

Anna Woźniak, M.Sc.
Michał Boroń, M.Sc.
Renata Zbieć-Piekarska, MD, PhD
Magdalena Spólnicka, MD, PhD
Central Forensic Laboratory of the Police

Application of high-throughput DNA sequencing technology in forensic genetics

Summary

The turn of the 20th and 21st centuries marks the beginning of high-throughput DNA sequencing methods, which, owing to increasing efficiency and gradual cost reduction, have led to the revolutionization of biomedical research. This article discusses the most popular next generation sequencing technologies and their practical application in forensic genetic analysis.

Key words: next generation sequencing, NGS, MPS, sequencing by synthesis, application of MPS in forensics

Introduction

In the 1980s, the first DNA sequencing methods were developed. The team led by A.M. Maxam and W. Gilbert proposed a chemical method (Maxam, Gilbert, 1977), while F. Sanger's and A.R. Coulson's team developed a chain termination method (Sanger, Nicklen, Coulson, 1977). Due to its greater simplicity, lower toxicity and high potential for further improvement, the technique developed by Sanger's team became widely used and for a long time remained the so-called gold standard in DNA sequencing. Improvements to the original protocol included the replacement of isotopic labelling of dideoxynucleotides with fluorescent dyes and automation of the sequencing product analysis process (Franca, Carrilho, Kist, 2002; Heather, Chain, 2016). This method, nowadays considered belonging to the first generation of sequencing methods, has led to breakthrough research in genetics, including human genetics, in particular to the determination of the first complete sequences of mitochondrial and nuclear genomes (Anderson et al., 1981; Van Dijk et al., 2014). Sanger sequencing technology is one of the most reliable methods of DNA sequence analysis, but its significant limitation is low throughput. The growing demand for large scale DNA sequence analysis has led to the development of several technologies commonly referred to as Next Generation Sequencing (NGS) or Massively Parallel Sequencing (MPS). Next generation sequencing may involve analysis of a group of amplified

template DNA fragments, which is defined as Second Generation Sequencing (SGS), or analysis of individual DNA molecules (without the need for their earlier amplification), called Third Generation Sequencing (TGS) (Kotowska, Zakrzewska-Czerwińska, 2010; Piątkowski et al., 2013).

Areas of application in DNA analysis

Taking into account the scale of sequencing and the target area, three basic sequencing technologies can be distinguished: whole genome, exome and specific genomic fragments.

1. Genome sequencing

Whole Genome Sequencing (WGS) technology is the most comprehensive method of identifying hereditary diseases or characterizing non-inherited mutations causing cancer in humans. Aside from humans, also genomes of animal, plant, fungal or microbial species can be sequenced, both those already known, owing to the availability of reference genomes, as well as previously unexplored ones. In *de novo* sequencing, the genome is constructed from scratch without using a reference sequence. Special bioinformatic tools are used for this purpose.

2. Exome sequencing

In the case of exome sequencing, the area covered by the analysis are exons, i.e. gene-forming genome

regions. They occupy only about 1.5% of all genomic DNA, but encompass most of the variants responsible for diseases. Sequencing costs are relatively low and the amount of data that can be obtained in this way is enormous (www.genome.gov).

3. Targeted resequencing

In targeted resequencing, only specific gene sets or genomic regions are sequenced. This approach saves time and money by analysing data for specific DNA regions only, such as exons, selected genes, fragments of non-coding DNA or mitochondrial DNA. Targeted sequencing is also used to analyse regions linked with phenotypic traits. This method makes it possible to obtain greater coverage of the sequences analysed and to identify rare variants or different types of mutations at reduced costs of analysis.

Basic methods of next generation sequencing

1. Second generation sequencing

In these technologies, thousands or millions of amplified DNA fragments are sequenced and the incorporation of nucleotides into newly formed strands occurs cyclically during parallel reactions. SGS methods are relatively fast because sequencing and detection occur simultaneously. The instruments are sensitive to quantitative reading of the signal from a large number of identical molecules attached at specific locations. Manufacturers provide servers and software dedicated for reading and analysis of large amounts of complex measurement data, and most sequencers are nowadays compact in size. Commercially available technologies and platforms for second generation sequencing differ in terms of the technical solutions applied and the method of detection. This translates into differences in reading length, process efficiency, speed of reaction, experiment cost, and the type and frequency of errors induced during reading. Second generation methods have limitations resulting from the use of PCR and require the use of novel alignment and assembly algorithms. For a reliable analysis it is also necessary to obtain an adequate coverage, i.e. the minimum number of readings necessary for a proper analysis of the fragments (amplicons) to be examined (Yang, Xie, Yan, 2014; Płoski, 2016; Gupta, Gupta, 2014). As part of the improvement of SGS techniques, methods and instruments appeared on the market which are now being withdrawn, e.g. pyrosequencing (Roche) or the method based on oligonucleotide ligation – SOLID. New companies offering other solutions are also entering the market. Since 2016, high-throughput sequencers have been offered by BGI from China. Currently, the most commonly used MPS technology consists of sequencing by synthesis, with detection based on the measurement of fluorescence or the amount of hydrogen ions released during the reaction.

