of salinity [Yim *et al.* 2014]. These studies illustrate that potential applicability of NGS seem vast.

Animal breeding

NGS was applied in the analysis of domestication in horses. Sequencing of several horse populations indicated the 10,193,421 SNPs and 1,361,948 insertion/deletion polymorphisms. Data analysis demonstrated a significant enrichment of private mutations in the coding regions of genes involved in primary metabolism, anatomical structures and morphogenesis [Metzger *et al.* 2014].

The source of genetic diversity in mammals in terms of the copy number variations (CNV) is the basis for biological functions that are associated with phenotypic traits. CNVs are an important form of genetic variation, which together with SNP has become an important source of information on the genetic variance and provides insight into inheritance of complex traits [Jiang *et al.* 2014]. Biological material from two cattle breeds: Holstein and Korean beef cattle Hanwoo has been sequenced using NGS. Deletion of 6.811 CNV was identified between the analysed individuals. The highest number of deleted copies was detected in 30 genes responsible for the functioning of the nervous system [Shin *et al.* 2014]. Detection of CNV by NGS has been used in the study of several swine breeds. Meishan pigs were analysed with a whole-genome

assembly comparison, whole-genome shotgun sequence detection and NGS. These complex analyses revealed the number of genes encoding proteins of variable copy number and confirmed two genes (*AADAT* and *ZNF622)* with high CNV values. These studies provide insight into the variability of the pig genome [Jiang *et al.* 2014]. The pig model has also been used to detect structural variants (SV) using NGS. SVs constitute rearrangements of the genome, which are often associated with genetic disorders. Chinese and European pig breeds were analysed. NGS detected 35 SV-related genes playing an important role in reproduction. SV analysis using NGS has been widely used in phylogenetic analyses of pig breeds. Identified SVs might be associated with a majority of genetic changes and might be used in cladistic analysis (a systematic method of analysis that detects and verifies phylogenetic relationships between taxa). This analysis points to the fact that SVs are significantly enriched or depleted with different functions of the genome [Zhao *et al.* 2016].

Next generation sequencing has also been used in poultry production in the search for genes that determine egg shell quality. From the economic point of view, broken egg shells cause great losses; additionally, they also pose a hazard to human health. Analysis of egg shell thickness and strength was carried out on samples taken from the oviducts of 49-wk-old Rhode Island hens. Apart from genome analysis, the transcriptome was examined for features and mutations responsible for shell calcification [Zhang *et al.* 2015]. A total number of 14234 active genes was identified, including 889 genes determining egg shell strength. These genes are related to the degree of mineralisation and are involved in calcium pathways. Subsequently the chicken genome was re-sequenced, which demonstrated 3 671 919 SNPs and 509 035 point mutations [Zhang *et al.* 2015].

Targeted Next Generation Sequencing of DNA

Some biological applications do not require whole-genome information. For this purpose, techniques have been developed that focus on specific sections of the genome in contrast to the whole-genome approach. Targeted sequencing was developed as a novel diagnostic tool to detect mutations [Nijman *et al.* 2014] and replace Sanger sequencing in diagnostics [Sikkema-Raddatz *et al.* 2013]. This method is based on the direct selection of genomic sub-regions or smaller sets of genes and their selective sequencing. Targeted sequencing is particularly useful for applications that take into account different parts of the genome [Altmüller *et al.* 2014]. Costs may be reduced thanks to the smaller amounts of information which need to be processed.

There are several techniques of targeted enrichment for NGS analyses (discussed below). When selecting an optimal methodology several parameters should be taken into consideration: the enrichment factor, mistakes in reading, specificity and uniformity of coverage of the entire region. Another important issue is connected with the length of the DNA fragment and the number of repetitions that it includes. Techniques used to prepare DNA fragments for sequencing differ depending on the

length of the target fragments [Mamanova *et al.* 2010]. For small DNA regions, ranging from 10 to 100 kb, the most effective methods that ensure deep and uniform coverage of selected regions, are based on PCR. The most commonly applied methods are: (a) long-range PCR, which is easy to automate and optimise [Jia *et al.* 2014]; (b) micro-droplet PCR, in which two types of microdroplets are used - one containing a sample of genetic material and the other, consisting of primers [Komori *et al.* 2011]; (c) multiplex PCR [Mertes *et al.* 2011; Zheng *et al.* 2014] or highly multiplexed PCR, which produces a very large number of short fragments in parallel using hundreds of primers. The amplified product must be purified and prepared in the form of an appropriate library for NGS analysis. For DNA fragments greater than 500 kb it is recommended to use hybridisation [Mertes *et al.* 2011]. A nucleic acid originating from the sample is hybridised to the pre-processed DNA fragments complementary to the target region. This technique is mostly used to identify genetic mutations in exons. The most recommended technique for sequences from 100 to 500 kb is based on molecular inversion probes (MIPs) [Niedzicka *et al.* 2016]. Probes used for enrichment by circularisation comprise a single-stranded DNA oligonucleotide which contains two complementary sequences of the target sections of the genomic fragment. At the end it contains two sequences which are complementary to the target genomic fragment in the reverse linear order. The specific hybridisation between these fragments generates a circular DNA structure. Subsequently, they are converted into closed circles by gap filling and ligation reactions [Hiatt *et al.* 2013, Stefan *et al.* 2016].

Whole Exome Sequencing (WES)

The most common application of targeted sequencing is Whole Exome Sequencin [Rabbani *et al.* 2014]. WES targets only coding regions, which represent less than 2% of the genetic code, but mutations in their areas are much more dangerous than in the remaining 98%, being non-coding sequences [St Laurent *et al.* 2014]. WES allows to focus on genes that may have the greatest impact on the phenotype. WES may also be used to determine the function of individual genes, identify new SNPs and their underlying phenotypic effects. To date more than 6,000 presumably single-gene disorders have been described in the human, although for most of them their genetic explanation is unknown [Rabbani *et al.* 2014]. WES may also be used to create clinical tools by identifying genetic variations responsible for Mendelian disorders and other genetic diseases in humans [Rabbani *et al.* 2014]. Commercially available kits for exome sequencing are only available for the human and the mouse. Attempts are being made to create reliable and efficient tools, which could be used to analyse exomes of other experimental animal models such as dogs [Broeckx *et al.* 2015]. In 2014 results of genetic analyses of bladder cancer, prostatic and histiocytic sarcomas in domestic dogs were published. Due to the high similarity of those cancers in humans and in dogs, efforts have been made towards biomedical applications of the dog model [Davis and Ostrander 2014]. Identification of DNA sequences of tumour cells compared to

normal cells provides insight into causes of malignancy. The differences between the sequences indicate genes involved in abnormal tissue growth and predisposition to spontaneous mutation of cells [Yan *et al.* 2011]. Increasingly, scientists decide to search for an analogy between mechanisms of cancers in easily accessible tissues of animal models. A unique set of molecular changes induced by each cancer facilitates an individualised approach to its treatment. The NGS technique can characterise the cancer genome, which combined with sequencing the whole patient exome might increase the chances for successful therapy. In these studies the mouse is used as an in vivo model to test the proposed treatment strategy [Garralda *et al.* 2014].

