

## A Review of Gene Sequencing in Infertility

Jie ZHAO<sup>†</sup>, Ri-Na SHA<sup>†</sup>, Fang LIU, Chen DU, Bai DAI, Yan-Bing ZHANG, Ru-Gang TIAN, Jiang-Feng HE,  
and Xiu-Juan CHEN

Accepted July 31, 2020

Published online September 01, 2020

Issue online September 30, 2020

Review article

ZHAO J., SHA R.-N., LIU F., DU C., DAI B., ZHANG Y.-B., TIAN R.-G., HE J.-F., CHEN X.-J. 2020.  
A Review of gene sequencing in infertility. *Folia Biologica (Kraków)* 68: 97-105.

Assisted reproductive technology (ART) is the technology used primarily for infertility treatments in order to achieve pregnancy. With the increasing demands of preimplantation genetic testing (PGT), the safety inspection of in-vitro fertilization (IVF), the prevention of pathogenesis in sterility, the application of gene sequencing technology in ART, and the diagnosis of infertility-related diseases have significantly increased. The development and principles of gene sequencing technology, the application of gene sequencing in PGT, the application of whole genome sequencing in ART-conceived children, and the application of whole genome sequencing in infertility-related disorders have been summarized in this review.

Key words: infertility; gene sequencing; preimplantation genetic testing; in-vitro fertilization.

Jie ZHAO, Fang LIU, Chen DU, Bai DAI, Yan-Bing ZHANG, Xiu-Juan CHEN<sup>✉</sup>, *Reproductive Medicine Center, the Affiliated Hospital of Inner Mongolia Medical University, Hohhot, China.*

Ri-Na SHA, *Pathology department, the Affiliated Hospital of Inner Mongolia Medical University, Hohhot, China.*

Ru-Gang TIAN, Jiang-Feng HE<sup>✉</sup>, *Inner Mongolia Academy of Agricultural and Animal Husbandry Science, Hohhot, China.*

<sup>†</sup> These authors contributed equally to this work.

<sup>✉</sup> To whom correspondence should be addressed: 13704781032@163.com (Jiang-Feng He);  
90098687@sina.com (Xiu-Juan Chen)

Since the DNA double helix structure was discovered in 1953, people began to understand that DNA fragments were responsible for genetic effects and that the gene is able to control biological properties (CHIU *et al.* 2020). Genes can not only faithfully reproduce themselves and maintain biological characteristics, but they can also mutate, leading to the occurrence of diseases that may be inherited by future generations. Therefore, it is necessary for researchers of the life sciences to reveal the complexity and diversity of genomes by gene sequencing and by further access to or utilization of genetic information. As gene sequencing technology has been widely used in many disease mechanism studies and clinical diagnoses, it will have extensive applied values in biomedicine, gene therapy, and clinic practice (THENMOZHI *et al.* 2020).

Infertility is a global problem, affecting an estimated one in four couples in developing countries. The incidence of infertility has not changed significantly over the past two decades, although it is diffi-

cult to estimate because of the inconsistent definitions used and a lack of standardization in demographic surveys (MASCARENHAS *et al.* 2012; WILTSHIRE *et al.* 2020). Through traditional semen analysis and other classical diagnostic tests, approximately one-third of infertility cases are attributable to male factors, one-third to female factors, and the remaining third of couples are unable to determine the cause, also known as unexplained or idiopathic infertility (ISIDORI *et al.* 2006; THONNEAU *et al.* 1991). Some infertility cases of unknown causes may be due to an unknown disease in the men, women, or couples (i.e., men, women and women). Among them, the combination of male and female factors can lead to infertility in couples. So, in general, men are just as likely as women to be infertile (GNOTH *et al.* 2005; TURNER *et al.* 2020).

### Gene Sequencing Technology

Gene sequencing is the process of determining the precise order of nucleotides within individual genes.

sRNA from *Escherichia coli* alanine and a yeast was the first nucleic acid molecule to be sequenced in 1965 (SANGER *et al.* 1965). A critical transition technique (the ‘plus and minus’ method) was introduced for DNA sequencing in 1975, which completely dominated gene sequencing for over 30 years (MAXAM & GILBERT 1977). In the 1970’s, Sanger and Maxam-Gilbert developed the ‘base-specific chemical cleaving method’ (DEWEY *et al.* 2014) and the ‘enzymatic chain termination dideoxy method’ (BOYCOTT *et al.* 2013), which greatly improved the speed, efficiency and accuracy of DNA sequencing technology. The ‘dideoxy method’ and the ‘chemical cleaving method’ were considered to be ‘first-generation’ sequencing technology. It is notable that the completion of the Human Genome Project (HGP) was achieved as consequence of this so-called first-generation sequencing technology.

