Czasopismo Polskiego Towarzystwa Andrologicznego
Journal of Polish Society of Andrology
Postępy Andrologii *Online*
Advances in Andrology *Online*
http://www.postepyandrologii.pl

NEXT-GENERATION SEQUENCING ANALYSIS IN MALE INFERTILITY DIAGNOSTICS

Marzena Skrzypczak-Zielinska*, Monika Fraczek*

Institute of Human Genetics, Polish Academy of Sciences, Poznan, Poland

*Corresponding authors:
Marzena Skrzypczak-Zielinska, Institute of Human Genetics, Polish Academy of Sciences,
str. Strzeszynska 32, 60–479 Poznan, Poland
tel.: +48 61 65 79 201, e-mail: marzena.skrzypczak-zielinska@igcz.poznan.pl

Monika Fraczek, Institute of Human Genetics, Polish Academy of Sciences,
str. Strzeszynska 32, 60–479 Poznan, Poland
tel.: +48 61 6579 231, e-mail: monika.fraczek@igcz.poznan.pl

Otrzymano/received: 22.12.2021 r. • Zaakceptowano/accepted: 31.01.2022 r.

DOI: 10.26404/PAO_2353-8791.2021.06

Marzena Skrzypczak-Zielinska – PhD, DSc, graduated from Biotechnology at the University of Life Sciences in Poznan, Poland. After graduating, she was internship within the Leonardo da Vinci Programme at the Institute of Human Genetics of the Leipzig University, where she also finished her PhD study and received the doctor title of medical sciences. Since 2010, she has been an assistant professor at the Institute of Human Genetics Polish Academy of Sciences in Poznan, Poland. She authored and co-authored over 40 international scientific publications. Her research focuses on human molecular genetics, pharmacogenetics, and microbiota analysis. Membership in organizations: Polish Society of Human Genetics, Polish College of Laboratory Medicine, and National Chamber of Laboratory Diagnosticians.

Abstract

Genetic analyses are essential in determining the causes of reproductive failure in couples. Currently, hereditary factors are considered responsible for at least 15% of male infertility cases. Karyotype analysis, along with Y-chromosome microdeletion screening and cystic fibrosis transmembrane conductance regulator (*CFTR*) gene mutation testing, has become the standard. Chromosome aberrations occur in 7% of infertile men, 30 times more frequently than in the general population. The cause of male infertility is also microdeletions of the Y chromosome or aberrations and mutations of genes responsible for male sexual development, e.g., located in the Yp11.2 region, and *CFTR* gene mutations analysis is performed by patients with bilateral absence or obstruction of the vas deferens. However,

in recent years, modern technologies have significantly increased the possibilities of genetics in diagnosing male infertility. Next-generation sequencing (NGS) is one of them. It has many applications depending on the size and type of the analyzed sequence of genetic material. In this article, we provide an overview of different next-generation sequencing applications, i.e., whole-genome sequencing (WGS), whole exome sequencing (WES), gene panel (targeted) sequencing, RNA sequencing (RNA-Seq), epigenome sequencing as well as metagenomic analysis of semen microbiota, in male infertility research and diagnostics. Through NGS studies, increased diagnostic efficiency is observed, especially in men with nonobstructive azoospermia and idiopathic infertility, where the cause of the disorder is a rare gene variant. Currently, the barrier to introducing next-generation sequencing into the routine diagnosis of male infertility is primarily the amount of data generated and problems with their interpretation, and the indication of clinically relevant gene mutations. An obstacle to developing the next-generation sequencing gene panel is the many genes involved and the participation of rare variants in spermatogenesis. Nevertheless, we expect next-generation sequencing analyses to be the future of broadly understood molecular and genetic diagnostics in infertility research. Shortly, they will probably completely replace the existing standard tools.
Keywords: male infertility, next generation sequencing (NGS), genetics, semen, microbiota

Skróty / Abbreviations

AR – androgen receptor gene; *AURKC* – aurora kinase C gene; AZF – azoospermia factor; CBAVD – congenital bilateral absence of the vas deferens; CF – cystic fibrosis; *CFTR* – cystic fibrosis transmembrane conductance regulator gene; CNVs – copy number variations; *DMRT1* – doublesex and mab-3 related transcription factor-1; *DPY19L2* – Dpy-19 like 2; *FSHB* – follicle-stimulating hormone subunit beta gene; *FSHR* – follicle-stimulating hormone receptor gene; GWAS – genome-wide association studies; *KAL1* – Kallmann syndrome 1 gene; *KLHL10* – kelch-like family member 10 gene; lncRNA – long non-coding RNA; miRNAs – microRNA; mRNA – messenger RNA; *NANOS1* – nanos C2HC-type zinc finger 1 gene; NGS – next-generation sequencing; *NR5A1* – nuclear receptor subfamily 5 group A member 1 gene; PCR – polymerase chain reaction; PGD – preimplantation genetic diagnostics; piRNA – piwi-interacting RNA; rRNA – ribosomal RNA; *SEPT12* – septin 12 gene; siRNA – small-interfering RNA; SNVs – single nucleotide variants; SRY – sex determining region Y; *SYCP3* – synaptonemal complex protein 3 gene; *TEX11* – testis expressed 11 gene; tRNA – transfer RNA; VUS – variants of uncertain significance; WES – whole exome sequencing; WGS – whole genome sequencing

Introduction

In about half of the cases, the reproductive failure of couples is of male origin. Male subfertility/infertility is due to abnormalities in spermatogenesis leading to one or a combination of low sperm count, poor sperm motility, or abnormal morphology. The etiology of these various pathological semen phenotypes is multifactorial and still not fully understood. Currently, genetic factors are detected in approximately 10-15% of male infertile patients (*Ambulkar et al.*, 2014). However, it can be seen, as molecular biology advances, and new techniques are developed, that genetic analysis becomes more and more important in explaining the causes of male infertility. Modern technologies open up new fields of research in diagnostics. Next-generation sequencing (NGS) is one such technology, which has revolutionized nearly every area of biotechnology and medicine. NGS supports detection and understanding of genomics variations, disease mechanisms, and immunity through numerous applications and the resulting data, which helps develop better diagnostics and therapy. What are the benefits of NGS in the diagnosis of male infertility? What applications of the NGS technique are used in this field? What was the genetic diagnostic of male infertility before the NGS era, and what can we expect soon? This paper aims to answer the above questions and provide an overview of NGS analysis in male infertility research.

Male infertility genetic diagnostics before the NGS era

So far, routinely applied genetic tests for diagnostic purposes of male infertility include primarily karyotyping, Y-chromosome microdeletion screening, and cystic fibrosis transmembrane conductance regulator (*CFTR*) gene mutation testing. Detailed genetic diagnostic scheme in male infertility is presented in Figure 1. Procedures for those tests are officially approved and unified in guidelines of the European Association of Urology, American Urologic Association, and American Society of Reproductive Medicine (*Jungwirth et al.*, 2019; *Practice Committee of the American Society for Reproductive Medicine*, 2015; *Schlegel et al.*, 2021).