Amplicon sequencing

Coding regions are not the only sequences which are important for the functioning of the organism. For targeted sequencing of selected DNA regions, other targeted methods may be applied, e.g. amplicon sequencing. This sequencing method provides high coverage for specific genes of interest. Using targeted sequencing very complex samples may be analysed in one experiment, along with comparative sequencing of many individuals at the same time. Targeted sequencing has contributed greatly to SNP discovery. A higher level of targeted sequencing is made possible by sequencing amplicons - DNA fragments which are generated by PCR [Jia *et al.* 2014]. Amplicon analysis is particularly useful for the discovery of rare somatic mutations and metagenomic projects. They are based on genome analysis of all organisms inhabiting a particular ecological niche, e.g. when other analytical methods may not be used due to difficult conditions. Such analyses have been used for extensive studies of the genetic material of all microorganisms without their prior culturing in profiling the microbial population in aquatic ecosystems or soil from areas inhabited by people [Tan *et al.* 2015]. In 2015 molecular methods to assess drinking water quality were presented [Vierheilig *et al.* 2015]. Samples containing bacteria naturally occurring in the water, feces, soil and sediment were analysed. A total of 454 amplicons of bacterial 16S rRNA were sequenced. It was shown that markers identified by NGS methods are a valuable tool in monitoring water quality and effective pathogen detection. Another application of targeted NGS is connected with SNP discovery and genotyping. An example in this respect may be provided by the analysis of pathogen resistance to antiparasitic agents. It has been suspected that SNPs cause structural changes in the coding sequence of certain amino acids. The object of the study were the larvae of equine roundworms, the most virulent and the most common pathogens of horses and foals [Tydén *et al.* 2014]. Pooled samples from 100 000 eggs were tested. Due to the unilateral treatment it was observed that populations of nematodes started to develop resistance to many substances. This phenomenon has also been observed in other intestinal parasites found in cattle and sheep. It has been demonstrated that pathogen resistance is associated with mutations of a single nucleotide and may be disseminated in populations by natural gene flow [Bryant *et al.* 2012]. This fact is a big problem for veterinary treatment because resources for safe and effective medicines

are limited. For this reason, attempts have been made to detect mutational changes in the sequences responsible for nematode resistance. At the same time it was analysed which treatment least contributes to this resistance.

Estimation of animal performance based on their genotype is widely used in the selection of breeding animals. The underlying concept is to design molecular diagnostic tests for genes responsible for the phenotypic manifestation of quantitative traits. A powerful tool for the detection of many complex variations associated with a particular trait at the genomic level is a Genome-Wide Association Study (GWAS). This method was used to study the genetic impact on milk production in dairy cattle [Jiang . *et al.* 2014]. Economically this trait is the most important for this type of animals. For this reason to improve breeding efficiency it is essential to gain insight into genetic mechanisms that may contribute to milk production improvement. Although many QTLs for milk production have already been reported [Hu *et al.* 2013], only a few candidate genes have been identified. By means of resequencing some completely new genes associated with milk production traits have been discovered. Targeted resequencing yielded 66 SNPs dispersed in 53 genes [Jiang *et al.* 2014].

An approach was presented, which reveals evolutionary processes and prospects for population genomics based on marine fish species [Hemmer-Hansen *et al.* 2014]. Marine fish have many traits which make them attractive species to study the principle of evolution in natural populations. A large variety of fish provides a body of information at the molecular level concerning processes of adaptation to diverse conditions. Identification of regions in the genome that are involved in this adaptation provides insight into mechanisms of selection in natural populations [Radwan and Babik 2012; Seehausen *et al.* 2014]. The genome of fish is characterized by a high genetic variation, which has contributed to the discovery of large numbers of molecular markers. In 2012 a panel of 578 single nucleotide polymorphisms was identified in the Atlantic herring using transcriptome sequencing [Helyar *et al.* 2012].

Next Generation RNA Sequencing

Proteins are produced based on the information coded in the DNA sequence. However, deciphering the RNA sequence demonstrates the actual biological activity of the cell. There are many types of RNA molecules constituting the transcriptome, such as mRNA, tRNA, rRNA or non-coding RNA. Ribonucleic acid (RNA) takes highly different forms in terms of structure, size and above all, its functions. RNA is also considerably to be less stable than DNA. Due to this fact the study of the RNA sequence requires the implementation of complex and often complicated methods. Early attempts at gene expression were based on the Northern blot analysis [Alwine *et al.* 1977]. However, those experiments were very laborious and required isolation of huge amounts of RNA. The development of quantitative PCR facilitated more efficient analyses. The combination of Sanger sequencing and microarray techniques made it possible to characterise the gene expression level for thousands of transcripts

in various cell types or physiological conditions [Schena *et al.* 1995]. In subsequent years, a number of modifications and improvements to techniques based on microarrays have been made [Morozova *et al.* 2009]. A breakthrough occurred when the RNA sequencing technique was designed facilitating a quantitative analysis of gene expression [Delseny *et al.* 2010]. The third generation sequencing has contributed to direct RNA sequencing, while deep-sequencing has facilitated the analysis of the nucleotide sequence for small sRNA [Shendure and Ji 2008]. New sequencing technologies contribute to our understanding of the function of non-coding, short fragments of RNA. At present, the importance and function of non-coding RNA is not fully understood, but it is suspected that ncRNA molecules may play a role in carcinogenesis [Stahlhut Espinosa and Slack 2006].

Application of NGS of animal transcriptome

Next Generation RNA sequencing additionally provides a large body of information, such as the appearance of the RNA structure, forms of alternative splicing, or the presence of SNPs in the transcribed regions [Mardis 2008a]. RNAseq is useful not only in mapping of the transcriptome, but also provides insight into biological processes or mechanisms of adaptive evolution [Meisel *et al.* 2010] and immune response mechanisms [Qian *et al.* 2014; Wang *et al.* 2016]. RNA sequencing is performed on such platforms as Roche454 or Illumina. In 2007 Roche was the company which developed the first RNAseq technique based on pyrosequencing. Currently, the most commonly used RNAseq technique is the chemical synthesis used by Illumina. In 2015 this technique provided 90% of all RNA sequences worldwide.