a. Synthesis by sequencing with fluorescence detection
In this type of sequencing, the preparation of libraries is a multi-stage process. The preparation of nucleic acid for MPS sequencing, called MPS library design, has replaced the previously used labor-intensive and time-consuming DNA cloning into bacterial vectors. Initially, DNA is amplified using a set of specific primers. During this process, DNA fragments are amplified and elongated by synthetic oligonucleotides called sequence tags or adapters, which perform identification and control functions at the sequencing stage. Then, magnetic beads are used to purify and normalize libraries, i.e. select fragments according to their size and equalize the concentration of individual libraries in relation to each other. The libraries obtained are combined together, denatured and diluted accordingly (Fig. 1a). The prepared libraries are pooled together and placed in a sequencer. The following reactions lead to the hybridization of single-stranded DNA fragments to complementary oligonucleotides attached to the wells in the surface of a glass plate (flow cell), and then to DNA amplification. In the bridge PCR, a polymerase elongates nascent strand after nucleotide addition. After the denaturation stage, the template strand is washed out. The cycles of annealing, elongation and denaturation are continuously repeated until clusters containing thousands of copies of the template fragment are formed. The last PCR cycle ends with flushing out one strand of DNA and blocking further elongation by adding a blocking dideoxynucleotide (ddNTP) to the 3' end (Fig. 1b). After preparing clusters with copies of target DNA molecules, the machine starts sequencing, which consists of incorporating fluorescently labelled nucleotides. Each tagging color is assigned to a specific base and acts as a reversible terminator. When a tagged nucleotide is incorporated, fluorescence is detected using a CCD camera. Because of the multiple DNA copies attached within a given area, the signal emitted by the cluster is strong enough to be captured by the camera. Flushing out the fluorescent group unlocks the possibility of incorporating the subsequent nucleotide. The cycles are repeated multiple (several hundred) times, whereby each cycle results in the incorporation of one of the four nucleotides present in the solution (Fig. 1c). This type of sequencing is used in Illumina sequencers (Heather, Chain, 2016; Van Dijk et al., 2014; Kotowska, Zakrzewska-Czerwińska, 2010; Piątkowski et al., 2013; Płoski, 2016; Gupta, Gupta, 2014).

b. Sequencing by synthesis with semiconductor-based detection and pH measurement

The construction of libraries is similar to the previous technology. The library amplification stage is designed differently, namely, it takes place outside the sequencer in the form of emulsion PCR. Within an emulsion of oil and water, droplets are forming containing a bead, DNA polymerase and other reagents necessary for the reaction. DNA amplification is carried out on beads

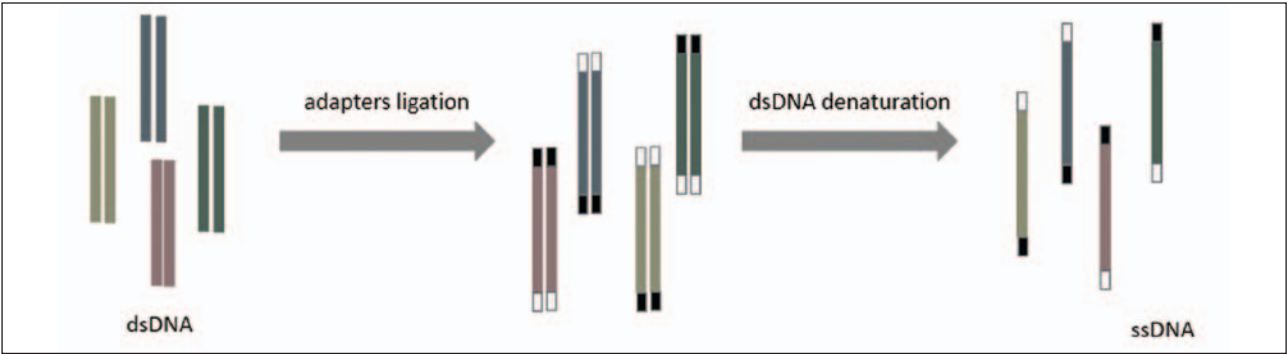


Fig. 1a. NGS library preparation in Illumina® technology.

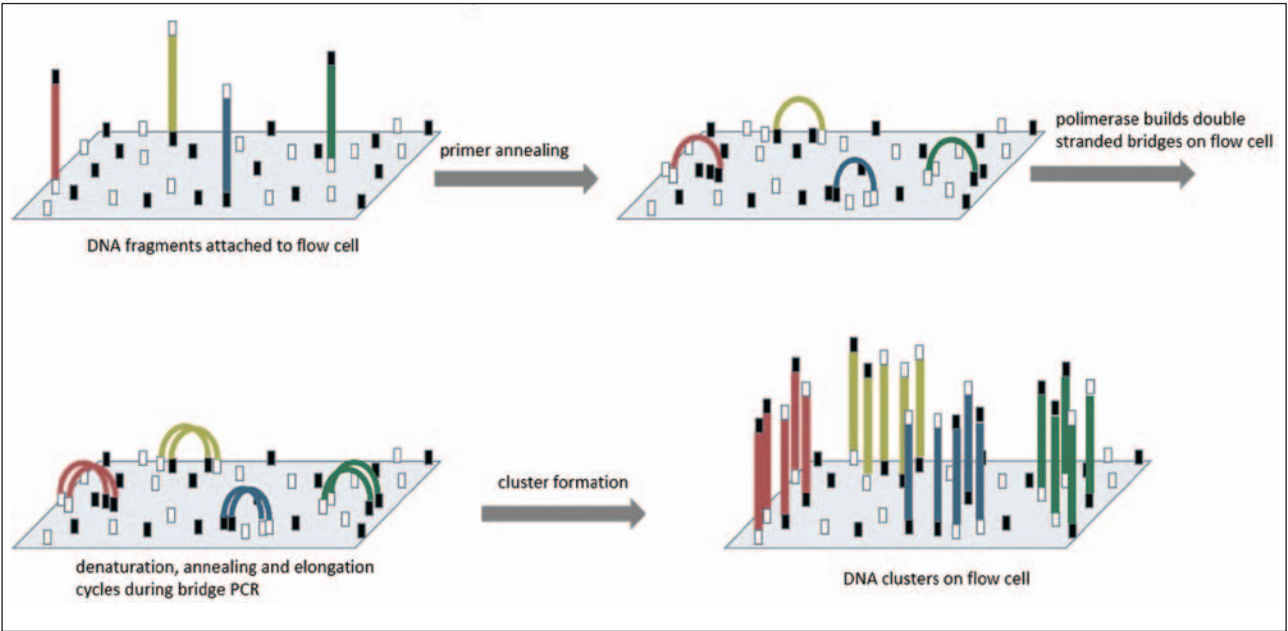


Fig. 1b. Bridge PCR and clusters formation on flow cell.