After the HGP, the application of nucleic acid sequencing has significantly increased in genomic function studies, with a resultant demand for highly efficient, rapid, and low-cost DNA sequencing platforms. Next-generation sequencing (NGS), or second-generation sequencing, is a generic term used to describe several high-throughput DNA sequencing technologies (GASSNER 2020). Also known as massive (or massively) parallel or deep sequencing, the technology can sequence DNA and RNA faster and more cheaply than the first generation of Sanger sequencing that was previously used, and even larger than the size of a genome. In the past decade, this development has revolutionized genomics and molecular biology (MARDIS 2011). The high demand for low cost sequence data has driven the development of NGS technologies that can produce 1000 or millions of sequences concurrently. NGS relies on massively parallel sequencing and imaging techniques to yield several 100’s of millions to several 100’s of billions of DNA bases per run (SHENDURE & JI 2008). Several NGS platforms, such as Roche 454 FLX Titanium (THUDI *et al.* 2012), Illumina Mi Seq and Hi Seq2500 (BENTLEY *et al.* 2008), and Ion Torrent PGM, have been developed and used recently (HE *et al.* 2014; QUAIL *et al.* 2012). High-throughput sequencing technologies are intended to lower the cost of DNA sequencing far beyond what is possible with the standard dye-terminator methods (SCHUSTER 2008). In ultra-high-throughput sequencing, as many as 500,000 sequencing-by-synthesis operations may be run in parallel (METZKER 2010a; QUAIL *et al.* 2012).

All NGS strategies follow a similar protocol for DNA template preparation, where universal adapters are ligated at both ends of randomly sheared DNA fragments. They also rely on the cyclic interrogation of millions of clonally amplified DNA molecules immobilized on a synthetic surface to generate up to several billions of sequences in a massively parallel fashion. Sequencing is performed in an iterative manner,

where the incorporation of one or more nucleotides is followed by the emission of a signal and its detection by the sequencer (METZKER 2010b). Most NGS platforms are able to generate reliable sequences and display near perfect coverages of GC-rich, neutral and moderate AT-rich genomes. However, there are key differences between the quality of sequence data and the applications it will support (QUAIL *et al.* 2012).

NGS technologies commercialized by Illumina generate shorter reads, ranging from 50 to 300 bp, with sequencing throughputs ranging from 1.5 to 600 G bp depending on the platform being used. Several instruments are commercialized by Illumina, ranging from the bench top Mi Seq sequencer to the high-throughput Hi Seq2500 sequencer. The Illumina sequencing technology combines clonal amplification of a single DNA molecule with a cyclical sequencing-by-synthesis approach. PCR amplification is performed using a solid phase amplification protocol to generate up to 1,000 copies of an original molecule of DNA, grouped together into a cluster. Sequencing is performed with proprietary reversible fluorescent terminator deoxyribonucleotides, in a series of cycles consisting of single base extension, fluorescence detection (where the nature of the signal is used to determine the identity of the base being incorporated) and cleavage of both the fluorescent label and of the chemical moieties at the 3 hydroxyl position to allow for the next cycle to occur (HE *et al.* 2014).

The application of NGS technologies highlights the striking impact of these massively parallel platforms on genotyping, which have expanded from previously focused readouts from a variety of DNA preparation protocols to a genome-wide scale and have fine-tuned their resolution to single base precision. This method, takes advantage of DNA polymerases to incorporate four different fluorescently labeled deoxyribonucleotide triphosphates (dNTPs) into a DNA template strand during DNA synthesis. The nucleotides are identified by fluorophore excitation at the point of incorporation. The DNA sequence is then analyzed by proprietary software on the sequencing platform. NGS sequencing platforms are typically represented by GSFLX Titanium from Roche, Solexa Genome Analyzer from Illumina, and SOLiD/Ion Torrent PGM from Life Sciences. Sequencing through synthesis is the most widely adopted technology. The principle of this approach is that a fluorescently labeled, reversible terminator is imaged as each dNTP is added, and then cleaved to allow the incorporation of the next base, which can allow the millions of DNA sequences to be read, thereby making a deep sequencing of transcriptome at a low cost and at high speed achievable (WANG *et al.* 2014).

NGS technologies have been under continuous development and improvement with the onset of the 21st century. By eliminating the major NGS drawback-short read length, the revolution is continu-

ing as scientists move into the era of single molecule sequencing. This is often referred to as ‘third-generation’ sequencing. Several platforms have been developed to provide low-cost and high-throughput RNA sequencing, such as true single molecular sequencing (tSMS), single molecule real-time (SMRT), fluorescence resonance energy transfer (FRET), DNA strand transits through nanopores, and microscopy-based techniques.