Currently, sequencing of total RNA, including coding and non-coding forms, measures the quantitative expression of individual genes and identifies known or novel characteristics of the transcriptome. Among other things total RNAseq provides comprehensive transcriptome profiling, looking at their entire area of biomarkers and greater understanding of the phenotypic effects. Researchers benefit from the advantages brought by RNAseq studies of many domestic animal species or experimental animal models, e.g. the chicken. This species promotes understanding of animal development mechanisms and is used in functional genomics. In 2014 an analysis of transcripts isolated from embryonic chicken hearts compared with five other tissues of adult individuals was carried out. This study revealed more than 9,000 new isoforms, including examples of an alternative transcription start site of lysoforms [Thomas *et al.* 2014]. Research in the field of poultry molecular biology is also an example of improving the efficacy of protecting animals against pathogens. Similar analyses are conducted in other animal species. The knowledge of mRNA or protein structures associated with meat quality facilitates the construction of practical biomarkers to assess and predict individual traits [Ma *et al.* 2014]. Various animal tissues are subjected to transcriptome analysis. For example, in dairy cattle transcripts of somatic cells of milk and mammary tissue were analysed in relation to

lactation [Cánovas *et al.* 2014]. In 2010 results of RNA sequencing of IVF embryos were published, presenting differentially expressed genes in various stages of embryo development and alternatively spliced forms of gene sets [Huang *et al.* 2010]. Attempts are also made at the comparative analysis of changes in gene expression between populations resistant and non-resistant to pathogens, e.g. nematodes [Capuco *et al.* 2011]. Total RNAseq has also been used to detect changes in the intestines of diseased or infected animals [Li and Schroeder 2012; Nalpas *et al.* 2013; Scholey *et al.* 2013]. RNAseq is also used for analysis of reproduction in cattle and several other species by determining gene expression levels in such tissues as bovine endometrium [Mamo *et al.* 2012], ovarian cysts [Walsh *et al.* 2012] or the blastocyst [Driver *et al.* 2012]. Similar studies were performed on porcine gonads [Esteve-Codina *et al.* 2011] and the hypothalamus [Pérez-Montarelo *et al.* 2014]. Research has also been conducted on some fish species, e.g. the Zebrafish [Collins *et al.* 2012]. In the last three years transcriptomics of fish has evolved rapidly. The analyses focused on important species playing a significant role in environmental ecosystems. These studies provided a considerable body of valuable information, including the discovery of new SNPs or determining post-translational modifications. Results of these studies made it possible to explain many adaptive mechanisms or helped to determine water toxicity and the impact of pollution on organisms [Qian *et al.* 2014]. Another interesting application of RNAseq is connected with determining genes that are particularly active during long-term training of horses [McGivney *et al.* 2010]. Experiments were conducted in purebred horses. A total of 92 differentially expressed transcripts in the muscle tissue were detected before and after long-term exercise. The downregulation of the myostatin gene was observed after exercise. A panel of candidate genes was proposed for traits associated with processes influenced by physical activity in muscle cells, directly connected with the development and structure of muscles and metabolic processes. Horses, in particular the thoroughbred are the object of many studies. In 2010 the RNA sequencing method was used to analyse the transcriptome of several tissues, including cartilage, placenta, testes, cerebellum and embryonic tissue, facilitating a comprehensive functional and structural analysis of a number of genes in horses [Coleman *et al.* 2010]. Many scientists focused their attention on analysing and discovering new variants of transcripts associated with functional traits or good health of horses.

An excellent model to monitor human diseases or observe evolutionary mechanisms is provided by primates. Currently, biomedical studies in this field consist in collecting information concerning the RNA sequence to supplement databases. In 2013 ten tissues of macaques were analysed, yielding valuable data which verified the previous information on the transcripts in this species [Zhang *et al.* 2013]. Then, based on the complete and reliable data, it will be possible to analyse molecular mechanisms of many diseases. Although RNAseq is a relatively new technique, we have been observing its exceptionally dynamic growth and in the near future we can expect many exciting discoveries in animal studies.

Summary

With a combination of several previously developed strategies and implementation of new, more accurate protocols, a specialized branch of genetic testing was created, which allows us to analyse the entire genome. New approaches facilitate the analysis of smaller fragments of the nucleic acid sequence, which has many advantages, particularly saving time and costs. Next generation targeted sequencing is a precise technique, which is more often used by researchers. Currently, the apparatus for sequencing reactions is much more common and available even for small laboratories. We are seeing an extremely dynamic development of research in each sector. Researchers focus on solving mysteries of the animal world. It may be done by deciphering the most powerful code which nature has created. Therefore current research topics include the analysis of evolution, protection of endangered species, or applications meeting human needs and molecular tools for livestock animals. Scientists carry out work not only on the most common animal models, such as the fruit fly, rat, pig or the dog, but also less known species. Scientists are looking for analogies between other species and the human - in accordance with the concept assuming that human health is closely linked to the health of animals and that all the species share the same environment. In practice, successful applications of next-generation sequencing are found everywhere where isolation of genetic material in the form of DNA or RNA is possible. To date huge amounts of data have been generated, which contain information on the structure of nucleotide sequences. However, the greatest challenge is to understand their functional significance.

Acknowledgments. *The authors thank Mikołaj Kozłowski for technical assistance in literature review.*

REFERENCES

1. AHMADIAN A., GHARIZA DEH B., GUSTAFSSON A.C., STERKY F., NYRĚN P., UHLĚN M., LUNDEBERG J., 2000 – Single-Nucleotide Polymorphism Analysis by Pyrosequencing. ***Analytical Biochemistry*** 280, 103-110.
2. ALFÖLDI J., LINDBLAD-TOH K., 2013 – Comparative genomics as a tool to understand evolution and disease. ***Genome Research*** 23, 1063-1068.
3. ALTMÜLLER J., BUDDE B.S., NÜRNBERG P., 2014 – Enrichment of target sequences for next-generation sequencing applications in research and diagnostics. ***Biological Chemistry*** 395, 231-237.
4. ALWINE J.C., KEMP D.J., STARK, G.R., 1977 – Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. ***Proceedings of the National Academy of Sciences*** 74, 5350-5354.
5. ARCHIBALD A.L., COCKETT N.E., DALRYMPLE B.P., FARAUT T., KIJAS J.W., MADDOX J.F., MCEWAN J.C., HUTTON ODDY V., RAADSMA H.W., WADE C., WANG J., WANG W., WANG J., 2010 – The sheep genome reference sequence: a work in progress. ***Animal Genetics*** 41, 449-453.
6. BERRY D.P., BERMINGHAM M.L., GOOD M., MORE S.J., 2011 – Genetics of animal health and disease in cattle. ***Irish Veterinary Journal*** 64, 5.