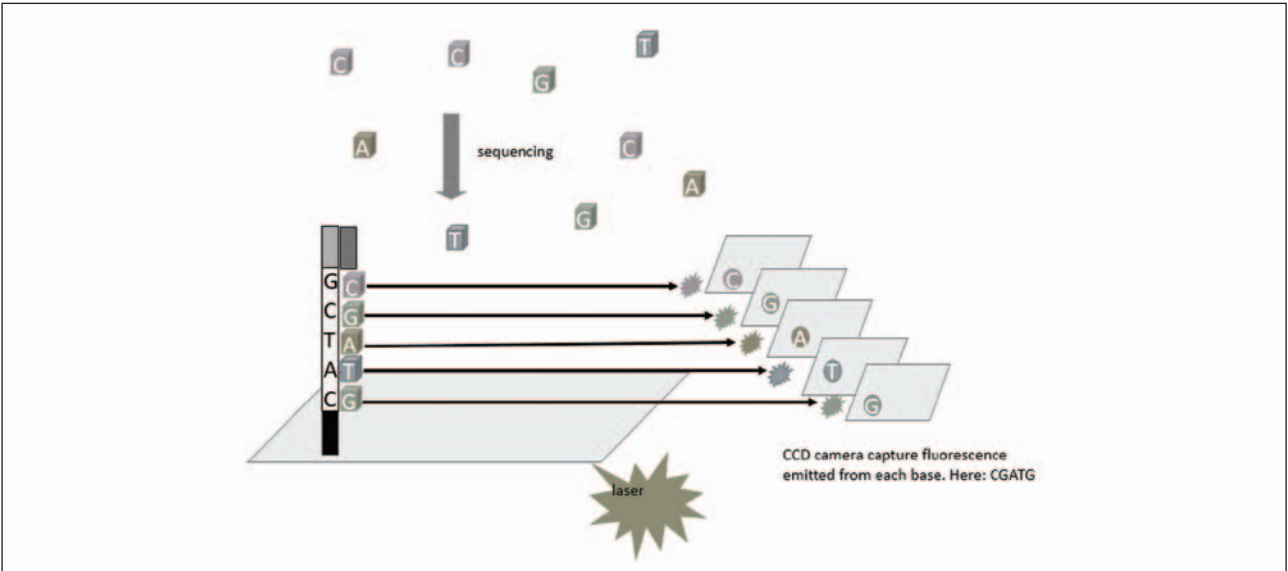


Fig. 1c. Sequencing by synthesis in Illumina® technology. Four fluorescence-labeled nucleotides are present. Fluorescence is measured with CCD camera after each cycle of nucleotide incorporation.

according to the principle: one DNA fragment – one bead (Fig. 2a). Next, uncoated, empty beads are largely rinsed away (so-called enrichment), while DNA-containing beads are transferred to a special plate called a chip. The complexes bead-multiplied DNA fragments match the size of the wells in the chip, underneath which the sensors are located measuring changes in pH. Once the chip is manually placed in the sequencer, the sequencing process takes place. Nucleotides are added individually, and when a nucleotide complementary to the template strand is incorporated, the hydrogen ions released are measured (Fig. 2b). The sequencer records the signal, whose strength depends on the number of nucleotides incorporated in a given cycle. In wells where the subsequent nucleotide is not complementary to that currently present in the solution, there is no incorporation or change in pH. The challenge of this method lies in correctly determining the number of bases in long homopolymeric stretches, while its advantages are the elimination of part of the recording equipment (laser, camera), simplification of the entire process, shortening the reading time and reducing the costs of sequencing. This type of sequencing is used in Ion Torrent's instruments such as PGM, Ion Proton and the latest S5, which cooperates with a compatible robot that prepares libraries and loads samples onto the chip, called Ion Chef (Van Dijk et al., 2014; Piątkowski et al., 2013; Płoski, 2016).

2. Third generation sequencing

Third generation sequencing involves reading a sequence of only one DNA molecule without amplifying molecules by PCR. Sequencing takes place in real

time, hence it is much faster than sequencing by synthesis, and the sequences read out are many times longer. The disadvantage of this technology is, for the time being, high reading error rate.

a. SMRT sequencing in ZMW optical fibre wells

Preparation of DNA for sequencing takes place in several stages. The first stage is the fragmentation of DNA into 10 kbp segments, followed by the repair of damage caused by fragmentation, the repair of DNA ends and its purification. In the next stage, two hairpin adapters are added to the blunt ends (it allows to read both strands multiple times) and the SMRTbell™ Templates sequencing primers are attached to the resulting complexes. Finally, a polymerase molecule is added to each complex. Sequencing takes place on a Single Molecule Real Time Sequencing (SMRT) plate with Zero-Mode Waveguides (ZMW) wells, where the complexes enter by diffusion or carried by the beads. The size of the well allows only one complex to fit into it. ZMW is a hole made in a 100 nm thick metal membrane placed on glass. The small diameter of the hole does not allow the light emitted by the laser underneath the hole to pass through its entire length. It is dispersed at 1/3 of the depth, in the place where the polymerase is affixed. Each of the four nucleotides is labelled with a fluorophore, which after excitation with laser light generates a flash of light. All types of nucleotides are present in the solution in which the plate is immersed. They flow into the wells on the principle of diffusion and flow out if there is no complementarity. The laser does not excite the fluorescent tags until they reach the bottom. The incorporation of nucleotide by polymerase increases the time of its exposure to