Karyotyping

In infertile men, the assessment of karyotype aberrations was started in 1966 and after G-banding modification remains a standard to date (*Van Assche et al.*, 1996) (Figure 2). A survey of pooled data from eleven publications, including 9,766 infertile men, demonstrated the presence of chromosome aberrations in 5.8% (*Johnson*, 1998). Some results indicate those karyotype aberrations occur with an even higher frequency of about 7–8% among infertile men, which is 20–30-times more

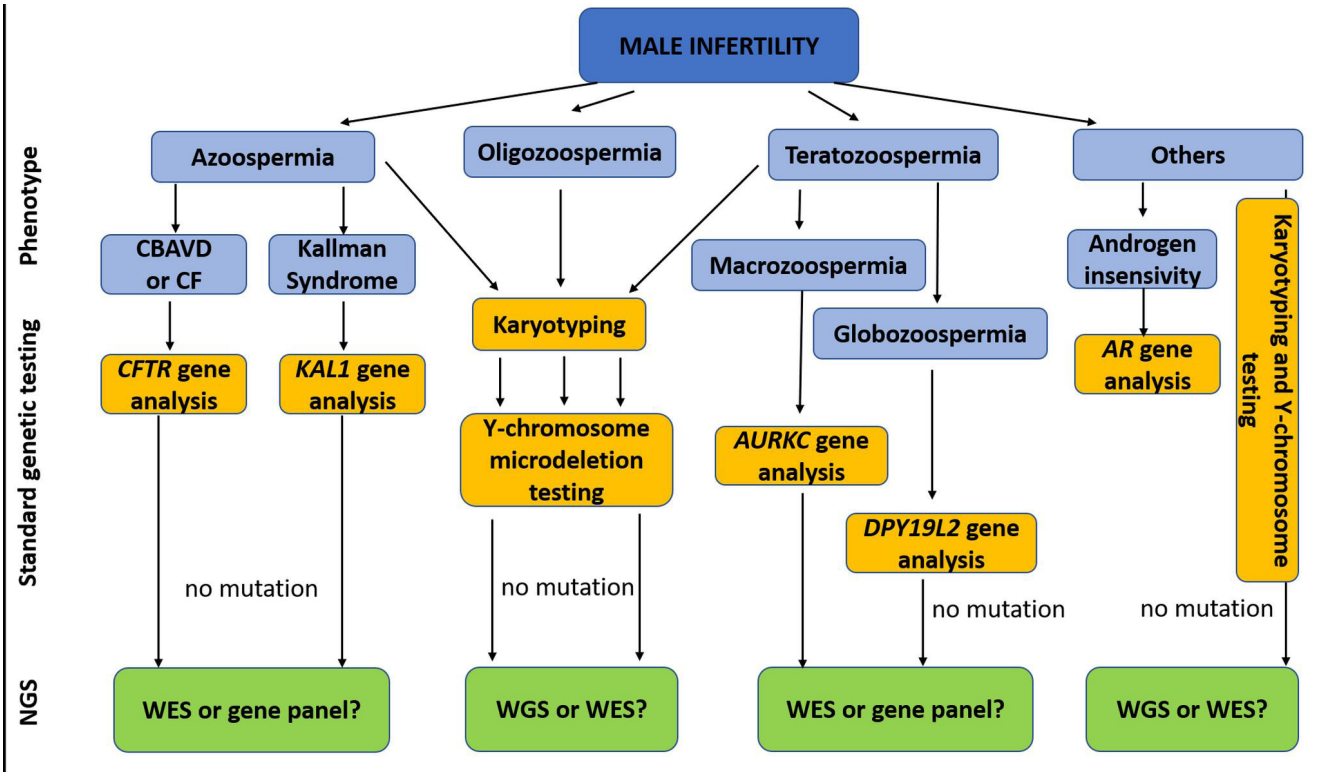

Fig. 1. Male infertility genetic diagnostics scheme in pre- and NGS time. *AR* – androgen receptor gene; *AURKC* – aurora kinase C gene; CBAVD – congenital bilateral absence of the vas deferens; CF – cystic fibrosis; *CFTR* – cystic fibrosis transmembrane conductance regulator gene; *DPY19L2* – Dpy-19 like 2; *KAL1* - Kallmann syndrome 1 gene; NGS – next-generation sequencing; WES – whole exome sequencing; WGS – whole genome sequencing (details in the text)

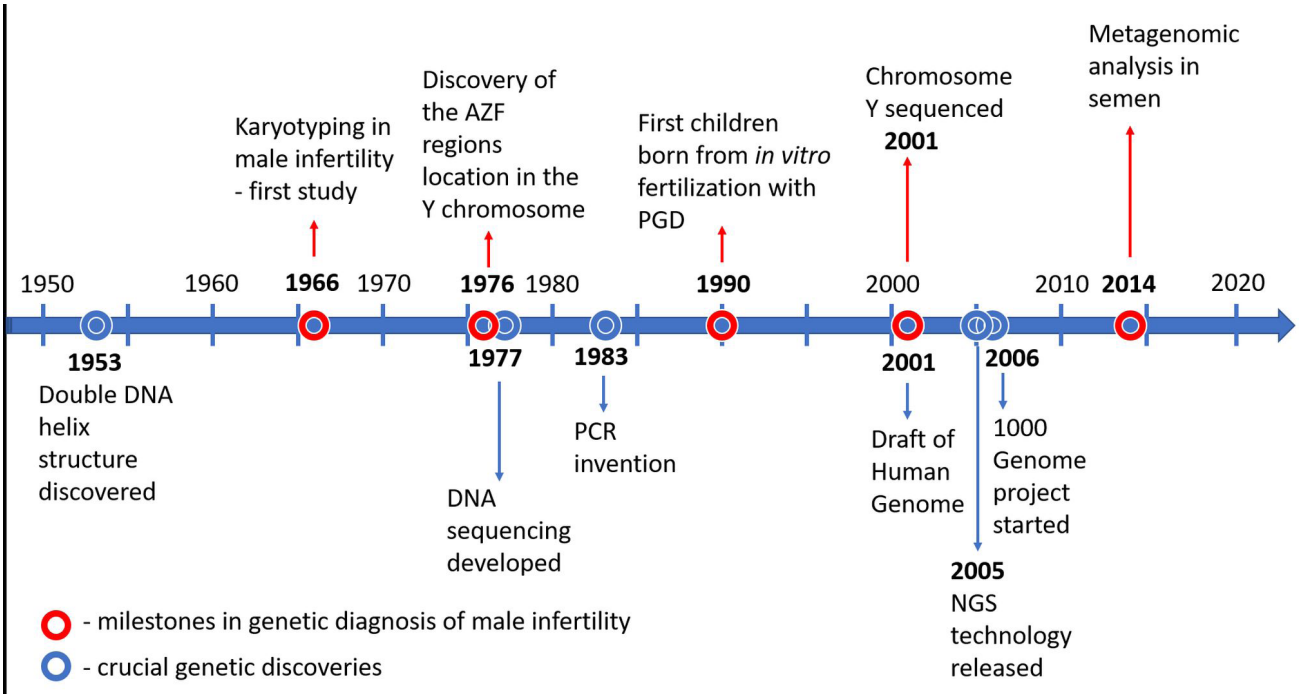

Fig. 2. The main discoveries over time influencing the improvement of genetic diagnosis of male infertility. AZF – azoospermia factor; NGS – next-generation sequencing; PCR – polymerase chain reaction; PGD – preimplantation genetic diagnostics

frequently than in the general population (*Ravel et al.*, 2006). Chromosome abnormalities are divided into numerical or structural and can affect both autosomes (accounted for 4.2%) and sex chromosomes (for 1.5%) (*Krausz et al.*, 2015). The most commonly detected chromosomal defect in infertile men with nonobstructive azoospermia is Klinefelter syndrome, which contains an aberrant supernumerary X chromosome 47,XXY in the karyotype 80–90% of cases and its variants like mosaic 47,XXY/46,XY. The remaining Klinefelter syndrome patients have karyotype 48,XXXY, 48,XXYY (*Bojesen et al.*, 2003). Another sex chromosome aberration

is 46,XX male syndrome, known as de la Chappelle syndrome, occurring when Y chromosomal material with the sex-determining region Y (*SRY*) gene is translocated onto another, usually autosomal, chromosome. Karyotyping also enables the detection of balanced or unbalanced translocations, insertions, deletions, and inversions. It is worth emphasizing that balanced structural chromosomal aberrations (inversions and translocations) are more common in infertile men with oligozoospermia (*Krausz* and *Riera-Escamilla*, 2018). On the other hand, the conventional karyotyping is very laborious and has a significant limitation, relatively poor resolution, which enables the detection of aberration larger than five megabases (Mb) (*Van Assche et al.*, 1996; *Pelzman* and *Hwang*, 2021).