7. BROECKX B.J., COOPMAN F., VERHOEVEN G., BOSMANS T., GIELEN I., DINGEMANSE W., SAUNDERS J.H., DEFORCE D., VAN NIEUWERBURGH F., 2015 – An heuristic filtering tool to identify phenotype-associated genetic variants applied to human intellectual disability and canine coat colors. ***BMC Bioinformatics*** 16, 391.
8. BRYANT J., CHEWAPREECHA C., BENTLEY S.D., 2012 – Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. ***Future Microbiology*** 7, 1283-1296.
9. CÁNOVAS A., RINCÓN G., BEVILACQUA C., ISLAS-TREJO A., BRENAUT P., HOVEY R.C., BOUTINAUD M., MORGENTHALER C., VAN KLOMPENBERG M.K., MARTIN P., MEDRANO J.F., 2014 – Comparison of five different RNA sources to examine the lactating bovine mammary gland transcriptome using RNA-Sequencing. ***Scientific Reports*** 4, 5297.
10. CAPUCO A.V., BINELLI M., TUCKER H.A., 2011 – Neither bovine somatotropin nor growth hormone-releasing factor alters expression of thyroid hormone receptors in liver and mammary tissues. ***Journal of Dairy Science*** 94, 4915-4921.
11. CHALLIS G.L., 2014 – Exploitation of the Streptomyces coelicolor A3 (2) genome sequence for discovery of new natural products and biosynthetic pathways. ***Journal of Industrial Microbiology & Biotechnology*** 41, 219-232.
12. CHO Y.S., HU L., HOU H., LEE H., XU J., KWON S., OH S., KIM H.M., JHO S., KIM S., SHIN Y.A., KIM B.C., KIM H., KIM C.U., LOU S.J., JOHNSON W.E., KOEPFLI K.P., SCHMIDT-KÜNTZEL A., TURNER J.A., MARKER L., HARPER C., MILLER S.M., JACOBS W., BERTOLA L.D., KIM T.H., LEE S., ZHOU Q., JUNG H.J., XU X., GADHVI P., XU P., XIONG Y., LUO Y., PAN S., SHIN Y.A., GOU C., CHU X., ZHANG J., LIU S., HE J., CHEN Y., YANG L., YANG Y., HE J., LIU S., WANG J., KIM C.H., KWAK H., KIM J.S., HWANG S., KO J., KIM C.B., KIM S., BAYARLKHAGVA D., PAEK W.K., KIM S.J., O'BRIEN S.J., WANG J., BHAK J., 2013 – The tiger genome and comparative analysis with lion and snow leopard genomes. ***Nature Communications*** 4.
13. COLEMAN S.J., ZENG Z., WANG K., LUO S., KHREBTUKOVA I., MIENALTOWSKI M.J., SCHROTH G.P., LIU J., MACLEOD J.N., 2010 – Structural annotation of equine protein□coding genes determined by mRNA sequencing. ***Animal Genetics*** 41, 121-130.
14. COLLINS J.E., WHITE S., SEARLE S.M., STEMPLE D.L., 2012 – Incorporating RNA-seq data into the zebrafish Ensembl genebuild. ***Genome Research*** *22*, 2067-2078.
15. SEQUENCING CONSORTIUM S. ELEGANS, 1998 – Genome sequence of the nematode C. elegans: a platform for investigating bilogy. ***Science*** 282, 2012-2018.
16. DASTJERDI A., ROBERT C., WATSON M., 2014 – Low coverage sequencing of two Asian elephant (Elephas maximus) genomes. ***GigaScience*** 3, 12.
17. DAVIS B.W., OSTRANDER E.A., 2014 – Domestic dogs and cancer research: A breed-based genomics approach. ***ILAR Journal*** 55, 59-68.
18. DEKKERS J.C.M., 2012 – Application of genomics tools to animal breeding. ***Current Genomics*** 13, 207-212.
19. DELSENY M., HAN B., HSING Y.I., 2010 – High throughput DNA sequencing: the new sequencing revolution. ***Plant Science*** 179, 407-422.
20. DIAZ-SANCHEZ S., HANNING I., PENDLETON S., D'SOUZE D., 2013 – Next-generation sequencing: the future of molecular genetics in poultry production and food safety. ***Poultry Science*** 92, 562-572.
21. DRIVER A. M., PEŃAGARICANO F., HUANG W., AHMAD K. R., HACKBART K. S., WILTBANK M. C., KHATIB H., 2012 – RNA-Seq analysis uncovers transcriptomic variations between morphologically similar in vivo-and in vitro-derived bovine blastocysts. ***BMC Genomics*** 13, 118.