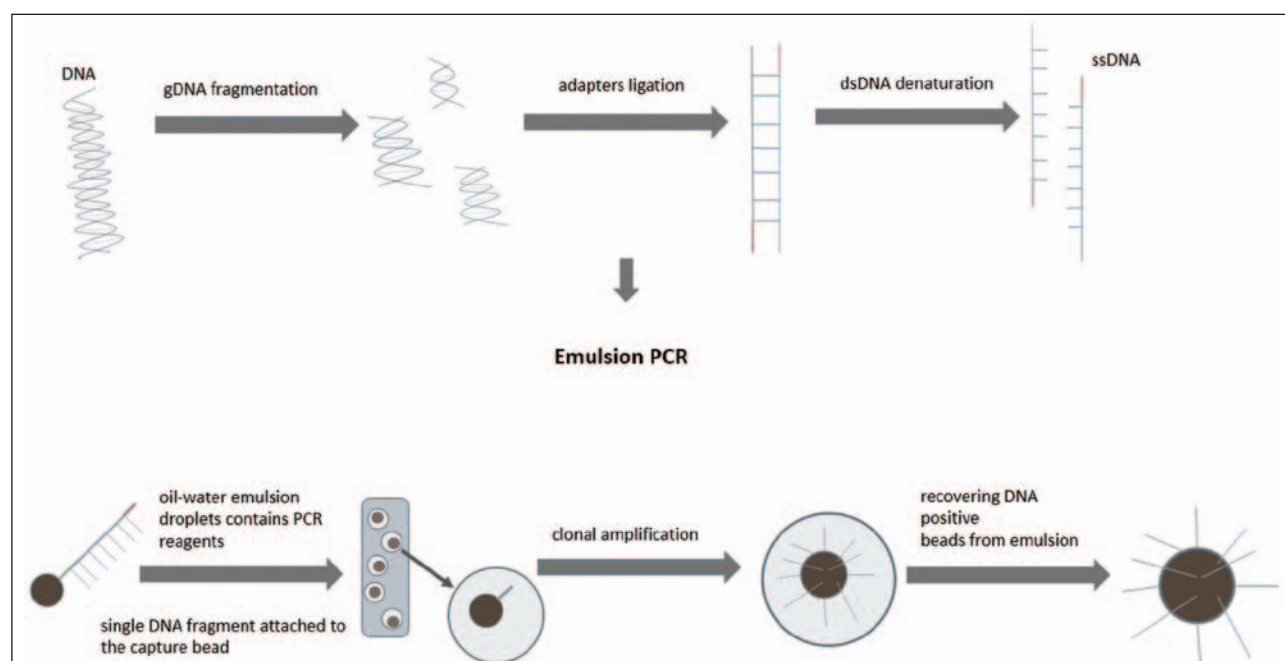


Fig. 2 a. NGS library preparation in Ion Torrent™ technology and clonal bead amplification (emulsion PCR).

laser light, resulting in a stronger emission of light in an appropriate colour for a given base (Fig. 3). The fluorescent flash is registered by the camera in real time. The tag is released after incorporation, which unlocks the possibility of another base being incorporated by the polymerase. This type of sequencing was introduced by Pacific Biosciences. It allows to read tens of thousands of base pairs within a few hours, as well as to detect epigenetic modifications in DNA (Heather, Chain, 2016; Van Dijk et al., 2014; Kotowska, Zakrzewska-Czerwińska, 2010; Piątkowski et al., 2013; Yang, Xie, Yan, 2014; Płoski, 2016; Gupta, Gupta, 2014; Rhoads, Au, 2015).

b. Sequencing in nanopores
Preparation of DNA samples consists of DNA fragmentation, end repair and adenylation. After the cleansing stage, the adapter with the structure of a hairpin is ligated to the poly(A) tail in double-stranded DNA. Next, the so-called molecular motor that unwinds the double helix and ensures optimal time of passing through the nanopore thanks to the immobilization of one of the strands, and the tether molecule affixing the free end of the DNA strand near the hole in the membrane are attached to DNA molecule. This technology is based on the natural phenomenon of transporting molecules through membranes, e.g.