Y-chromosome microdeletion testing

The discovery of the polymerase chain reaction (PCR) technique in 1983 and location of factors controlling spermatogenesis allowed the extension of genetic diagnostics in male infertility by Y chromosome microdeletions analysis (Figure 2). The defect in azoospermia factor (CBAVD – congenital bilateral absence of the vas deferens) regions containing three *loci*: AZFa, AZFb, and AZFc occurs in 3% to 15% of men with non-obstructive azoospermia and in 6% to 8% of severely oligozoospermic men depending on the patient selection criteria and technique used, where heterogeneous AZFc region deletions are the most common (*Abur et al.*, 2019; *Gunes* and *Esteves*, 2021). Y chromosome microdeletions rate in the general population is 0.025%. According to the European Academy of Andrology and European Molecular Genetics Quality Network guidelines, the screening of Y-chromosome microdeletions should include multiplex PCR to analyze two sequence-tagged sites *loci* in each region of the AZFa, AZFb, and AZFc regions (*Krausz et al.*, 2014). The recommended primer sets contain sY14 (SRY), ZFX/ZFY, sY84 and sY86 (AZFa), sY127 and sY134 (AZFb), and sY254 and sY255 (AZFc). The molecular diagnosis using these methodology enables the detection of over 95% of the deletions sufficient for routine screening (*Krausz et al.*, 2014).

CFTR gene mutation testing

The third part of standard genetics tests in male infertility refers to patients with bilateral absence or obstruction of the vas deferens and checking for mutations in the *CFTR* gene located in chromosome 7 (7q31.2). CBAVD accounts for 2–6% of male infertility individuals and up to 25% of patients of obstructive azoospermia (*Cui et al.*, 2020). *CFTR* gene variants also lead to the onset of cystic fibrosis (CF). The most common mutation of this gene is p.F508del which results in impaired folding and trafficking of the encoded protein. The frequency of deletion p.F508 in patients with CBAVD in Europe ranges from 35 to 74.6%, depending on the population (*Cui et al.*, 2020). The *CFTR* gene is relatively large, consisting of 24 exons (OMIM 602421); hence sequence analysis is often performed in stages or using screening methods. A faster alternative is to use targeted sequencing with NGS technology, as described in a separate section of this manuscript.

From the observations so far it appears that complete diagnostic work-up (including also genetic testing) provides an explanation of the causes of male infertility in about 60% up to 72% of primary testicular failure cases (*Krausz et al.*, 2015; *Kothandaraman et al.*, 2016; *Cannarella et al.*, 2019). In other words, in the rest 40% of the patients, the etiology remains unknown and is referred to as idiopathic infertility. Therefore, numerous patients do not know the underlying cause of their disease. According to the data described below, the detection of genetic factors can be increased using high-throughput techniques for DNA and RNA analysis. Such techniques include NGS and are most relevant to patients with severe azoospermia possessing rare gene mutations.

Next generation sequencing (NGS)

The first available NGS technology was released in 2005 and is based on the pyrosequencing method. Four other NGS methodologies were developed within the next five years that took over the market (*Van Dijk et al.*, 2014). NGS-based approaches have also quickly gained broad applicability in medicine, from genetic diagnosis and disease networks to drug discovery and pharmacogenomics (*Chaitankar et al.*, 2016).

Recent reports based on studies using the NGS method showed that the human genome is 200 000 base pairs longer than the first human genome sequencing data 20 years ago. This achievement was possible due to the current technological possibilities, which fill gaps in the genome reading containing challenged sequences, such as repeat-rich telomeres. The Telomere-to-Telomere (T2T) Consortium comprising 30 institutions indicated that the human genome is just over 3.2 billion base pairs, and they discovered about 115 new genes (*Reardon*, 2021). The most extensive NGS equipment can sequence this amount of DNA in about one day instead of 13 years as it was the case in the first human genome sequencing (started on October 1, 1990, and completed in April 2003; Figure 2). Also, the cost has sharply decreased over the 20 years from about $ 100 million to just about $ 1,000 (data from the *National Human Genome Research Institute*, https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data). However, diagnostics will certainly not be used for a long time from genome-wide analysis due to the amount of data generated and the requirement of enormous computing power resources

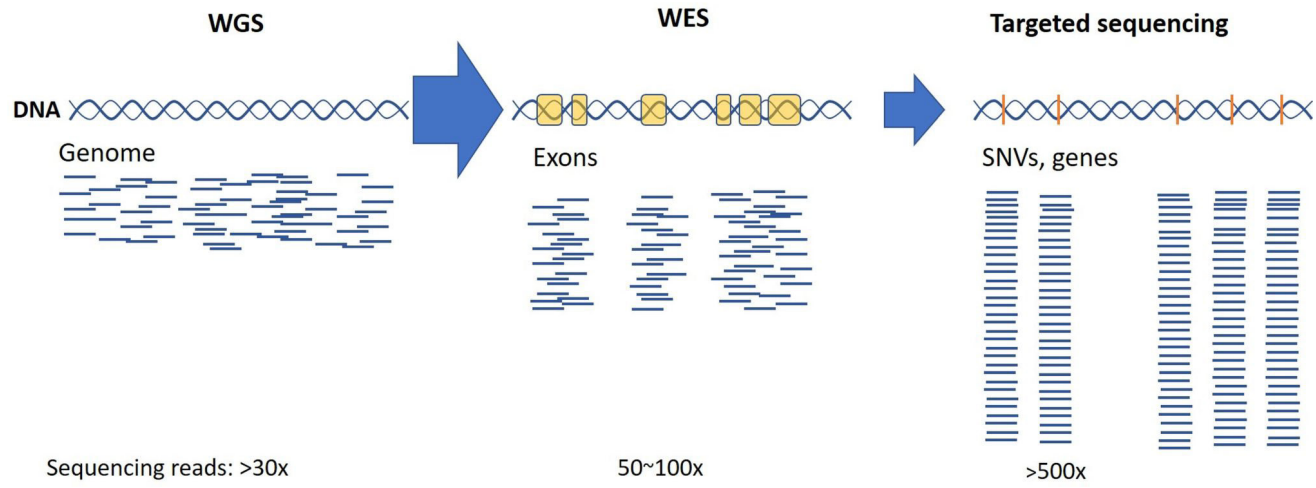

Fig. 3. Next-generation sequencing main applications. SNVs – single nucleotide variants; WES – whole exome sequencing; WGS – whole genome sequencing (details in the text)