22. EDEA Z., DADI H., KIM S.W., DESSIE T., LEE T., KIM H., KIM J.J., KIM K.S., 2013 – Genetic diversity, population structure and relationships in indigenous cattle populations of Ethiopia and Korean Hanwoo breeds using SNP markers. ***Frontiers in Genetics*** 4, 35.
23. ESTEVE-CODINA A., KOFLER R., PALMIERI N., BUSSOTTI G., NOTREDAME C., PÉREZ-ENCISO M., 2011 – Exploring the gonad transcriptome of two extreme male pigs with RNA-seq. ***BMC Genomics*** 12, 552.
24. FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F., KERLAVAGE A.R., BULT C.J., TOMB J.F., DOUGHERTY B.A., MERRICK J.M., et al., 1995 – Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. ***Science*** 269, 496-512.
25. FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., 1995 – The minimal gene complement of Mycoplasma genitalium. ***Science*** 270, 397-403.
26. GALIBERT F., ALEXANDRAKI D., BAUR A., BOLES E., CHALWATZIS N., CHUAT J.C., COSTER F., CZIEPLUCH C., DE HAAN M., DOMDEY H., DURAND P., ENTIAN K.D., GATIUS M., GOFFEAU A., GRIVELL L.A., HENNEMANN A., HERBERT C.J., HEUMANN K., HILGER F., HOLLENBERG C.P., HUANG M.E., JACQ C., JAUNIAUX J.C., KATSOULOU C., KARPFINGER-HARTL L., et al., 1996 – Complete nucleotide sequencing of Saccharomyces cerevisiae chromosome X. ***The EMBO Journal*** 15, 2031-2049.
27. GARRALDA E., PAZ K., LÓPEZ-CASAS P.P., JONES S., KATZ A., KANN L.M., LÓPEZ-RIOS F., SARNO F., AL-SHAHROUR F., VASQUEZ D., BRUCKHEIMER E., ANGIUOLI S.V., CALLES A., DIAZ L.A., VELCULESCU V.E., VALENCIA A., SIDRANSKY D., HIDALGO M., 2014 – Integrated next-generation sequencing and avatar mouse models for personalized cancer treatment. ***Clinical Cancer Research*** 20, 2476-2484.
28. GARZA D.R., DUTILH B.E., 2015 – From cultured to uncultured genome sequences: metagenomics and modeling microbial ecosystems. Cellular and Molecular ***Life Sciences*** 72, 4287-4308.
29. GOMEZ-ESCRIBANO J.P., CASTRO J.F., RAZMILIC V., CHANDRA G., ANDREWS B., ASENJO J.A., BIBB M.J., 2015 – The Streptomyces leeuwenhoekii genome: de novo sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2. ***BMC Genomics*** 16, 485.
30. GROENEN M.A.M., ARCHIBALD A.L., UENISHI H., TUGGLE C.K., TAKEUCHI Y., ROTHSCHILD M.F., et al., 2012 – Analyses of pig genomes provide insight into porcine demography and evolution. ***Nature*** 491, 393-398.
31. HEATHER J.M., CHAIN B., 2016 – The sequence of sequencers: The history of sequencing DNA. ***Genomics*** 107, 1-8.
32. HELYAR S.J., LIMBORG M.T., BEKKEVOLD D., BABBUCCI M., VAN HOUDT J., MAES G.E., BARGELLONI L., NIELSEN R.O., TAYLOR M.I., OGDEN R., CARIANI A., CARVALHO G.R., FISH POP TRACE CONSORTIUM, PANITZ F., 2012 – SNP discovery using next generation transcriptomic sequencing in Atlantic herring (Clupea harengus). ***PLoS One*** 7, e42089.
33. HEMMER-HANSEN J., THERKILDSEN N.O., PUJOLAR J.M., 2014 – Population genomics of marine fishes: next-generation prospects and challenges. ***The Biological Bulletin*** 227, 117-132.
34. HIATT J.B., PRITCHARD C.C., SALIPANTE S.J., O'ROAK B.J., SHENDURE J., 2013 – Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. ***Genome Research*** 23, 843-854.
35. HILLIER L.W., MILLER W., BIRNEY E., WARREN W., HARDISON R.C., PONTING C.P., INTERNATIONAL CHICKEN GENOME SEQUENCING CONSORTIUM, et al., 2004 – Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. ***Nature*** 432, 695-716.

36. HU Z.L., PARK C.A., WU X.L., REECY J.M., 2013 – Animal QTLdb: an improbed database tool for livestock animal QTL/association data dissemination in the post-genome era. ***Nucleid Acids Research*** 41, 871-879.
37. HUANG W., YANDELL B.S., KHATIB H., 2010 – Transcriptomic profiling of bovine IVF embryos revealed candidate genes and pathways involved in early embryonic development. ***BMC Genomics*** 11, 23.
38. IKEDA H., ISHIKAWA J., HANAMOTO A., SHINOSE M., KIKUCHI K., SHIBA T., SAKAKI Y., HATTORI M., OMURA S., 2003 – Complete genome sequence and comparatibe analysis of the industrial microorganism Streptomyces avermitilis. ***Nature Biotechnology*** 21, 526-531.
39. JIA H., GUO Y., ZHAO W., WANG K., 2014 – Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer. ***Scientific Reports*** 4, 5737.
40. JIANG J., WANG J., WANG H., ZHANG Y., KANG H., FENG X., WANG J., YIN Z., BAO W., ZHANG Q., LIU J., 2014 – Global copy number analyses by next generation sequencing provide insight into pig genome variation. ***BMC Genomics*** 15, 593.
41. JIANG L., LIU X., YANG J., WANG H., JIANG J., LIU L., HE S., DING X., LIU J., ZHANG Q., 2014 – Targeted resequencing of GWAS loci reveals novel genetic variants for milk production traits. ***BMC Genomics*** 15, 1105.
42. JOHNSON R.N., HOBBS M., ELDRIDGE M.D.B., KING A.G., COLDAN D.J., WILKINGS M.R., et al., 2014 – The Koala Genome Consortium. ***Technical Reports of the Australian Museum*** 24, 91-92.
43. KOMORI H.K., LAMERE S.A., TORKAMANI A., HART G.T., KOTSOPOULOS S., WARNER J., SAMUELS M.L., OLSEN J., HEAD S.R., ORDOUKHANIAN P., LEE P.L., LINK D.R., SALOMON D.R., 2011 – Application of microdroplet PCR for larg-scale targeted bisulfate sequencing. ***Genome Research*** 21, 1738-1745.
44. LI R., FAN W., TIAN G., ZHU H., HE L., CAI J., et al., 2010 – The sequence and de novo assembly of the giant panda genome. ***Nature*** 463, 311-317.
45. LI R.W., SCHROEDER S.G., 2012 – Cytoskeleton remodeling and alterations in smooth muscle contractility in the bovine jejunum during nematode infection. ***Functional & Integrative Genomics*** 12, 35-44.
46. LINDBLAD-TOH K., WADE C.M., MIKKELSEN T.S., KARLSSON E.K., JAFFE D.B., KAMAL M., et al., 2005 – Genome sequence, comparative analysis and haplotype structure of the domestic dog. ***Nature*** 438, 803-819.
47. MA J., YANG J, ZHOU L., REN J., LIU X., ZHANG H., YANG B., ZHANG Z., MA H., XIE X., XING Y., GUO Y., HUANG L., 2014 – A splice mutation in the PHKG1 gene causes high glycogen content and low meat quality in pig skeletal muscle. ***PLoS Genetics*** 10, e1004710.
48. MAMANOVA L., COFFEY A.J., SCOTT C.E., KOZAREWA I., TURNER E.H., KUMAR A., HOWARD E., SHENDURE J, TURNER D.J., 2010 – Target-enrichment strategies for next-generation sequencing. ***Nature Methods*** 7, 111-118.
49. MAMO S., MEHTA J.P., FORDE N., MCGETTIGAN P., LONERGAN P., 2012 – Conceptus-endometrium crosstalk during material recognition of pregnancy in cattle. ***Biology of Reproduction*** 87, 1-9.
50. MARDIS E.R., 2008a – The impact of next generation sequencing technology on genetics. ***Trends in Genetics*** 24, 133-141.
51. MARDIS E.R., 2008b – Next-Generation DNA Sequencing Methods. ***Annual Review of Genomics and Human Genetics*** 9, 387-402.
52. MARROQUIN J., 2007 – The Language of God: A Scientific Presents Evidence for Belif by Francis Collins. ***Proceedings (Baylor University Medical Centre)*** 20, 197-199.