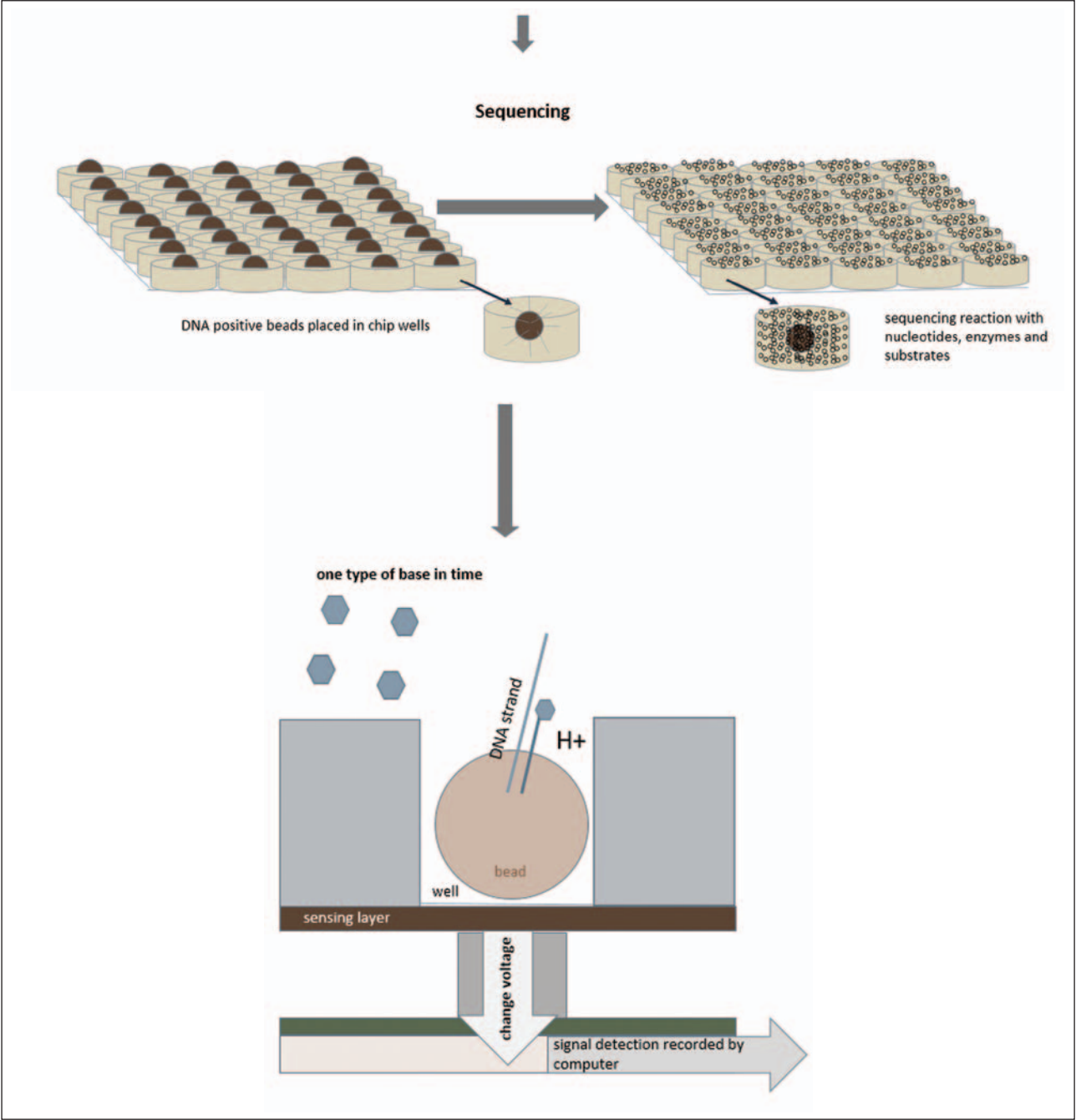


Fig. 2 b. Sequencing by synthesis in Ion Torrent™ technology with detection based on pH change.

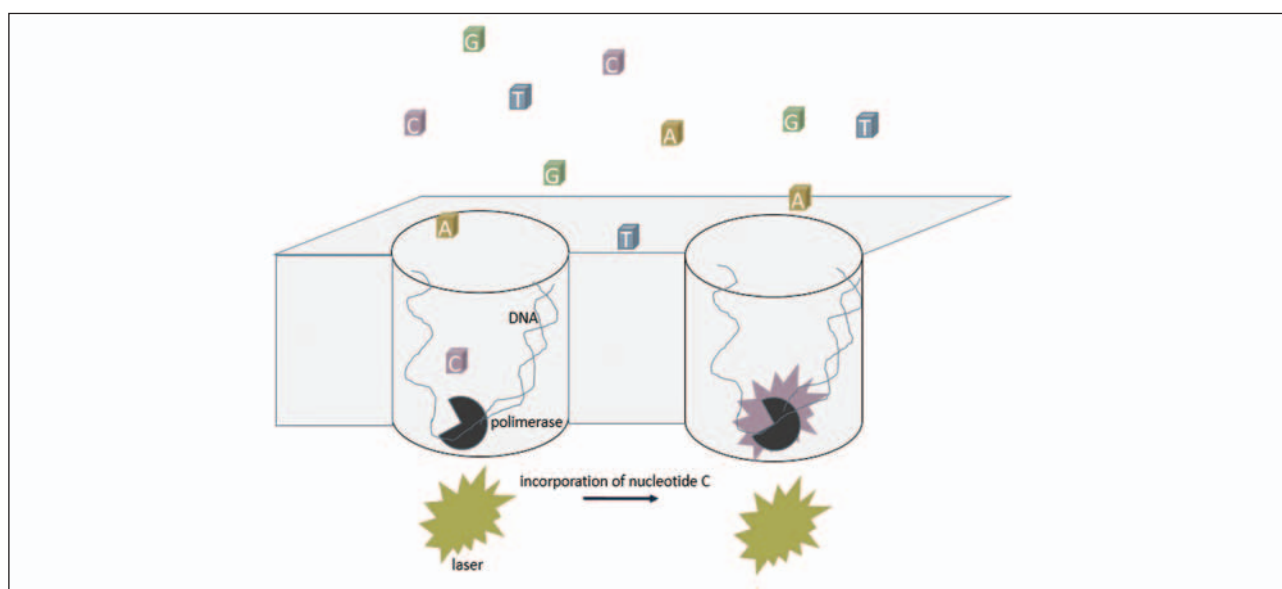


Fig. 3. Single molecule – SMRT sequencing in PacBio technology. Each type of base emitted different color light after laser excitation.

cell membranes with ion channels. Membranes can be of biological origin (protein, e.g. phi29, alpha-hemolysine), synthetic or mixed. Synthetic nanopores are holes of 1–100 nm diameter, formed by chemical etching, or bombarding thin films of materials (e.g. polymers, silicones, compounds of silicon, aluminium, molybdenum or graphene) with ions or electron beam from a transmission electron microscope. Nanopores have different properties: thickness, particle transfer rate, adhesiveness to DNA, etc. The membrane is selected to be as thin as possible, which makes it possible to read one base at a time. The electrically impermeable membrane is immersed in the solution, and the electrical voltage allows the ions to pass through the pore. A concentration gradient directs the particles towards the nanopore. The closer to the pore, the higher the concentration and the electrical conductivity. The charged particles move towards oppositely charged electrode and the DNA strand, due to its charge, moves in the solution in the direction of the positively charged electrode (Fig. 4). When the DNA molecule enters the nanopore, the ion flow is partially blocked and the electric current value decreases. The pace at which the molecule passes through the pore can be modified. It is slowed down by a decrease in temperature or current intensity and an increase in salt concentration. The flow of particles through smaller diameter pores is faster. The kinetics of DNA transfer depends on the length and conformation of the strands. Each base has a different conductivity, hence the sequence reading is based on its measurements (Feng et al., 2015).

One or more DNA molecules can be sequenced within one experiment. Readings can be up to several hundred thousand base pairs. This type of sequencing

is used in Oxford Nanopores instruments. The device in which the sequencing process takes place is a matchbox-sized chip with sensors measuring potential, electrodes and one or more nanopore plates called MinION and PromethION. This method can also be used to identify epigenetic changes in DNA (Heather, Chain, 2016; Kotowska, Zakrzewska-Czerwińska, 2010; Gupta, Gupta, 2014; Laver et al., 2015; Regalado, 2014).