With the advance of gene sequencing technologies, the application of whole genome sequencing has been widely used in many fields of biology, including genetic studies of diseases (MARTIN *et al.* 2013). Over 4000 genes responsible for rare monogenic diseases have been identified (PALINI *et al.* 2013). Whole exome sequencing provides new approaches to characterize individual genomic landscapes and identify

disease-related information, such as simple nucleotide variations (SNVs), insertions and deletions (INDELS), copy number variation (CNV), structural variation (SV) (ZONG *et al.* 2012), RNA expression profiling difference (LU *et al.* 2012), and the base 5-hydroxymethylcytosine (HOU *et al.* 2013). Ultimately this will facilitate the discovery of rare-disease-causing genes and increase therapeutic opportunities.

#### The Application of Gene Sequencing in Preimplantation genetic testing (PGT)

PGT is a test performed to analyze the DNA from oocytes (polar bodies) or embryos (cleavage stage or blastocyst) for HLA-typing or for determining genetic abnormalities (Figure 1). These include: PGT for aneuploidies (PGT-A); PGT for monogenic/single gene

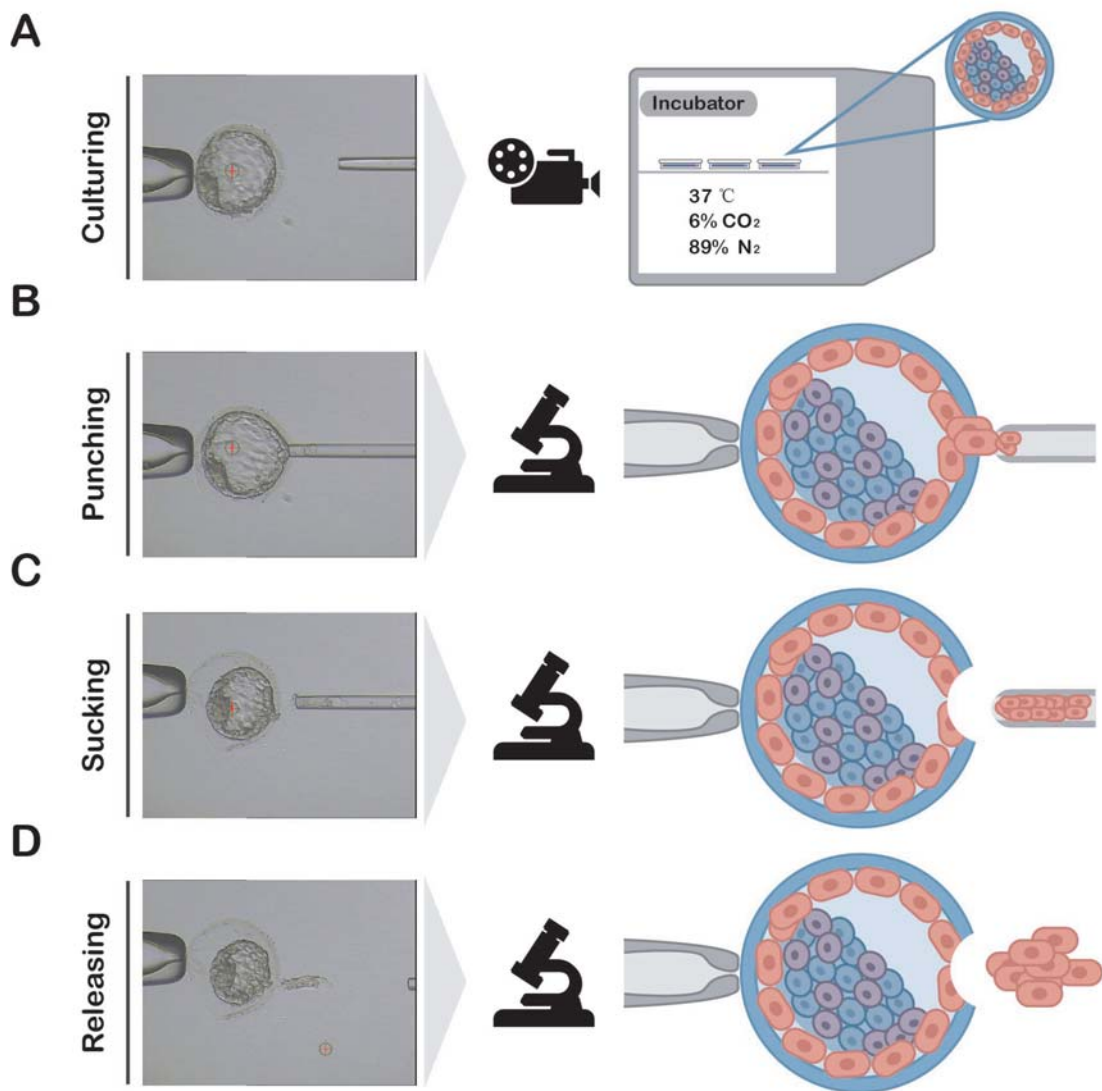


Fig. 1. The most commonly used method for blastocyst biopsy. (A) Embryos are cultured to the blastocyst stage in a standard incubator. (B) The zona pellucida is drilled on the opposite side of the Inner Cells Mass. (C) A small number of cells are sucked into a pipette and then removed from the blastocyst by alternating the stretching and firing a laser. (D) The biopsy fragment is released.