53. MASSER D.R., STANFORD D.R., FREEMAN W.M., 2015 – Target DNA methylation analysis by next-generation sequencing. ***Journal of Visualized Experiments: JoVE*** 96.
54. MAXIM A.M., GILBERT W., 1977 – A new method for sequencing DNA. ***Proceedings of the National Academy of Sciences*** 74, 560-564.
55. MCGIVNEY B.A., MCGETTIGAN P.A., BROWNE J.A., EVANS A.C., FONSECA R.G., LOFTUS B.J., LOHAN A., MACHUGH D.E., MURPHY B.A., KATZ L.M., HILL E.W., 2010 – Characterization of the equine skeletal muscle transcriptome identifies novel functional response to exercise training. ***BMC Genomics*** 11, 398.
56. MCKAY S.D., SCHNABEL R.D., MURDOCH B.M., MATUKUMALLI L.K., AERTS J., COPPIETERS W., CREWS D., NETO E.D., GILL C.A., GAO C., MANNEN H., WANG Z., VAN TASSELL C.P., WILLIAMS J.L., TAYLOR J.F., MOORE S.S., 2008 – An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. ***BMC Genetics*** 9, 37.
57. MEISEL R.P., HILLDORFER B.B., KOCH J.L., LOCKTON S., SCHAEFFER S.W., 2010 – Adaptive evolution of genes duplicated from the Drosophila pseudoobscura neo-X chromosome. ***Molecular Biology and Evolution*** 27, 1963-1978.
58. MERTES F., ELSHARAWY A., SAUER S., VAN HELVOORT J.M., VAN DER ZAAG P.J., FRANK A., NILSSON M., LEHRACH H., BROOKES A.J., 2011 – Targeted enrichment of genomic DNA regions for next-generation sequencing. ***Briefings in Functional Genomics*** 10, 374-386.
59. METZGER J., TONDA R., BELTRAN S., AGUEDA L., GUT M., DISTL O., 2014 – Next generation sequencing gives an insight into the characteristics of highly selected breeds versus non-breed horses in the course of domestication. ***BMC Genomics*** 15, 562.
60. MOROZOVA O., HIRST M., MARRA M.A., 2009 – Applications of new sequencing technologies for transcriptome analysis. ***Annual Review of Genomics and Human Genetics*** 10, 135-151.
61. NALPAS N.C., PARK S.D.E., MAGEE D.A., TARAKTSOGLOU M., BROWNE J.A., CONLON K.M., RUE-ALBRECHT K., KILLICK K.E., HOKAMP K., LOHAN A.J., LOFTUS B.J., GORMLEY E., GORDON S.V., MACHUGH D.E., 2013 – Whole-transcriptome, high-throughput RNA sequence analysis of the bovin macrophage response to Mycobacterium bovis infection in vitro. ***BMC Genomics*** 14, 230.
62. NIEDZICKA M., FIJARCZYK A., DUDEK K., STUGLIK M., BABIK W., 2016 – Molecular Inversion Probes for targeted resequencing in non-model organisms. ***Scientific Reports*** 6: 24051.
63. NIJMAN I.J., VAN MONTFRANS J.M., HOOGSTRAAT M., BOES M.L., VAN DER CORPUT L., RENNER E.D., VAN ZON P., VAN LIESHOUT S., ELFERINK M.G., VAN DER BURG M., VERMONT C.L., VAN DER ZWAAG B., JANSON E., CUPPEN E., PLOOS VAN AMSTEL J., VAN GIJN M.E., 2014 – Targeted next-generation sequencing: a novel diagnostic tool for primary immunodeficiencies. ***Journal of Allergy and Clinical Immunology*** 133, 529-534.
64. ORLANDO L., GINOLHAC A., ZHANG G., FROESE D., ALBRECHTSEN A., STILLER M., SCHUBERT M., CAPPELLINI E., PETERSEN B., MOLTKE I., JOHNSON P.L.F., FUMAGALLI M., VILSTRUP J.T., RAGHAVAN M., KORNELIUSSEN T., MALASPINAS A.S., VOGT J., SZKLARCZYK D., KELSTRUP C.D., VINTHER J., DOLOCAN A., STENDERUP J., VELAZQUEZ A.M.V., CAHILL J., RASMUSSEN M., WANG X., MIN J., ZAZULA G.D., SEGUIN-ORLANDO A., MORTENSEN C., MAGNUSSEN K., THOMPSON J.F., WEINSTOCK J., GREGERSEN K., RŘED K.H., EISENMANN V., RUBIN C.J., MILLER D.C., ANTCZAK D.F., BERTELSEN M.F., BRUKAN S., AL-RASHIED K.A.S., RYDER O., ANDERSSON L., MUNDY J., KROGH A., GILBERT M.T.P., KJĆER K., SICHERITZ-PONTEN T., JENSEN L.J., OLSEN J.V., HOFREITER M., NIELSEN R., SHAPIRO B., WANG J., WILLERSLEV E., 2013 – Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. ***Nature*** 499, 74-78.
65. PAREEK C.S., SMOCZYNSKI R., TRETYN A., 2011 – Sequencing technologies and genome sequencing. ***Journal of Applied Genetics*** 52, 413-435.