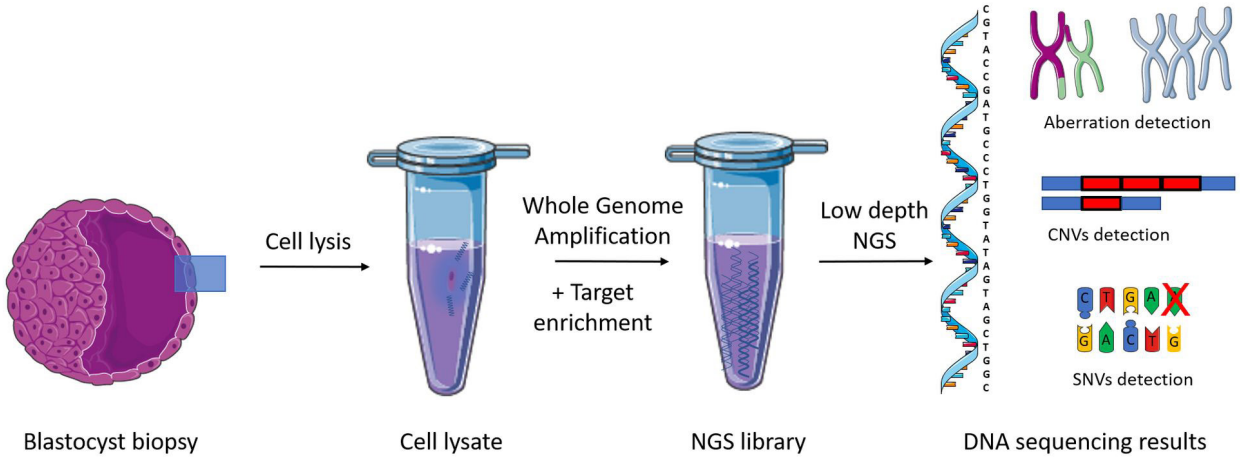

Fig. 4. Preimplantation genetic testing using NGS. CNVs – copy number variations; NGS – next-generation sequencing; SNVs – single nucleotide variants (details in the text)

for their analysis and interpretation. Genetic diagnostics with the use of NGS is limited to selected more functionally informative elements of the human genome like protein-coding regions (exome), which is called whole exome sequencing (WES), or sequences of selected genes (targeted sequencing). Narrowing down the sequencing target allows for more readings exclusively for regions of interest (*loci*) (Figure 3).

Whole genome sequencing in male infertility

Currently, the sequencing of entire genomes is most often used in scientific research for genome-wide association studies (GWAS) to select genes and single nucleotide variants (SNVs) that significantly differentiate a group of patients from healthy people. This NGS application has enormous clinical potential to indicate the possible genetic cause of the occurrence of a given phenotype, in this case, male fertility disorder, without a prior hypothesis of genomic location (*Aston*, 2014). From a genetic point of view, it can be concluded that NGS, due to its high potential to discover unknown gene variants, its flexibility, and no need for reference material, is replacing the earlier microarray technique. Microarrays are based on the DNA hybridization property and have been used in male infertility research since 2009 (*Aston* and *Carrell*, 2009, *Aston et al.*, 2010, *Hu et al.*, 2012) whereas the GWAS in male infertility with NGS was first published in 2013 (*Xu et al.*, 2013). It is challenging to discuss the obtained results in detail because of a large amount and diversity of the observed correlations. However, overall, the data obtained in the GWAS studies mentioned above indicated an association from several (*Xu et al.*, 2013, *Hu et al.*, 2012) to over 170 *loci* connected with azoospermia or oligospermia (*Aston et al.*, 2010); and in all the studies, there is an indication to verify these results in further investigations on large groups of patients. For this purpose, e.g., targeted sequencing can be used by constructing a panel of genes at a pre-diagnostic stage.

Another essential aspect of whole-genome sequencing (WGS) is its use for karyotyping. NGS could combine the reading of the entire nucleotide sequence with a resolution of 1 base pair with quantification of DNA material, i.e., numerical aberration of chromosomes. The research shows that the low-coverage WGS can also reliably detect balanced translocations and precisely map breakpoints compared with conventional procedures. Therefore, WGS and other NGS applications can replace cytogenetic methods in analyzing chromosomal aberrations, including the diagnosis of clinically balanced translocation carriers (*Liang et al.*, 2017). An example of a karyotype assessment with the use of NGS in terms of infertility is preimplantation genetic diagnostics (PGD) as part of the *in vitro* procedure to reduce the risk of severe chromosomal defects in the embryo (before a transfer) and in the detection of monogenic diseases and disorders related to the occurrence of translocation in the offspring (*Łukaszuk et al.*, 2015; *Fiorentino et al.*, 2014) (Figure 4). Currently, PGD research aims to validate the sensitivity of the NGS methodology concerning cytogenic methods, like the detection of chromosomal aneuploidies and mosaicism degree in preimplantation embryos. *García-Pascual* and co-workers (*García-Pascual et al.*, 2020) demonstrated that specificity and sensitivity for preimplantation genetic testing for aneuploidy were both 100% for whole chromosomes, segments (≥10 Mb), small rearrangements (del/dup ≥a6 Mb), and mosaicism degree, wherein the thresholds established for mosaicism were: euploid embryos (<30% aneuploidy), low mosaic (from 30% to <50%), high mosaic (50–70%) or aneuploid (>70%). In the future, NGS can be also promising approach towards a non-invasive preimplantation genetic testing of aneuploidy in the embryonic cell free DNA released to the culture media at blastocyst stage (*Rubio et al.*, 2019).

Whole exome sequencing in male infertility

Whole-exome sequencing is a widely used NGS method for genetic studies, primarily for disease gene identification and clinical diagnosis that involves capturing and sequencing only the protein-coding regions (the exomes) of the genome. The human exome represents less than 2% of the genome but contains about 85% of known disease-related variants, making this application more economical compared to whole-genome sequencing (*Choi et al.*, 2009). Therefore, WES enables the effective identification of coding variants and splice site variants.

This analysis is especially justified in nonobstructive azoospermia or severe oligozoospermia patients who have undergone numerous laboratory tests and are still idiopathic. *Fakhro* and co-workers (*Fakhro et al.*, 2018) observed that the WES analysis explains the genetic etiology in a significant part of these patients (>50% in the family and >10% sporadic). The results of WES research on a large cohort of 285 males with severe oligospermia and nonobstructive azoospermia, in 10.5% (n = 30) confirmed chromosomal aberrations while 24.2% (n = 69) of the cases had a potential monogenic form of male infertility. Therefore, the authors concluded, that the standard approach to male infertility using cytogenetic karyotyping and chromosome Y microdeletion testing overlooked the significant contribution of monogenic causes of male infertility. Moreover, they added 33 novel candidate genes to the list of about 400 plausible causal genes of male infertility (*Alhathal et al.*, 2020). It is worth emphasizing that the WES, like the WGS, provides an opportunity to simultaneously detect SNVs, copy number variations (CNVs) and chromosomal aberration. The identification of rare variants in various genes underlying male infertility, on the one hand, presents problems in the development of an effective diagnostic scheme, and on the other hand, proves the benefits and effectiveness of WES. On the contrary, the disadvantage of this NGS application is the possibility of finding additional gene defects unrelated to infertility, such as cancer predisposition or early-onset neurodegenerative disease. Hence, very importantly, the implementation of WES and WGS also should always be prudent and preceded by genetic counseling (*Ghieh et al.*, 2021).

Panel gene (targeted) sequencing in male infertility

Next generation sequencing, apart from extensive genomic analyses, also offers applications for the analysis of selected DNA fragments, e.g. important genes or gene regions associated or probably associated with a disease or phenotype. Panel gene sequencing (called targeted sequencing) is dedicated especially to the diagnostic purposes or to the second stage (after WGS or WES) of research on various diseases. In the laboratory procedure for the preparation of genetic material for sequencing, a sample enrichment step is required, in which the PCR amplification of the target gene sequences is performed. The scheme of the procedure is shown in Figure 5, which in practice takes an average of a few days in a laboratory, depending on the number of targeted genes and the number of analyzed samples.