Application of high-throughput DNA sequencing methods in forensic genetics

MPS technology introduces a range of new possibilities for DNA testing. Some of them are directly applicable in forensics, because they are an extension and combination of the methods of individual identification used so far. This includes the profiling of STR (Short Tandem Repeat) markers, both autosomal and located on sex chromosomes, Single Nucleotide Polymorphism (SNP) markers and mitochondrial DNA (mtDNA). Particularly promising is the use of high-throughput sequencing at the criminal investigation stage for the so-called forensic DNA phenotyping (FDP). Methods of predictive DNA analysis require the analysis of hundreds and even thousands of polymorphisms scattered throughout the human genome and therefore the development of these methods is closely related to the implementation of genomic research results in forensic genetics. DNA phenotyping refers primarily to the prediction of physical appearance, but often also includes the analysis of biogeographical origin and human age. Such information can be of intelligence value. Increasing the information value of the biological material by extracting additional data, even from difficult and degraded samples, is one of the key advantages

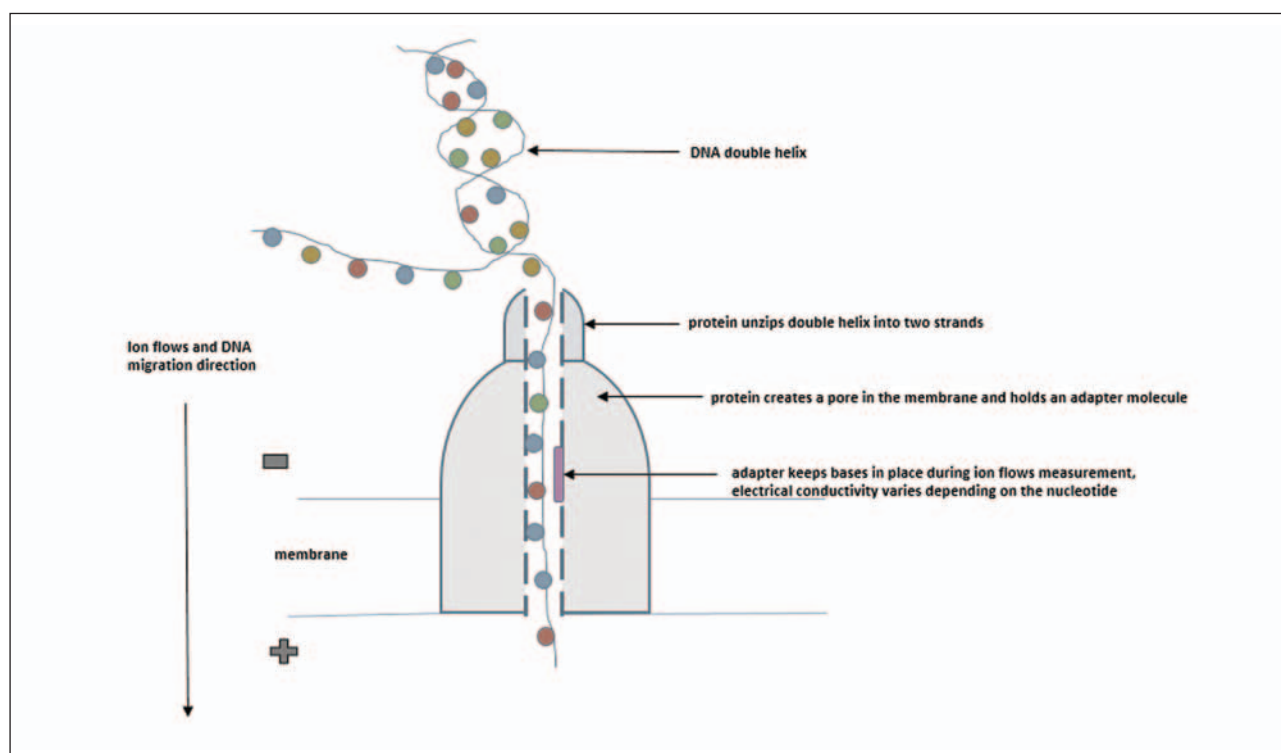


Fig. 4. Single molecule sequencing in Oxford Nanopores technology.

of modern sequencing technologies and in the future NGS may be routinely used in police operations.

1. Use of NGS

a. Individual identification: STR and SNP markers

Currently, the analysis of variability of non-coding DNA fragments, mainly microsatellite sequences called STR markers is used for genetic identification. DNA profiling may concern STR loci located on autosomal and sex chromosomes: X and Y. Autosomal STR markers are used for genetic identification in forensic laboratories using multiplex PCR and capillary electrophoresis technologies. Profiles obtained in the course of analyses are entered into national DNA databases and searched on the basis of national and international legal regulations. The basic STR set is the 13 loci system from CODIS (Combined DNA Index System) developed by the FBI, together with the amelogenin gene marker allowing for gender determination. The Y-STR analysis is particularly useful for establishing male relatives or detecting male components in sex crimes. Highly conservative mutations of the Y chromosome, inherited in subsequent generations make it possible to assign an individual to a specific haplogroup. The worldwide Y-STR Haplotype Reference Database (Y-HRD) contains information on identified haplotypes and their prevalence in different populations. Profiling with the use of NGS technology already allows for simultaneous examination of several dozen STR markers and several hundred SNP markers in one reaction for several dozen samples. By obtaining data on the number of repetitions

and sequences of the DNA markers tested, it is possible to increase the discriminatory power of testing. This is due to the presence of isoalleles, i.e. alleles of identical length, but with different DNA sequences. The discriminatory power is also increased by the analysis of DNA polymorphism of the regions flanking STR markers. The presence of isoalleles in the mixture may facilitate the process of its deconvolution, as well as enable the assessment of the origin of mutations (from the mother or father) during identification or paternity studies. SNP-type identification marker sets are particularly useful in situations where DNA is degraded. The 50–100 SNP set has a discriminatory power equal to that obtained for 10–16 loci STR set (Loveliness et al., 2017; Pakstis et al., 2009; Gill et al., 2004).