66. PĚREZ-MONTARELO D., MADSEN O., ALVES E., RODRIGUEZ M.C., FOLCH J.M., NOGUERA J.L., GROENEN M.A.M., FERNĂNDEZ A.I., 2014 – Identification of genes regulating growth and fatness traits in pig through hypothalamic transcriptome analysis. ***Physiological Genomics*** 46, 195-206.
67. PONTIUS J.W., MULLIKIN J.C., SMITH D.R., AGENCOURT SEQUENCING TEAM, LINDBLAD-TOH K., GNERRE S., CLAMP M., CHANG J., STEPHENS R., NEELAM B., VOLFOVSKY N., SCHÄFFER A.A., RICHA A., NARFSTRÖM K., MURPHY W.J., GIGER U., ROCA A.L., ANTUNES A., MENOTTI-RAYMOND M., YUHKI N., PECON-SLATTERY J., JOHNSON W.E., BOURQUE G., TESLER F., NISC COMPARATIVE SEQUENCING PROGRAM, O'BRIEN S.J., 2007 – Initial sequence and comparative analysis of the cat genome. ***Genome Research*** 17, 1675-1689.
68. QIAN X., BA Y., ZHUANG Q., ZHONG G., 2014 – RNA-Seq technology and its application in fish transcriptomics. ***Omics: a Journal of Integrative Biology*** 18, 98-110.
69. RABBANI B., TEKIN M., MAHDIEH N., 2014 – The promise of whole-exome sequencing in medical genetics. ***Journal of Human Genetics*** 59, 5-15.
70. RADWAN J., BABIK W., 2012 – The genomics of adaptation. ***Proceedings of the Royal Society*** 279, 5024-5028.
71. REIS-FILHO J.S., 2009 – Next-generation sequencing. ***Breast Cancer Research*** 11, S12.
72. ROHLAND N., REICH D., MALLICK S., MEYER M., GREEN R.E., GEORGIADIS N.J., ROCA A.L., HOFREITER M., 2010 – Genomic DNA sequences from mastodon and wolly mammoth reveal deep speciation of forest and savanna elephants. ***PLoS Biology*** 8, e1000564.
73. RONAGHI M., 2001 – Pyrosequencing sheds light on DNA sequencing. ***Genome Research*** 11, 3-11.
74. RONAGHI M., KARAMOHAMED S., PETTERSSON B., UHLÉN M., NYRÉN P., 1996 – Real-time DNA sequencing using detection of pyrophosphate release. ***Analytical Biochemistry*** 242, 84-89.
75. RONAGHI M., UHLÉN M., NYRÉN P., 1998 – A sequencing method based on real-time pyrophosphate. ***Science*** 281, 363-365.
76. SANGER F., BROWNLEE G.G., BARRELL B.G., 1965 – A two-dimensional fractionation procedure for radioactive nucleotides. ***Journal of Molecular Biology*** 13, 373IN1-398IN4.
77. SANGER F., NICKLEN S., COULSON A., 1977 – DNA sequencing with chain-terminating inhibitors. ***Proceedings of the National Academy of Sciences*** 74: 5463-5467.
78. SCHADT E.E., TURNER S., KASARSKIS A., 2010 – A window into third-generation sequencing. ***Human Molecular Genetics*** 19, R227-R240.
79. SCHENA M., SHALON D., DAVIS R.W., BROWN P.P., 1995 – Quantitative monitoring of gene expression patterns with a complementary DNA microarrays. ***Science*** 270, 457-470.
80. SCHOLEY R.A., EVANS N.J., BLOWEY R.W., MASSEY J.P., MURRAY R.D., SMITH R.F., OLLIER W.R., CARTER S.D., 2013 – Identifying host pathogenic pathways in bovine digital dermatitis by RNA-Seq analysis. ***The Veterinary Journal*** 197, 699-706.
81. SEEHAUSEN O., BUTLIN R.K., KELLER I ., WAGNER C.E., BOUGHMAN J.W., HOGENLOHE P.A., PEICHEL C.L., SAETRE G.P., BANK C., BRÄNNSTRÖM A., BRELSFORD A., CLARKSON C.S., EROUKHMANOFF F., FEDER J.L., FISCHER M.C., FOOTE A.D., FRANCHINI P., JIGGINS C.D., JONES F.C., LINDHOLM A.K., LUCEK K., MAAN M.E., MARQUES D.A., MARTIN S.H., MATTHEWS B., MEITER J.I., MÖST M., MACHMAN M.W., NONAKA E., RENNISON D.J., SCHWARZER J., WATSON E.T., WESTRAM A.M., WIDMER A., 2014 – Genomics and the origin of species. ***Nature Reviews Genetics*** 15, 176-192.
82. SHAFFER H.B., MINX P., WARREN D.E., SHEDLOCK A.M., THOMSON R.C., VALENZUELA N., ABRAMYAN J., AMEMIYA C.T., BADENHORST D., BIGGAR K.K.,BORCHERT G.M., BOTKA C.W., BOWDEN R.M., BRAUN E.L., BRONIKOWSKI A.M., BRUNEAU B.G.,

BUCK L.T., CAPEL B., CASTOE T.A., CZERWINSKI M., DELEHAUNTY K.D., EDWARDS S.V., FRONICK C.C., FUJITA M.K., FULTON L., GRAVES T.A., GREEN R.E., HEARTY W., HARIHARAN R., HERNANDEZ O., HILLER L.W., HOLLOWAY A.K., JANES D., JANZEN F.J., KANDOTH C., KONG L., DE KONING A.P.J., LI Y., LITERMAN R., MCGAUGH S.E., MORK L., O'LAUGHIN M., PAITZ R.T., POLLOCK D.D., PONTING C.P., RADHAKRISHNAN S., RANEY B.J., RICHMAN J.M., ST JOHN J., SCHWARTZ T., SETHURAMAN A., SPINKS P.Q., STOREY K.B., THANE N., VINAR T., ZIMMERMAN L.M., WATTEN W.C., MARDIS E.R., WILSON R.K., 2013 – The western painted turtle genome, a model for the evolution of extreme physiological adaptations in a slowly evolving lineage. ***Genome Biology*** 14: R28.