Nearly 2,300 genes are believed to be involved in spermatogenesis (*Krausz* and *Riera-Escamilla*, 2018). Knowledge in this area is constantly expanding. *Cannarella* and co-authors presented a broad list of the most critical genes responsible for spermatogenic failure with their detailed descriptions. They also proposed a phenotype-based approach taking into account flagella abnormalities, macrozoospermia, acephalic spermatozoa, globozoospermia, oligozoospermia, oligoasthenozoospermia, asthenozoospermia, nonobstructive azoospermia, and oligoasthenoteratozoospermia

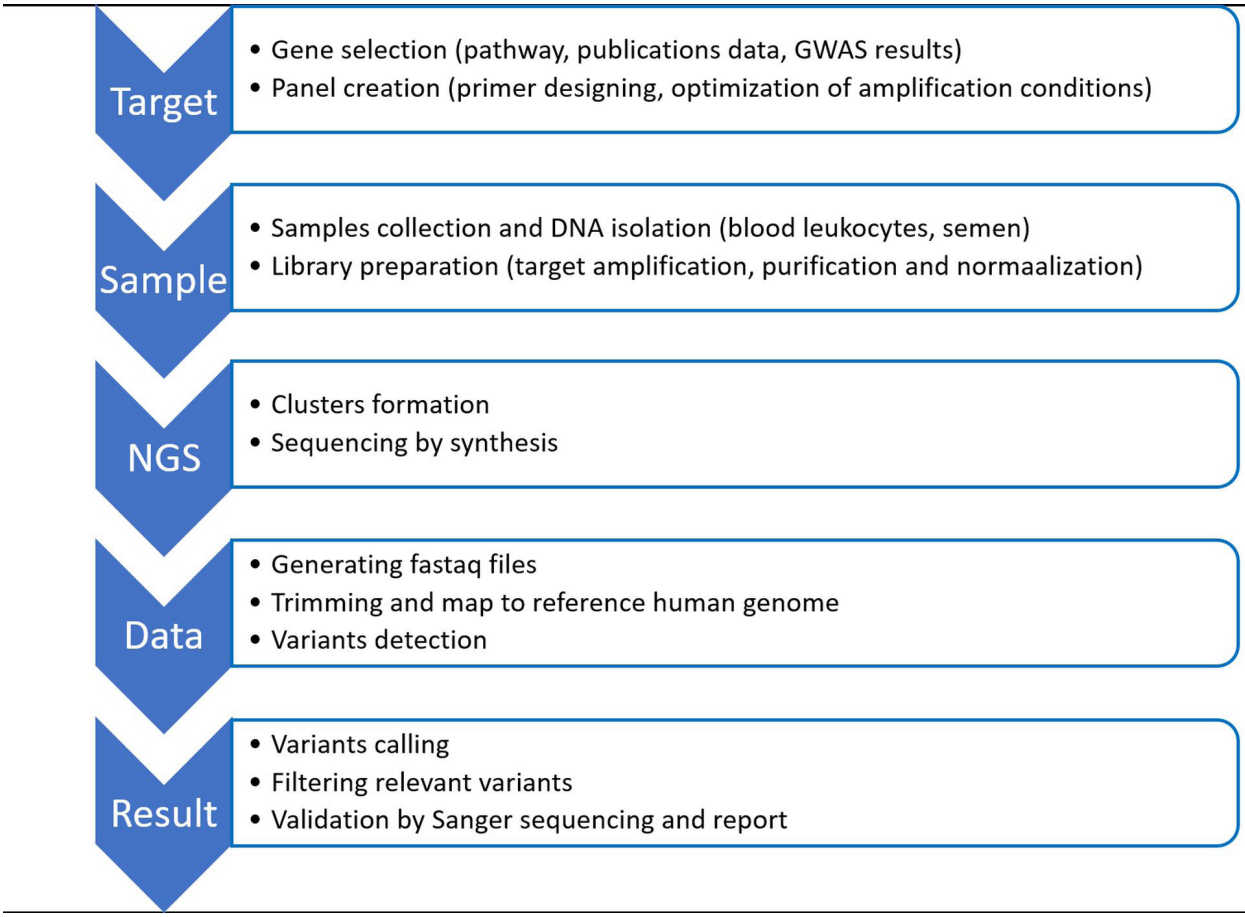

Fig. 5. Targeted sequencing workflow and bioinformatic pipeline. GWAS – genome-wide association studies; NGS – next-generation sequencing (details in the text)

to identify targeted genes (*Cannarella et al.*, 2019). After careful analysis of the data in the literature, the authors argue that only part of this enormous group of genes, numbering about 60, is responsible for spermatogenesis failures. These are their proposals for a panel that needs to be tested on a large group of azoospermia and severe oligozoospermia patients to verify this choice (*Cannarella et al.*, 2019). In contrast, recently, in another review, a list of 38 candidate genes involved in monogenic causes of human nonobstructive azoospermia was compiled. The authors classified and described these genes according to the related testicular histology underlying the non-obstructive azoospermic phenotype (*Cioppi et al.*, 2021). Based on the list presented in the review by *Cannarella* and co-workers (*Cannarella et al.*, 2019), the pre-diagnostic gene panel was constructed in practice and contained 65 genes (*Precone et al.*, 2021). In this research, variants of potential causative importance (two pathogenic, three variants of uncertain significance (VUS), and three variants with high impact) were identified in 10 out of 12 infertile patients (83%). Interestingly, almost half of those variants belong to the cytoplasmic dynein genes. However, this study has serious limitations due to a very small group of patients (*Precone et al.*, 2021).

In research, in a cohort of 241 infertile idiopathic males with a negative result in standard genetic testing, a panel of the following nine genes were analyzed: androgen receptor (*AR*), follicle-stimulating hormone receptor (*FSHR*), follicle-stimulating hormone subunit beta (*FSHB*), kelch-like family member 10 (*KLHL10*), nuclear receptor subfamily 5 group A member 1 (*NR5A1*), nanos C2HC-type zinc finger 1 (*NANOS1*), septin 12 (*SEPT12*), synaptonemal complex protein 3 (*SYCP3*) and testis expressed 11 (*TEX11*). The authors identified 19 missense variants in 23 patients, constituting nearly 10% of the study group. Five identified variants (26.3%, 5/19) were classified as likely benign, two (10.5%, 2/19) as benign, eleven (57.9%, 11/19) as VUS, and one variant (5.3%, 1/19) as pathogenic (*Rocca et al.*, 2020). The results are deliberately referred to as potential variants of male infertility due to the difficulty in interpreting new genetic changes. The causal mutation ratio in earlier studies of *NR5A1*, doublesex and mab-3 related transcription factor-1 (*DMRT1*), and *TEX11* genes sequencing in 80 azoospermic patients increased from 5% in standard diagnostics testing up to 25% (*Tüttelmann et al.*, 2018). The data obtained justify the utility of applying panel gene sequencing for infertility diagnosis to find new genetic variants potentially linked to male infertility.

It is difficult to predict how much the yield of genetic testing will increase for the different male infertility phenotypes. However, it is realistic to expect the overall

diagnostic efficiency, independently of subtypes of male infertility, will rise from the current 4% (*Tüttelmann et al.*, 2018) to more than 10% in the coming few years, as *Oud* and co-authors announced (*Oud et al.*, 2019). The consensus of European Society of Human Genetics and the European Society for Human Reproduction and Embryology includes an opinion that NGS gene panels may soon become a useful tool for the identification of additional causes of male infertility, and thus improve genetic- and reproductive counselling, facilitate patient stratification and therefore enable more precise assisted reproductive technology approaches (*Harper et al.*, 2018). It should be borne in mind that knowledge in this field is still at the stage of discovery and verification, which genes can actually be the source of frequent mutations and constitute a potential marker of male infertility.