b. Mitochondrial DNA analysis

NGS technology is also applicable to mtDNA analysis, especially in the case of small amounts of material or degraded DNA. By sequencing the entire mitochondrial genome instead of just hypervariable polymorphic HVI/HVII regions, which are usually tested by using Sanger's technology, the discriminatory power increases and the problem of hypervariable regions shared by populations from the same ancestor is eliminated. Additionally, next generation sequencing increases heteroplasmy detection more than twice compared to Sanger sequencing. On the basis of mtDNA analysis, it is possible to establish a relationship in the female line, which is particularly useful when comparing samples originating from victims of mass disasters or unidentified

individuals with living relatives on the mother's side. The presence of numerous mtDNA haplogroups has been revealed, many of which show continental distribution and are a source of biogeographical data. The population database of the mtDNA control region EMPOP contains data on hypervariable regions, the coding region and the entire mitochondrial genome. This database allows to verify the correctness of the results obtained and to assign the sequence tested to a specific haplogroup (Holland, Makova, McElhoe, 2018; Just, Irwin, Parson, 2015; Shih et al., 2018).

c. Prediction of appearance and biogeographical origin

One of the most promising and important directions of development nowadays is the prediction of physical appearance (DNA phenotyping) and biogeographical origin (ancestry) of a sample donor. The tools used for this purpose are mostly based on the analysis of SNP markers. Special panels of the best markers of these traits have been developed and published to aid the prediction of both appearance and origin. In the case of phenotype prediction, the testing consists in genotyping markers linked to the genes responsible for human appearance and building mathematical models allowing for the analysis of the data obtained. Until now, the best known are the markers of pigmentation traits. The colour of iris, hair or skin tone is the result of interaction of several genes with high heritability. These include, but are not limited to: *HERC2*, *OCA2*, *SLC24A4*, *SLC45A2*, *SLC24A5*, *TYR*, *IRF4*, *MC1R*, *TYRP1*, *TPCN2*, *ASIP*, *KITLG*, *EXOC2*. These genes are not equivalent and some are responsible for a higher percentage of variation of a given trait than others, e.g. *HERC2*, *OCA2* for eye pigmentation and *MC1R* for hair pigmentation. Several predictive systems have been developed and validated for forensic purposes. They allow, after determining the genotype of specific markers, to predict the pigmentation of eyes, hair or even skin tones, e.g. IrisPlex (eye colour prediction: blue, intermediate, brown), HirisPlex (eye colour prediction: blue, intermediate, brown; hair colour: blond, red, brown, black; hair shade: light, dark) or Snipper (eye colour: brown, blue, greenish-blue; hair colour: blond, red, brown, black; hair shade: light, dark; skin tone: light, medium, dark). Work on genetic identification of such traits as male baldness (e.g. *AR*/*EDA2R* genes on X chromosome), hair graying, hair morphology (e.g. *EDAR*, *FGFR2*, *TCHH* genes), facial shape (e.g. *PAX3*, *PRDM16*, *TP63* genes) or height (Kayser, 2015) is underway. The assessment of human biogeographic origin is based mainly on the analysis of AIMs (Ancestry Informative Markers) markers of the SNP and InDel type. The analysis of origin on the basis of markers obtained from Y chromosome or mtDNA allows to reconstruct the migration pathways and population distribution, and, in relation to an individual, to combine his/her physical appearance with origin,

depending upon the percentage of admixture of another population.

Commercial forensic kits produced by market-leading companies such as ThermoFisherScientific or Illumina allow for genotype identification and/or prediction of hair, or eye pigmentation, or origin using NGS technology. For example, Illumina's FGx kit provides information on 27 autosomal STR markers, 24 Y chromosome STR markers, 7 X chromosome STR markers, 94 SNP-type identification markers, 22 SNP markers used to determine eye and hair color, and 56 SNP markers used to determine origin in a matter of days (<https://www.illumina.com/>).

d. Prediction of human age

Thanks to MPS technology, it is now possible to determine the age of a person on the basis of biological traces deposited with an accuracy of a few years. This is all the more important because age influences the appearance of a person, and thus the result of predictive analysis of phenotypic traits is associated with such changes as hair loss or graying. The tests use epigenetic markers – regions of the genome that change their methylation profile with age. Based on the analysis of these changes and mathematical modelling, human chronological age is predicted. Age markers are usually located on CpG islands within gene promoter regions and tend to hypomethylate or hypermethylate over time. Several hundred potentially useful epigenetic markers have been published. Some of these markers are used in the estimation of biological age, while others determine chronological age. The most frequently used are sets composed of several markers, which allow to estimate the age of a person with the accuracy of about ± 3 to 5 years on average from blood, saliva or sperm samples (Zbieć-Piekarska et al., 2015).

e. Identification of human tissues and body fluids: mRNA and miRNA markers

Body fluids and tissues are distinguished from each other by sequencing mRNA biomarkers, which allows to obtain the so-called mRNA profile. This approach is based on differentiated expression of particular genes in cells of different organs (e.g. lungs, brain, kidneys, skin) or in substances of interest from the point of view of forensic research (e.g. blood, saliva, semen, menstrual blood, vaginal secretion), a number of mRNA markers allowing their identification were discovered. What is important, simultaneous extraction of RNA and DNA from samples allows to perform both mRNA identification and classic STR profiling. Samples can be multiplexed, which speeds up the time of analysis, and the procedure is also suitable for degraded evidence samples.

MicroRNA (miRNA) biomarkers are also considered as potentially useful for the differentiation of substances such as urine, blood or saliva. MiRNA molecules are short (20–25 nucleotides) non-coding RNA structures

that are more resistant to degradation, which in the case of evidence originating from different environments and stored under different conditions is particularly important. A number of tissue specific miRNAs have been identified, but the methodology needs further development to ensure reproducible results (Hanson et al., 2018; Silva et al., 2014).

f. Microbiome

The sequencing of RNA and DNA of microorganisms is a source of information on the biogeographical origin of the evidence tested and can be used in an investigation.

Based on the knowledge of human microbiome, it is possible to identify a person, identify a part of the body from which the evidence originates and estimate time of death. The sequencing of genetic material also allows for the identification of the type of bacteria or viruses in cases involving the use of biological weapons (Lilje et al., 2013).