defects (PGT-M); and PGT for chromosomal structural rearrangements (PGT-SR) (ZEGERS-HOCHSCHILD *et al.* 2017). As an assisted reproductive technology (ART), PGT was applied and resulted in the success of the world's first PGT test-tube baby in 1990 and it has become an important technology in clinical applications (GRECO *et al.* 2020; LI *et al.* 2015).

NGS is a rapidly developing approach applied for PGT. With increasing applications in clinical practices, this new approach offers the process of removing a single-cell from an in vitro fertilization embryo for genetic testing. In comparison with fluorescence in situ hybridization and DNA chip testing, the NGS technique allows for the detection of an embryonic chromosomal abnormality rapidly with higher accuracy, higher resolution, more sensitivity, and efficiency (BOYCOTT *et al.* 2013). It can detect copy number variations more than 1 Mb at the single cell level (LI *et al.* 2015) and can identify individual single-nucleotide variations (SNVs) with no false positives detected (ZHANG *et al.* 2015). Furthermore, NGS could potentially support more 'personalized' procedures in patients at risk of genetic disease (PETERS *et al.* 2015).

The success of gene sequencing requires a large population of cells, the gene expression information derived from clinical studies reflects the expression from the population level, i.e. the majority of cell types among millions of cells are typically analyzed in bulk. However, the cell phenotypic difference is due to cell heterogeneity (KONG *et al.* 2012) and gene expression noise (SUZUKI & BIRD 2008). To understand gene expression from a single cell for PGT, genomic sequencing must be done on a single cell.

Some common resources used for PGT include blastomere cells, trophoctoderm cells, first and second-polar bodies, and blastocoel fluid (ZHAO *et al.* 2013). The quantity of genomic DNA present in an individual cell is insufficient to directly sequence; therefore, genome-wide amplification (WGA) has been used as the first step in single-cell DNA amplification. However, allele dropout (ADO) and the amplification bias (PA) phenomenon still restrict low genome coverage of WGA-based PGT at 40%. To address this limitation, WGA efficiency must be increased. A new technique has been developed to obtain extensive coverage and uniform amplification of genomes (SUZUKI & BIRD 2008). This method, termed multiple annealing and looping-based amplification cycles (MALBAC), leads to 93% genome coverage at an average 253 sequencing depth for DNA sequencing from a single cell. This method was also used to sequence sperm to phase the personal genome and map recombination events at a high resolution (HALLIDAY *et al.* 2004). With this method, the genome of the oocyte pronucleus can be deduced by sequencing the triads of the first and second polar bodies (HANSEN *et al.* 2013), thereby avoiding mater-

nal genetic congenital birth defects in newborns. MALBAC-based preimplantation genomic screening makes it possible to accurately select normal fertilized eggs for embryo transfer (FENG *et al.* 2008; KITZMAN *et al.* 2012; ZHENG *et al.* 2013), therefore, the sequencing of MALBAC-amplified material from a single cell increases the number of healthy, live babies born (EHRICH 2011). In addition, NGS could potentially support more 'personalized' procedures in patients at risk of genetic disease (PETERS *et al.* 2015).

*De novo* mutations, occurring in a gamete (egg or sperm) of one of the parents or in a developing fetus, include single nucleotide polymorphisms (SNPs) and short multibase insertion/deletions (INDELS). This will cause a number of genetic diseases, such as Down syndrome, severe intellectual disability, autism, epileptic encephalopathies, and many other congenital disorders. However, due to the mutation being absent in the blood DNA of the unaffected parents, the screening of both parents could fail to detect such a mutation. By using gene sequencing, the preimplantation genomic screening of an embryo prior to being transferred to the uterus allows for the detection of *de novo* mutations and methylation pattern. PALOMAKI *et al.* (2012) developed a new process using advanced WGA to detect single base *de novo* mutations from IVF blastocyst biopsies, providing highly sensitive and specific screening. In addition, WGA can also be used to understand the profile of DNA methylation (QIN *et al.* 2007). This method facilitates the evaluation of ART outcomes and improves clinical diagnoses.