At the end of this subsection, it is worth noting that targeted sequencing also provides a valuable tool to uncover the structural variability of highly dynamic regions of the human genome, such as the Y chromosome. According to recent research on semen samples from 222 infertile patients with abnormal semen parameters, after standard diagnosis, the Y chromosome structural variations in 31.88% of the cases have been overlooked (*Liu et al.*, 2021). The NGS method can replace the standard Y chromosome testing and enables the detection of new CNVs; nevertheless, its clinical significance remains often unknown.

RNA sequencing and epigenetics analysis in male infertility

RNA-sequencing (RNA-Seq) with NGS is becoming a common tool for transcriptome profiling. A transcriptome is a complete set of transcripts at a given time in analyzed cells, containing the messenger RNA (mRNA), long non-coding RNAs (lncRNA), and small RNAs including microRNA (miRNA), small-interfering RNA (siRNA), transfer RNA (tRNA), ribosomal RNA (rRNA) and piwi-interacting RNA (piRNA). The RNA-Seq provides comprehensive information on the quantitative level of individual genes and structural changes in the transcriptome, including the gene fusions and alternative splicing (*Wang et al.*, 2009). MiRNA molecules are profusely abundant in the seminal plasma of men and participate in the control of each stage of male germ cell differentiation during spermatogenesis (*Kotaja*, 2014). Therefore, the expression of various RNA molecules could be helpful as potential biomarkers of male infertility (*Barceló et al.*, 2018).

Previous scientific results of broad studies using RNA-Seq demonstrated significant differences in spermatic miRNA expression profiles correlated with sperm DNA fragmentation index (*Li et al.*, 2020) or the occurrence of azoospermia (*Tang et al.*, 2018; *Cheung et al.*, 2019). A very advanced study of the transcriptome with the co-expression analysis of various types of RNA particles conducted on 25 patients with asthenozoospermia and 25 male normal healthy individuals indicated lncRNA-mRNA-miRNA regulation networks that may be involved in the pathological mechanisms of asthenozoospermia (*Lu et al.*, 2020). However, relatively few of these experiments in male infertility research have been described.

It is worth noting that the regulation of transcription belongs to important epigenetic mechanisms after DNA methylation, remodeling of chromatin, and histone tail modifications. Epigenetic biomarkers, which can incorporate information from the genetic background, lifestyle, and environmental factors, are extremely interesting and potentially crucial in elucidating the cause of idiopathic male infertility (*Gunes* and *Esteves*, 2021). Available epigenetic studies in male infertility mainly concern human sperm DNA methylome analysis because DNA methylation is a heritable epigenetic modification of cytosine residues within CpG dinucleotides and less frequently of adenine residues in non-CpG sites. This pattern is determined using bisulfite conversion sequencing, which can be done as target enrichment or whole-genome bisulfite sequencing. This approach makes it possible to evaluate, e.g., the DNA methylation profile of human sperm during aging. *Cao* and co-authors observed thousands of age-related and sperm-specific epigenetic alterations in their investigation. These findings show massive human sperm DNA methylation dynamics during paternal aging, which can subsequently affect genes potentially related to diseases in offspring (*Cao et al.*, 2020). From a laboratory perspective, global analysis of methylation profile in sperm is challenging, as many external factors can distort the results. However, it may help identify modifiable risk causes of male infertility in the future.

Semen microbiota in male infertility

The diagnosis of infertility uses the NGS also for the characteristics of the microbiome of the reproductive system. As the human body contains more microbes than human cells, the microbiome (the 'second human genome') has a vast potential to influence human physiology (*Tsonis et al.*, 2021). To date, the use of NGS supports the explanation of the functional, quantitative, and mechanical aspects of the complex microorganism-host interactions, particularly those concerning the gut microbiome. Bacterial identification methods performed in previous studies were either PCR-based or culture methods. NGS metagenomic analysis based on bacterial 16S rRNA gene sequencing revolutionized microbiome research. A significant advantage of NGS over traditional methods is that it enables the detection and classification of all microorganisms according to taxa down to the species level. This technology has changed the knowledge of the percentage of microbiota, as it also detects and characterizes taxa that are uncultured

in vitro (they do not survive outside the host organism). It is particularly surprising that quantitatively as many as 97% of the microorganisms of the semen microbiota are uncultured or unknown species (unpublished own research). The sperm microbiome is an area of growing scientific interest due to important implications for male reproductive health, the health of couples, and even the health of the offspring by transmitting microorganisms to the partner and offspring (*Chen et al.*, 2018; *Baud et al.*, 2019, *Altmäe et al.*, 2019).

The first analysis approach using NGS technology to investigate the associations between bacterial communities and semen quality was performed in 2014 for 96 semen samples (*Weng et al.*, 2014). The results showed that the most abundant genera among all the samples were *Lactobacillus* (19.9%), *Pseudomonas* (9.85%), *Prevotella* (8.51%), and *Gardnerella* (4.21%). The proportion of *Lactobacillus* and *Gardnerella* was significantly higher in the normal samples. At the same time, *Prevotella* abundance was significantly higher in the low-quality semen samples. An indication of a particular species of bacteria impacting sperm function may facilitate the development of new therapies (e.g., probiotics). However, the current data are often not repetitive, inconclusive, and concern small groups of patients. Therefore more research is needed (*Farahani et al.*, 2021).

Final remarks

The implementation of NGS analysis in male infertility diagnostics is currently the subject of intense debate. This technology offers many benefits, opens up new possibilities, and makes diagnostics more effective, but on the other hand, there are barriers to its introduction into everyday practice. This technology, through improvements, enables not only reading the sequence of genetic material and the detection of SNVs but also quantification to detect CNVs and chromosomal aberrations, including balanced translocations. The standard approach to male infertility using classical karyotyping and chromosome Y microdeletion testing overlooks the monogenic causes of male infertility. WGS and other NGS applications will probably soon replace cytogenetic methods in analyzing chromosomal aberrations by raising the resolution and enriching diagnostics to detect rare point variants underlying male infertility. A barrier against this breakthrough is generating a lot of data in NGS, which needs enormous computing power resources and experienced bioinformatics staff. NGS panel genes seem to be a solution to this problem; however, there are no selected genes that can be used diagnostically. The clinical significance of many detected variants often remains unknown. In approximately 40% of male patients, the etiology of reproduction failure remains undetermined. NGS research also in sperm epigenetics and microbiota may significantly contribute to explaining the cause of idiopathic male infertility.

References

Abur U., Gunes S., Ascı R., Altundag E., Akar O.S., Ayas B. et al.: Chromosomal and Y-chromosome microdeletion analysis in 1,300 infertile males and the fertility outcome of patients with AZFc microdeletions. Andrologia. 2019, 51, e13402. doi: 10.1111/and.13402. PMID: 31650616.

Alhathal N., Maddirevula S., Coskun S., Alali H., Assoum M., Morris T. et al.: A genomics approach to male infertility. Genet Med. 2020, 22, 1967–1975. doi.org/10.1038/s41436-020-0916-0. PMID: 32719396.

Ambulkar P.S., Sigh R., Reddy M., Varma P.S., Gupta D.O., Shende M.R. et al.: Genetic risk of azoospermia factor (AZF) microdeletions in idiopathic cases of azoospermia and oligozoospermia in central indian population. J Clin Diagn Res. 2014, 8, 88–91. doi:10.7860/JCDR/2014/7680.4116. PMID: 24783090.