Summary

In the field of genetic testing and molecular analysis, new technologies are being replaced relatively quickly by even newer ones. On the one hand, high-throughput sequencers with the ability to analyze multiple genomes simultaneously are being developed, while on the other, researchers are looking for individual changes in DNA sequences that can be linked to specific characteristics of human appearance. Recently, significant portions of the sequencing and library preparation protocols have been automated, which allows to perform analyses with minimal human involvement. The precision of reading single DNA molecules is increasing, and the availability of small portable devices substantiates the vision of mobile laboratories performing the most complex analyses on the spot, allowing to determine the appearance and age of the suspect, and thus adding relevant information to the investigation.

Source of figures: authors

Bibliography

- Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J., Staden, R., Young, I.G. (1981). Sequence and organization of the human mitochondrial genome. *Nature*, 290(5806).
- Feng, Y., Zhang, Y., Ying, C., Wang, D., Du, C. (2015). Nanopore-based Fourth-generation DNA Sequencing Technology. *Genomics, Proteomics & Bioinformatics*, 13(1).
- Franca, L.T.C., Carrilho, E., Kist, T.B.L. (2002). A review of DNA sequencing techniques. *Quarterly Reviews of Biophysics*, 35(2).
- Gill, P., Werrett, D.J., Budowle, B., Guerrieri, R. (2004). An assessment of whether SNPs will replace STRs in national DNA databases – joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDM). *Science and Justice: Journal of the Forensic Science Society*, 44(1).
- Gupta, A.K., Gupta, U.D. (2014). Next generation sequencing and its applications. W: A.S. Verma (ed.), *Animal Biotechnology: Models in Discovery and Translation*. Amsterdam–Boston: Elsevier.
- Hanson, E., Ingold, S., Haas, C., Ballantyne, J. (2018). Messenger RNA biomarker signatures for forensic body fluid identification revealed by targeted RNA sequencing. *Forensic Science International: Genetics*, 34.
- Heather, M., Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*, 107(1).
- Holland, M.M., Makova, K.D., McElhoe, J.A. (2018). Deep-coverage MPS analysis of heteroplasmic variants within the mtGenome allows for frequent differentiation of maternal relatives. *Genes* (Basel), 9(3).
- Just, R.S., Irwin, J.A., Parson, W. (2015). Mitochondrial DNA heteroplasmy in the emerging field of massively parallel sequencing. *Forensic Science International: Genetics*, 18.
- Kayser, M. (2015). Forensic DNA Phenotyping: Predicting human appearance from crime scene material for investigative purposes. *Forensic Science International: Genetics*, 18.
- Kotowska M., Zakrzewska-Czerwińska J. (2010). Kurs szybkiego czytania DNA – nowoczesne techniki sekwencjonowania (A course in rapid DNA reading – modern sequencing techniques). *Biotechnologia*, 4(91).
- Laver, T., Harrison, J., O'Neil, P.A., Moore, K., Farbos, A., Paszkiewicz, K., Studholme, D.J. (2015). Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomolecular Detection and Quantification*, 3.
- Lilje, L., Lillsaar, T., Rätsep, R., Simm, J., Aaspõllu, A. (2013). Soil sample metagenome NGS data management for forensic investigation. *Forensic Science International: Genetics*, 4(1).
- Loveliness, D., Dennis, S., Salvador, J., Calacal, G., De Ungria, M. (2017). Comparison of two massively parallel sequencing platforms using 83 single nucleotide polymorphisms for human identification. *Scientific Reports*, 7, Article number: 398 (2017); doi:10.1038/s41598-017-00510-3.
- Maxam, A.M., Gilbert, W. (1977). A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74.
- Pakstis, A.J., Speed, W.C., Fang, R., Hylan, F.C.L., Furtado, M.R., Kidd, J.R., Kidd, K.K. (2009). SNPs

- for a universal individual identification panel. *Human Genetics*, 127(3), doi: 10.1007/s00439-009-0771-1.
17. Piątkowski, J., Skalniak, A., Bodzioch, M., Pach, D., Hubalewska-Dydejczyk, A. (2013). Wprowadzenie do sekwencjonowania ludzkiego genomu w diagnostyce (Introduction to human genome sequencing in diagnostics). *Przegląd Lekarski*, 70(7).
 18. Płoski, R. (2016). Next generation sequencing – general information about the technology, possibilities, and limitations. W: U. Demkow, R. Płoski (ed.), *Clinical Applications for Next-Generation Sequencing*. London: Academic Press.
 19. Regalado, A. (2014). Radical new DNA sequencer finally gets into researchers hands. *Biomedicine*, 17 września.
 20. Rhoads, A., Au, K. (2015). PacBio sequencing and its applications. *Genomics, Proteomics & Bioinformatics*, 13(5).
 21. Sanger, F., Nicklen, S., Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74.
 22. Shih, S.Y., Bose, N., Goncalves, A., Erlich, H., Calloway, C. (2018). Applications of probe capture enrichment next generation sequencing for whole mitochondrial genome and 426 nuclear SNPs for forensically challenging samples. *Genes* (Basel), 9(1).
 23. Silva, S., Lopes, C., Teixeira, A.L., Sousa, M., Medeiros, R. (2014). Forensic miRNA: Potential biomarker for body fluids. *Forensic Science International: Genetics*, 14, <https://doi.org/10.1016/j.fsigen.2014.09.002>
 24. Van Dijk, E.L., Auger, H., Jaszczyszyn, Y., Thermes, C. (2014). Ten years of next-generation sequencing technology. *Trends in Genetics*, 30(9).
 25. Yang, Y., Xie, B., Yan, J. (2014). Application of next-generation sequencing technology in forensic science. *Genomics, Proteomics & Bioinformatics*, 12(5).
 26. Zbieć-Piekarska, R., Spólnicka, M., Kupiec, T., Parys-Proszek, A., Makowska, Ż., Pałeczka, A., Kucharczyk, K., Płoski, R., Branicki, W. (2015). Development of a forensically useful age prediction method based on DNA methylation analysis. *Forensic Science International: Genetics*, 17.

Translation Rafał Wierchośławski