WGA not only offers great potential in analyzing preimplantation embryos, but also poses substantial challenges regarding its clinical utilization, such as the difficulty with examining balanced translocations and expansion CGG(n) testing (e.g. fragile X-mental retardation-1 gene). Therefore, a better reference genome and the combination of traditional PGT techniques are needed to improve its diagnostic accuracy. In addition, the limitation in resolution of chromosome level PGT with DNA PGT needs to be overcome in the future. The possibility of performing WGA and NGS for PGT is just around the corner for IVF. With the breakthrough of sequencing technology and accurate sequencing data analysis, it will be possible to increase the chances of a successful pregnancy and live baby birth rate, as well as to reduce the miscarriage rate due to aneuploidy embryos and to avoid congenital birth defects.

#### The Application of Gene Sequencing in the safety assessment of ART progeny

In mammals, gametogenesis and early preimplantation development are two critical periods of epigenetic modification. A wave of global DNA demethylation

and remethylation during these two periods, which have an important role in fetoplacental development, are required for the expression of imprinted genes (WANG *et al.* 2012). ART-related manipulations to oocytes and embryos involve several processes, such as follicular stimulation, intracytoplasmic sperm injection (ICSI), and embryo culture. These ART procedures might predispose embryos to acquire imprinting errors and diseases, resulting from coincident timing between gametogenesis, embryo development periods, and the time windows of ART (WU *et al.* 2005). Several studies have shown that ART treatment increases the risk of live baby birth defects (HANSEN *et al.* 2013). ART might result in abnormal methylation, loss of methylation, aneuploidy, structural rearrangements of chromosomes, dynamic mutation of trinucleotide repeats (ZHENG *et al.* 2013), and X or Y-chromosome microdeletion (FENG *et al.* 2008). In addition, ICSI offspring are at higher risk of aneuploidy and structural rearrangements of the chromosomes than control groups and the incidence of Y-chromosome micro deletion in the AZF gene increased in ART children. Whole genome sequencing (WGS) diagnoses can detect these genetic diseases by examining a small quantity of embryo cells before a baby is born.

Chromosome abnormality is an inherited genetic disease and it is one of leading cause of congenital birth defects. However, due to the lack of a better therapy, the pregnancy is often terminated according to diagnostic testing. Currently amniocentesis and umbilical vein sampling are two main accurate and useful diagnostic antenatal tests for the diagnosis of chromosomal abnormalities. These are fairly easy procedures and are less traumatic to the fetus. The disadvantage of these two tests is that they are invasive and carry a small risk of miscarriage. Hence it is important that these tests are performed only in certain situations with an accurate diagnostic result. High-throughput sequencing in combination with amniocentesis and umbilical vein sampling from fetal cell DNA in maternal blood plasma has been widely used in the identification of trisomy 21, trisomy 18, trisomy 13, and Turner's syndrome.

There are two fetal genetic materials present in maternal circulation, including intact fetal cells and cell-free fetal DNA (cff-DNA). All fetal genomes are present in maternal blood in the form of cff-DNA (ALBERRY *et al.* 2007). They consist of short DNA fragments in maternal circulation, originating from apoptotic placental cells (trophoblast cells) in the embryo (LUN *et al.* 2008). In addition, cff-DNA can be detected within 4 weeks of gestation and will be gone from maternal circulation within 2 h after delivery (SUZUMORI *et al.* 2018). Therefore, cff-DNA in maternal plasma has been used for fetal aneuploidy, NIPT of paternal hereditary disease, and fetal X-chain disease sex identification. The presence of fetal cells

is a relative rarity in maternal blood, estimated to be at only one to two per ml in the first-trimester pregnancy, whereas cff-DNA is more abundant in maternal plasma in early pregnancy. Therefore, analysis of cff-DNA provides a method for a non-invasive prenatal diagnosis (NIPD). A variety of methods have been used for mutation detection with cff-DNA. As a up-to-date technique, gene sequencing is trended to be applied in clinical studies (PAPAGEORGIOU & PATSALIS 2013). It approaches high sensitivity and accuracy to perform prenatal screening for trisomy 21 (EHRICH 2011), trisomy 16, and trisomy 18 (PALOMAKI *et al.* 2012) using the NGS of cell-free DNA found in maternal circulation. Furthermore, a complete set of the genome sequences of a human fetus at 18.5 weeks of gestation promoted the analysis of monogenic diseases using noninvasive prenatal genetic diagnostics (KITZMAN *et al.* 2012).

A safe treatment is critical to implement the diagnosis of genetic diseases, ART can provide an effective treatment to evaluate and ensure its safety by sequencing. With technology advancing, the information of small gene mutation, including chromosomal aneuploidy of ART fetal, mark genetic changes, and small deletions in tiny genetic mutations, can be obtained from cff-DNA or fetal cells in the blood of pregnant women. Consequently, such information allows for a better assessment of the safety of ART offspring and a decrease of the risk of birth defects.