83. SHENDURE J., JI H., 2008 – Next-generation DNA sequencing. ***Nature Biotechnology*** 26, 1135-1145.
84. SHIN D.H., LEE H.J., CHO S., KIM H.J., HWANG J.Y., LEE C.K., JEONG J.Y., YOON D., KIM H., 2014 – Deleted copy number variation of Hanwoo and Holstein using next generation sequencing at the population level. ***BMC Genomics*** 15:240.
85. SIKKEMA-RADDATZ B., JOHANSSON L.F., DE BOER E.N., ALMOMANI R., BOVEN L.G., VAN DER BERG M.P., VAN DER SPAENDONCK-ZWARTS K.Y., VAN TINTELEN J.P., SIJMONS R.H., JONGBLOED J.D., SINKE R.J., 2013 – Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. ***Human Mutation*** 34, 1035-1042.
86. SIWEK M., SLAWINSKA A., RYDZANICZ M., WESOLY J., FRASZCZAK M., SUCHOCKI T., SKIBA J., SKIBA K., SZYDA J., 2015 – Identification of candidate genes and mutations in QTL regions for immune response in chicken. ***Animal Genetics*** 46, 247-254.
87. STAHLHUT ESPINOSA C.E., SLACK F.J., 2006 – The role of microRNA in cancer. ***Yale Journal of Biology and Medicine*** 79, 131-140.
88. ST LAURENT G., VYATKIN Y., KAPRANOV P., 2014 – Dark matter RNA illuminates the puzzle of genome-wide association studies. ***BMC Medicine*** 12, 97.
89. STEFAN C.P., KOEHLER J.W., MINOGUE T.D., 2016 – Targeted next-generation sequencing for the detection of ciprofloxacin resistance markers using molecular inversion probes. ***Scientific Reports*** 6, 25904.
90. TAN B.F., NG C.M., NSHIMYIMANA J.P., LOH L.L., GIN K.Y., THOMPSON J.R., 2015 – Next-generation sequencing (NGS) for assessment of microbial water quality: current progress, challenges, and future opportunities. ***Frontiers in Microbiology*** 6, 1027.
91. THOMAS S., UNDERWOOD J.G., TSENG E., HOLLOWAY A.K., 2014 – Long-read sequencing of chicken transcripts and identification of new transcript isoforms. ***PLoS One*** 9, e94650.
92. TYDÉN E., DAHLBERG J., KARLSBERG O., HÖGLUND J., 2014 – Deep amplicon sequencing of preselected isolates of Parascaris equorum in β-tubulin codons associated with benzimidazole resistance in other nematodes. ***Parasit Vectors*** 7, 410.
93. VENTER J.C., ADAMS M.D., MYERS E.W., LI P.W., MURAL R.J., *et al.* 2001 – The sequence of the human genome. ***Science*** 291, 1304-1351.
94. VIERHEILIG J., SAVIO D., LEY R.E., MACH R.L., FARNLEITNER A.H., REISCHER G.H., 2015 – Potential applications of next generation DNA sequencing of 16S rRNA gene amplicons in microbial water quality monitoring. ***Water Science and Technology*** 72, 1962-1972.
95. WALSH S.W., MEHTA J.P., MCGETTIGAN P.A., BROWNE J.A., FORDE N., ALIBRAHIM R.M., MULLIGAN F.J., LOFTUS B., CROWE M.A., MATTHEWS D., DISKIN M., MIHM M., EVANS A.C.O., 2012 – Effect of the metabolic environment at key stages of follicle development in cattle: focus on steroid biosynthesis. ***Physiological Genomics*** 44, 504-517.
96. WANG K., DEL CASTILLO C., CORRE E., ESPINOSA E.P., ALLAM B., 2016 – Clam focal and systemic immune responses to QPX infection revealed by RNA-seq technology. ***BMC Genomics*** 17, 146.

97. WANG Z., GERSTEIN M., SNYDER M., 2009 – RNA-seq: a revolutionary tool for transcriptomics. ***Nature Reviews Genetics*** 10, 57-63.
98. WILKIE G.S., DAVISON A.J., KERR K., STIDWORTHLY M.F., REDROBE S., STEINVACH F., DASTIJERDI A., DENK D., 2014 – First fatality associated with elephant endotheliotropic herpesvirus 5 in an Asian Elephant: pathological findings and complete viral genome sequence. ***Scientific Reports*** 4, 6299.
99. Yan X.J., XU J., GU Z.H., PAN C.M., LU G., SHEN Y., SHI J.Y., ZHU Y.M., TANG L., ZHANG X.W., LIANG W.X., MI J.Q., SONG H.D., LI K.Q., CHEN Z., CHEN S.J., 2011 – Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. ***Nature Genetics*** 43, 309-315.
100. YIM H.S., CHO Y.S., GUANG X., KANG S.G., JEONG J.Y., CHA S.S., OH H.M., LEE J.H., YANG E.C., KWON K.K., KIM Y.J., KIM T.W., KIM W., JEON J.H., KIM S.J., CHOI D.H., JHO S., KIM H.M., KO J., KIM H., SHIN Y.A., JUNG H.J., ZHENG Y., WANG Z., CHEN Y., CHEN M., JIANG A., LI E., ZHANG S., HOU H., KIM T.H., YU L., LIU S., AHN K., COOPER J., PARK S.G., HONG C.P., JIN W., KIM H.S., PARK C., LEE K., CHUN S., MORIN P.A., O'BRIEN S.J., LEE H., KIMURA J., MOON D.Y., MANICA A., EDWARDS J., KIM B.C., KIM S., WANG J., BHAK J., LEE H.S., LEE J.H., 2014 – Minke whale genome and aquatic adapration in cetaceans. ***Nature genetics*** 46, 88-92.
101. ZHANG Q., ZHU F., LIU L., ZHENG C.W., HOU Z.C., NING Z.H., 2015 – Integrating transcriptome and genome re-sequencing data to identify key genes and mutations affecting chicken eggshell qualities. ***PLoS one*** 10, e0125890.
102. ZHANG S.J., LIU C.J., SHI M., KONG L., CHEN J.Y., ZHOU W.Z., ZHU X., YU P., WANG J., YANG X., HOU N., YE Z., ZHANG R., XIAO R., ZHANG X., LI C.Y., 2013 – RhesusBase: a knowledgebase for the monkey research community. ***Nucleic Acids Research*** 41, D892-905.
103. ZHAO P., LI J., KANG H., WANG H., FAN Z., YIN Z., WANG J., ZHANG Q., WANG Z., LIU J.F., 2016 – Structural Variant Detection by Large-scale Sequencing Reveals New Evolutionary Evidence on Breed Divergence between Chinese and European Pigs. ***Scientific Reports*** 6:18501.
104. ZHENG Z., LIEBERS M., ZHELYAZKOVA B., CAO Y., PANDITI D., LYNCH K.D., CHEN J., ROBINSON H.E., SHIM H.S., CHMIELECKI J., PAO W., ENGELMAN J.A., IAFRATE A.J., LE L.P., 2014 – Anchored multiplex PCR for targeted next-generation sequencing. ***Nature Medicine*** 20, 1479-1484.
105. ZIMIN A.V., DELCHER A.L., FLOREA L., KELLEY D.R., SCHATZ M.C., PUIU D., HANRAHAN F., PERTEA G., VAN TASSELL C.P., SONSTEGARD T.S., MARÇAIS G., ROBERTS M., SUBRAMANIAN P., YORKE J.A., SALZBERG S.L., 2009 – A whole-genome assembly of the domestic cow Bos taurus. ***Genome Biology*** 10, R42.