Altmäe S., Franasiak J.M., Mändar R.: The seminal microbiome in health and disease. Nat Rev Urol. 2019, 16, 703–721. doi: 10.1038/s41585-019-0250-y. PMID: 31732723.

Aston K.I.: Genetic susceptibility to male infertility: news from genome-wide association studies. Andrology. 2014, 2, 315–321. doi: 10.1111/j.2047-2927.2014.00188.x. PMID: 24574159.

Aston K.I., Carrell D.T.: Genome-wide study of single-nucleotide polymorphisms associated with azoospermia and severe oligozoospermia. J Androl. 2009, 30, 711–725. doi: 10.2164/jandrol.109.007971. PMID: 19478329.

Aston K.I., Krausz C., Laface I., Ruiz-Castane E., Carrell D.T.: Evaluation of 172 candidate polymorphisms for association with oligozoospermia or azoospermia in a large cohort of men of European descent. Hum Reprod. 2010, 25, 1383–1397. doi: 10.1093/humrep/deq081. PMID: 20378615

Barceló M., Mata A., Bassas L., Larriba S.: Exosomal microRNAs in seminal plasma are markers of the origin of azoospermia and can predict the presence of sperm in testicular tissue. Hum Reprod. 2018, 33, 1087–1098. doi: 10.1093/humrep/dey072. PMID: 29635626.

Baud D., Pattaroni C., Vulliemoz N., Castella V., Marsland B.J., Stojanov M.: Sperm microbiota and its impact on semen parameters. Front Microbiol. 2019, 10, e234. doi: 10.3389/fmicb.2019.00234. PMID: 30809218.

Bojesen A., Juul S., Gravholt C.H.: Prenatal and postnatal prevalence of Klinefelter syndrome: a national registry study. J Clin Endocrinol Metab. 2003, 88, 622–626. doi: 10.1210/jc.2002-021491. PMID: 12574191.

Cao M., Shao X., Chan P., Cheung W., Kwan T., Pastinen T.B. et al.: High-resolution analyses of human sperm dynamic methylome reveal thousands of novel age-related epigenetic alterations. Clin Epigenetics. 2020, 12, 192. doi: 10.1186/s13148-020-00988-1. PMID: 3331763.

Cannarella R., Condorelli R.A., Duca Y., La Vignera S., Calogero A.E.: New insights into the genetics of spermatogenic failure: a review of the literature. Hum Genet. 2019, 138, 125–140. doi: 10.1007/s00439-019-01974-1. PMID: 30656449.

Chaitankar V., Karakülah G., Ratnapriya R., Giuste F.O., Brooks M.J., Swaroop A.: Next generation sequencing technology and genome wide data analysis: Perspectives for retinal research. Prog Retin Eye Res. 2016, 55, 1–31. doi: 10.1016/j.preteyeres.2016.06.001. PMID: 27297499.

Chen H., Luo T., Chen T., Wang G.: Seminal bacterial composition in patients with obstructive and non-obstructive azoospermia. Exp Ther Med. 2018, 15, 2884–2890. doi: 10.3892/etm.2018.5778. PMID: 29456693.

Cheung S., Parrella A., Rosenwaks Z., Palermo G.D.: Genetic and epigenetic profiling of the infertile male. PLoS ONE. 2019 14, e0214275. doi.org/10.1371/journal.pone.0214275. PMID: 30897172.

Cioppi F., Rosta V., Krausz C.: Genetics of azoospermia. Int J Mol Sci. 2021, 22, 3264. doi: 10.3390/ijms22063264. PMID: 33806855.

Choi M., Scholl U.I., Ji W., Liu T., Tikhonova I.R., Zumbo P. et al.: Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci USA. 2009, 106, 19096–19101. doi: 10.1073/pnas.0910672106. PMID: 19861545.

Cui X., Wu X., Li Q., Jing X.: Mutations of the cystic fibrosis transmembrane conductance regulator gene in males with congenital bilateral absence of the vas deferens: Reproductive implications and genetic counseling (Review). Mol Med Rep. 2020, 22, 3587–3596. doi: 10.3892/mmr.2020.11456. PMID: 33000223.

Fakhro K.A., Elbardisi H., Arafa M., Robay A., Rodriguez-Flores J.L., Al-Shakaki A. et al.: Point-of-care whole-exome sequencing of idiopathic male infertility. Genet Med. 2018, 20, 1365–1373. doi: 10.1038/gim.2018.10. PMID: 29790874.

Farahani L., Tharakan T., Yap T., Ramsay J.W., Jayasena C.N., Minhas S.: The semen microbiome and its impact on sperm function and male fertility: A systematic review and meta-analysis. Andrology. 2021, 9, 115–144. doi: 10.1111/andr.12886. PMID: 32794312.

Fiorentino F., Bono S., Biricik A., Nuccitelli A., Cotroneo E., Cottone G. et al.: Application of next-generation sequencing technology for comprehensive aneuploidy screening of blastocysts in clinical preimplantation genetic screening cycles. Hum Reprod. 2014, 29, 2802–2813. doi: 10.1093/humrep/deu277. PMID: 25336713.

García-Pascual C.M., Navarro-Sánchez L., Navarro R., Martínez L., Jiménez J., Rodrigo L. et al.: Optimized NGS Approach for Detection of Aneuploidies and Mosaicism in PGT-A and Imbalances in PGT-SR. Genes (Basel). 2020, 11(7), 724. doi: 10.3390/genes11070724. PMID: 32610655.

Ghieh F., Barbotin A.L., Leroy C., Marcelli F., Swierkowsky-Blanchard N., Serazin V. et al.: Will whole-genome sequencing become the first-line genetic analysis for male infertility in the near future? Basic Clin Androl. 2021, 31, 21. doi: 10.1186/s12610-021-00138-4. PMID: 34407766.

Gunes S., Esteves S.C.: Role of genetics and epigenetics in male infertility. Andrologia. 2021, 53, e13586. doi: 10.1111/and.13586. PMID: 32314821.

Harper J.C., Aittomäki K., Borry P., Cornel M.C., de Wert G., Dondorp W., on behalf of the European Society of Human Reproduction and Embryology and European Society of Human Genetics: Recent developments in genetics and medically assisted reproduction: from research to clinical applications. Eur J Hum Genet. 2018, 26, 12-33. doi: 10.1038/s41431-017-0016-z. PMID: 29199274.

Hu Z., Xia Y., Guo X., Dai J., Li H., Jiang Y. et al.: A genome-wide association study in Chinese men identifies three risk loci for non-obstructive azoospermia. Nat Genet. 2012, 44, 183-186. doi: 10.1038/ng.1040. PMID: 22197933.

Johnson M.D.: Genetic risks of intracytoplasmic sperm injection in the treatment of male infertility: recommendations for genetic counseling and screening. Fertil Steril. 1998, 70, 397. doi: 10.1016/s0015-0282(98)00209-x. PMID: 9757865.

Jungwirth A., Diemer T., Kopa Z., Krausz C., Minhas S., Tournaye H.: EAU Guidelines on Male Infertility. European Association of Urology. 2019, 1–48. [viewed: 20.12.202]. Available in: https://uroweb.org/wp-content/uploads/EAU-Guidelines-on-Male-Infertility-2019.pdf

Kotaja N.: MicroRNAs and spermatogenesis. Fertil Steril. 2014, 101, 1552–1562. doi: 10.1016/j.fertnstert.2014.04.025. PMID: 24882619.