#### The Application of Gene Sequencing in Infertility-Related Disorders

According to the investigation and statistics of the World Health Organization (WHO), infertility is mainly caused by premature ovarian failure (POF), polycystic ovarian syndrome (PCOS), endometriosis (EMS), and spermatogenesis. To test for genetic disorders, DNA sequencing is one of the most reliable, accurate methods to identify possible disease-causing mutations. NGS has been revolutionizing genetic research and diagnostics, especially in the area of human fertility problems. Advances in NGS have led to a better understanding of reproductive genetics by investigating infertility-related genes, such as the genes involved in premature ovarian failure, polycystic ovarian syndrome, endometriosis, and spermatogenesis.

POF was defined as the cessation of ovarian function before the age of 40 and is associated with gonadotropin serum (FSH) & GT; 40mIU/ml (GUNNING *et al.* 2019). Studies have reported that incidences of POF in women before they are 20 years old is 0.01%, before they are 30 years old is 0.1% (KOKCU 2010), and before they are 40 years old is 1%-2% (CABURET *et al.* 2014). The average age of onset is 23.3 years old (PERSANI *et al.* 2010). POF reduces the fertility of women of childbearing age. In addition, it has serious

health consequences, including an increase in autoimmune diseases, osteoporosis, infertility, ischemic heart disease, and cardiovascular disease. Genetic studies through gene sequencing indicated that a variety of genes are related to premature ovarian failure, including the Nanos Homolog 3 gene (NANOS3) (QIN *et al.* 2007), B-cell lymphoma-2 gene (Bcl-2), Human bone morphogenetic protein 15 gene (BMP-15), and growth differentiation factor 9 (GDF-9) (LAISSUE *et al.* 2006). NANOS3 is a functional gene that makes primordial germ cells (PGCs) migrate to the gonad and perform their reproductive function. NANOS3 can express in oocytes at different follicular development phases and consequently, takes an important role in the formation of germ cells. Bcl-2 is a kind of apoptosis gene that can even speed up the apoptosis and atresia of follicles if its expression is down-regulated (WANG *et al.* 2012). BMP-15 is only expressed in oocytes and the oocyte-specific BMP-15 may promote follicle growth *in vivo*, but prevent the premature luteinization. If the BMP-15 gene is down-regulated, follicles can singularly mature and consequently, this will lead to ovulation failure. The GDF-9 gene can normally regulate the expression of key enzymes which warrant the normal development of the cumulus-oocyte complex. In an experiment on mice, follicular development was restrained and could cease at the primary stage if GDF-9 genes were knocked out. In addition, many genes, such as the transforming growth factor III receptor gene (TGFB3), fox head box L2 (FOXL2), newborn ovary homeobox gene (NOBOX), factor in the germ line alpha (FIGLA), sal like 4 (SALL4), diaphanous homolog 2 (DIAPH2), natriuretic peptide C (NPPC), (zinc finger X) ZFX, the X-inactivation gene (XISF), and the FSH receptor gene, are related to premature ovarian failure and distributed on the autosome and X chromosome (QIN *et al.* 2007).

PCOS, which is characterized by irregular menstruation, infertility, hirsutism, and polycystic ovary morphology (PCOM), is a common disease which affects female fertility, and is prone to increase the risk of complications, such as diabetes, cardiovascular disease, and endometrial cancer, in the long term (BARRY *et al.* 2014). Genetic studies have shown that PCOS is caused by a multigene mutation accompanied by abnormalities of the endocrine and metabolic systems. These multigenes mainly include the aldosterone synthase gene, CYP11A gene, a variable number of tandem repeats (VNTR), the MCF2L2 gene, melatonin receptor gene 1B (MTNR1B), adiponectin gene, gonadotropin releasing hormone receptor (GnRH-R) gene, serum tumor necrosis factor alpha (TNF- $\alpha$ ), and insulin like growth factor 1 (IGF-1). These genes are closely related with the PCOS pathophysiological mechanism and mainly concentrate on insulin, sex hormones, and the type II diabetes mellitus (T2DM)-related pathway, providing a new direction for PCOS genetic and genomic research.

EMS affects 10-15% of women of childbearing age. Endometriosis may occur in 17-44% of all patients treated with assisted reproductive technology (ART) (REDWINE 1999). The likely impact of endometriosis on ART outcomes remains the opposite. The available data on endometria-mediated ovarian injuries are conflicting. A gene association study of EMS indicated that several regional mutations of chromosomes were correlated with morbidity and that the abnormal expressed gene is involved in cellular localization, ion channel, and material transport. The periodic expression of a few homeobox genes (HOX) in the endometrium has been shown to be important for implantation and decidualization. HOXA10 gene expression in the endometrium is significantly decreased in patients with EMS infertility (WU *et al.* 2005).

Studies of the correlation between sperm quality and spermatogenesis genes in infertile males indicated that abnormal epigenetic modifications, Y chromosome microdeletion, and genetic variation in the autosomes can affect sperm quality and spermatogenesis. Normal epigenetic modifications of sperm can ensure the normal formation of the sperm nucleus and trigger normal, embryonic development. However, abnormal epigenetic modifications of sperm may result in early miscarriage, a descendant's phenotype defect, and clinical illnesses after birth. Some researchers reported that the expression of transcription factor (Ets variant 5) and the methylation transferase gene (ZAMUDIO *et al.* 2011) were defective and consequently, led to a spermatogenesis disorder and even to infertility, and the decrease of methylation levels of immune-related genes and sperm imprinting genes. The increase of the promoter activity of the perm DAZ (deleted in azoospermia) gene can lead to oligospermia. Low methylation changes on CpG loci result in weak sperm; whereas the low methylation of the maternal imprinted gene (MEST) can lead to a change in sperm morphology and dynamic (ZHENG *et al.* 2013). The Follicle stimulating hormone (FSH)  $\beta$  subunit and the follicle stimulating hormone receptor (FSHR) play an important role in regulating spermatogenesis. FSH and FSHR can be used as genetic markers associated with infertility treatment and consequently, they can provide a basis for hormone therapy and the optimization the infertility treatment program.

At the genetic level, diseases related to infertility involve chromosome diseases, single gene diseases, point mutations of related genes, epigenetics, etc., therefore, gene sequencing technology is considered to be a very effective way to reach an infertility diagnosis. In addition, many diseases are the result of interactions between multiple genes, gene sequencing technology can help us to discover differentially expressed genes and the genes related to etiology. This would be very helpful in the study of the infertility mechanism and to design a treatment plan.

## Conclusions

As our scientific knowledge and experience accumulate, the ART success rate has significantly increased and the potential applications of this technology have widely expanded. With the increasing prevalence of ART pregnancies, there have also been many concerns, which have gradually been overcome through the improvement of ART over more than 40 years. Tens of thousands of infants were born via ART around the world a number that will continually increase in the future. Although ART can help plenty of couples to solve their fertility issues, some unsuccessful cases resulted from individual differences and offspring safety assessments that are still an issue. These challenges will be settled with the rapid development of new technology in the postgenomic age.

The gene sequencing technique has not only been applied to the whole genome in a single cell, but also to the transcriptome. With NGS, the dynamic change of the transcriptome of human oocytes and early embryos, as well as the differential expression of allelic genes can be discovered in different developmental phases. The basic process and theory of the cell cycle, gene regulation, and metabolic pathways can be clarified and the gene promoter region in early embryonic development will one day be understood. Gene sequencing technology has had a significant impact on the study of the developmental mechanism of embryos, the diagnosis of human fetal sex diseases, and the prevention and treatment of infertility.

## Acknowledgments

For this review we received financial support from the Inner Mongolia Natural Fund (2018LH03006, 2018MS03006, 2018LH08031); Inner Mongolia Education Department Fund (NJZY16118); Inner Mongolia Innovation Fund (2017CXJJM08).

## Author Contributions

Research concept and design: X.-J.C.; Collection and/or assembly of data: J.Z.<sup>†</sup>; Data analysis and interpretation: C.D., B.D.; Writing the article: R.-N.S.<sup>†</sup>, J.-F.H.; Critical revision of the article: F.L., Y.-B.Z., R.-G.T.