Kothandaraman N., Agarwal A., Abu-Elmagd M., Al-Qahtani M.H.: Pathogenic landscape of idiopathic male infertility: new insight towards its regulatory networks. NPJ Genom Med. 2016, 1, 16023. doi: 10.1038/npjgenmed.2016.23. PMID: 29263816.

Krausz C., Escamilla A.R., Chianese C.: Genetics of male infertility: from research to clinic. Reproduction. 2015, 150, R159-174. doi: 10.1530/REP-15-0261. PMID: 26447148.

Krausz C., Riera-Escamilla A.: Genetics of male infertility. Nat Rev Urol. 2018, 15, 369-384. doi: 10.1038/s41585-018-0003-3. PMID: 29622783.

Krausz C., Hoefsloot L., Simoni M., Tüttelmann F., *European Academy of Andrology, European Molecular Genetics Quality Network*: EAA/EMQN best practice guidelines for molecular diagnosis of Y-chromosomal microdeletions: state-of-the-art 2013. Andrology. 2014, 2, 5–19. doi: 10.1111/j.2047-2927.2013.00173.x. PMID: 24357628.

Liang D., Wang Y., Ji X., Hu H., Zhang J., Meng L. et al.: Clinical application of whole-genome low-coverage next-generation sequencing to detect and characterize balanced chromosomal translocations. Clin Genet. 2017, 91, 605–610. doi: 10.1111/cge.12844. PMID: 27491356.

Li L., Li H., *Tian Y., Hu M., Le F., Wang L. et al.*: Differential microRNAs expression in seminal plasma of normospermic patients with different sperm DNA fragmentation indexes. Reprod Toxicol. 2020, 94, 8–12. doi: 10.1016/j.reprotox.2020.03.007. PMID: 32259568.

Liu X., Zhang H., Zhang X., Zhang H., Jiang Y., Liu R. et al.: Y chromosome structural variation in infertile men detected by targeted next-generation sequencing. J Assist Reprod Genet. 2021, 38, 941–948. doi: 10.1007/s10815-020-02031-x. PMID: 33454900.

Lu H., Xu D., Wang P., Sun W., Xue X., Hu Y. et al.: RNA-sequencing and bioinformatics analysis of long noncoding RNAs and mRNAs in the asthenozoospermia. Biosci Rep. 2020, 40, BSR20194041. doi: 10.1042/BSR20194041. PMID: 32614449.

Łukaszuk K., Pukszta S., Wells D., Cybulska C., Liss J., Płóciennik Ł., *et al.*: Routine use of next-generation sequencing for preimplantation genetic diagnosis of blastomeres obtained from embryos on day 3 in fresh in vitro fertilization cycles. Fertil Steril. 2015, 10, 1031–1036. doi: 10.1016/j.fertnstert.2014.12.123. PMID: 25624194.

National Human Genome Research Institute https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data, entry date: 20.12.2021

Oud M.S., Volozonoka L., Smits R.M., Vissers L., Ramos L., Veltman J.A.: A systematic review and standardized clinical validity assessment of male infertility genes. Hum Reprod. 2019, 34, 932–941. doi.org/10.1093/humrep/dez022. PMID: 30865283.

Pelzman D.L., Hwang K.: Genetic testing for men with infertility: techniques and indications. Transl Androl Urol. 2021, 10, 1354–1364. doi: 10.21037/tau-19-725. PMID: 33850771.

Practice Committee of the American Society for Reproductive Medicine: Diagnostic evaluation of the infertile male: a committee opinion. Fertil Steril. 2015, 103, e18–25. DOI: 10.1016/j.fertnstert.2014.12.103. PMID: 25597249.

Precone V., Cannarella R., Paolacci S., Busetto G.M., Beccari T., Stuppia L. et al.: Male infertility diagnosis: Improvement of genetic analysis performance by the introduction of pre-diagnostic genes in a Next-Generation Sequencing custom-made panel. Front Endocrinol. 2021, 11, 605237. doi: 10.3389/fendo.2020.605237. PMID: 33574797.

Ravel C., Berthaut I., Bresson J.L., Siffroi J.P.: Genetics Commission of the French Federation of CECOS. Prevalence of chromosomal abnormalities in phenotypically normal and fertile adult males: large-scale survey of over 10,000 sperm donor karyotypes. Hum Reprod. 2006, 21, 1484–1489. doi: 10.1093/humrep/del024. PMID: 16484311.

Reardon S.: A complete human genome sequence is close: how scientists filled in the gaps. Nature. 2021, 594, 158–159. doi: 10.1038/d41586-021-01506-w. PMID: 34089035.

Rocca M.S., Msaki A., Ghezzi M., Cosci I., Pilichou K., Celeghin R., et al.: Development of a novel next-generation sequencing panel for diagnosis of quantitative spermatogenic impairment. J Assist Reprod Genet. 2020, 37, 753–762. doi: 10.1007/s10815-020-01747-0. PMID: 32242295.

Rubio C., Rodrigo L., Garcia-Pascual C., Peinado V., Campos-Galindo I., Garcia-Herrero S, et al.: Clinical application of embryo aneuploidy testing by next-generation sequencing. Biol Reprod. 2019, 101, 1083–1090. doi: 10.1093/biolre/ioz019. PMID: 30721942.

Schlegel P.N., Sigman M., Collura B., De Jonge C.J., Eisenberg ML, Lamb DJ, et al.: Diagnosis and treatment of infertility in men: AUA/ASRM Guideline Part I. J Urol. 2021, 205, 36–43. PMID: 33295257. doi: 10.1097/JU.0000000000001521.

Tang D., Huang Z., He X., Wu H., Peng D., Zhang L., et al.: Altered miRNA profile in testis of post-cryptorchidopexy patients with non-obstructive azoospermia. Reprod Biol Endocrinol. 2018, 16, 78. doi: 10.1186/s12958-018-0393-3. PMID: 30103742.

Tsonis O., Gkrozou F., Paschopoulos M.: Microbiome affecting reproductive outcome in ARTs. J Gynecol Obstet Hum Reprod. 2021, 50, 102036. doi: 10.1016/j.jogoh.2020.102036. PMID: 33307241.

Tüttelmann F., Ruckert C., Röpke A.: Disorders of spermatogenesis. Medgen. 2018, 30, 12–20. doi.org/10.1007/s11825-018-0181-7. PMID: 29527098.

Van Assche E., Bonduelle M., Tournaye H., Joris H., Verheyen G., Devroey P., et al.: Cytogenetics of infertile men. Hum Reprod. 1996, 11, 1–26. doi: 10.1093/humrep/11.suppl_4.1. PMID: 9147109.

Van Dijk E.L., Auger H., Jaszczyszyn Y., Thermes C.: Ten years of next-generation sequencing technology. Trends Genet. 2014, 30. 418–426. doi: 10.1016/j.tig.2014.07.001. PMID: 25108476.

Wang Z., Gerstein M., Snyder M.: RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009, 10, 57–63. doi: 10.1038/nrg2484. PMID: 19015660.

Weng S.-L., Chiu C.-M., Lin F.-M., Huang, W.C., Liang, C., Yang, et al.: Bacterial communities in semen from men of infertile couples: metagenomic sequencing reveals relationships of seminal microbiota to semen quality. PLoS One. 2014, 9, e110152. doi.org/10.1371/journal.pone.0110152. PMID: 25340531.

Xu M., Qin Y., Qu J., Lu C., Wang Y., Wu W., et al.: Evaluation of five candidate genes from GWAS for association with oligozoospermia in a Han Chinese population. PloS one. 2013, 8, e80374. doi.org/10.1371/journal.pone.0080374. PMID: 24303009.