<sup>†</sup> – These authors contributed equally to this work.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- ALBERRY M., MADDOCKS D., JONES M., ABDEL HADI M., ABDEL-FATTAH S., AVENT N., SOOTHILL P.W. 2007. Free fetal DNA in maternal plasma in anembryonic pregnancies: Confirmation that the origin is the trophoblast. *Prenat. Diagn.* **27**: 415-418. <https://doi.org/10.1002/pd.1700>
- BARRY J.A., AZIZIA M.M., HARDIMAN P.J. 2014. Risk of endometrial, ovarian and breast cancer in women with polycystic ovary syndrome: A systematic review and meta-analysis. *Hum. Reprod. Update* **20**: 748-758. <https://doi.org/10.1093/humupd/dmu012>
- BENTLEY D.R., BALASUBRAMANIAN S., SWERDLOW H.P., SMITH G.P., MILTON J., BROWN C.G., HALL K.P., EVERS D.J., BARNES C.L., BIGNELL H.R. et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**: 53-59. <https://doi.org/10.1038/nature07517>
- BOYCOTT K.M., VANSTONE M.R., BULMAN D.E., MACKENZIE A.E. 2013. Rare-disease genetics in the era of next-generation sequencing: Discovery to translation. *Nat. Rev. Genet.* **14**: 681-691. <https://doi.org/10.1038/nrg3555>
- CABURET S., ARBOLEDA V.A., LLANO E., OVERBEEK P.A., BARBERO J.L., OKA K., HARRISON W., VAIMAN D., BEN-NERIAH Z., GARCIA-TUNON I., FELLOUS M., PENDÁS A.M., VEITIA R.A., VILAIN E. 2014. Mutant cohesin in premature ovarian failure. *N. Engl. J. Med.* **370**: 943-949. <https://www.nejm.org/doi/10.1056/NEJMoa1309635>
- CHIU F.P., DOOLAN B.J., MCGRATH J.A., ONOUFRIDIS A. 2020. A decade of next-generation sequencing in genodermatoses: The impact on gene discovery and clinical diagnostics. *Br. J. Dermatol.* (online pre-print) <https://doi.org/10.1111/bjd.19384>
- DEWEY F.E., GROVE M.E., PAN C., GOLDSTEIN B.A., BERNSTEIN J.A., CHAIB H., MERKER J.D., GOLDFEDER R.L., ENNS G.M., DAVID S.P., PAKDAMAN N., ORMOND K.E., CALESHU C., KINGHAM K., KLEIN T.E., WHIRL-CARRILLO M., SAKAMOTO K., WHEELER M.T., BUTTE A.J., FORD J.M., BOXER L., IOANNIDIS J.P., YEUNG A.C., ALTMAN R.B., ASSIMES T.L., SNYDER M., ASHLEY E.A., QUERTERMOUS T. 2014. Clinical interpretation and implications of whole-genome sequencing. *JAMA* **311**: 1035-1045. <https://doi.org/10.1001/jama.2014.1717>
- EHRICH M., DECIU C., ZWIEFELHOFER T., TYNAN J.A., CAGASAN L., TIM R., LU V., MCCULLOUGH R., MCCARTHY E., NYGREN A.O., DEAN J., TANG L., HUTCHISON D., LU T., WANG H., ANGKACHATCHAI V., OETH P., CANTOR C.R., BOMBARD A., VAN DEN BOOM D. 2011. Noninvasive detection of fetal trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting. *Am. J. Obstet. Gynecol.* **204**: 205.e1-11. <https://doi.org/10.1016/j.ajog.2010.12.060>
- FENG C., WANG L.Q., DONG M.Y., HUANG H.F. 2008. Assisted reproductive technology may increase clinical mutation detection in male offspring. *Fertil. Steril.* **90**: 92-96. <https://doi.org/10.1016/j.fertnstert.2007.06.004>
- GASSNER C. 2020. Next-generation sequencing in blood group genomics: State of the art and perspectives. *Transfus. Med. Hemother.* **47**: 2-3. <https://doi.org/10.1159/000505463>
- GNOTH C., GODEHARDT E., FRANK-HERRMANN P., FRIOL K., TIGGES J., FREUNDL G. 2005. Definition and prevalence of subfertility and infertility. *Hum. Reprod.* **20**: 1144-1147. <https://doi.org/10.1093/humrep/deh870>
- GRECO E., LITWICKA K., MINASI M.G., CURSIO E., GRECO P.F., BARILLARI P. 2020. Preimplantation genetic testing: Where we are today. *Int. J. Mol. Sci.* **21**: 4381. <https://doi.org/10.3390/ijms21124381>
- GUNNING M.N., MEUN C., VAN RIJN B.B., MAAS A., BENSCHOP L., FRANX A., BOERSMA E., BUDDE R.P.J., APPELMAN Y., LAMBALK C.B., EIJKEMANS M.J.C., VELTHUIS B.K., LAVEN J.S.E., FAUSER B.C.J.M. 2019. Coronary artery calcification in middle-aged women with premature ovarian insufficiency. *Clin. Endocrinol. (Oxf)* **91**: 314-322. <https://doi.org/10.1111/cen.14003>

- HALLIDAY J., OKE K., BREHENY S., ALGAR E., J. AMOR D. 2004. Beckwith-wiedemann syndrome and ivf: A case-control study. *Am. J. Hum. Genet.* **75**: 526-528. <https://doi.org/10.1086/423902>
- HANSEN M., KURINCZUK J.J., MILNE E., DE KLERK N., BOWER C. 2013. Assisted reproductive technology and birth defects: a systematic review and meta-analysis. *Hum. Reprod. Update*: **19**: 330-353. <https://doi.org/10.1093/humupd/dmt006>
- HE J., ZHAO X., LAROCHE A., LU Z.X., LIU H., LI Z. 2014. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front. Plant. Sci.* **5**: 484. <https://doi.org/10.3389/fpls.2014.00484>
- HOU Y., FAN W., YAN L., LI R., LIAN Y., HUANG J., LI J., XU L., TANG F., XIE X.S., QIAO J. 2013. Genome analyses of single human oocytes. *Cell* **155**: 1492-1506. <http://dx.doi.org/10.1016/j.cell.2013.11.040>
- ISIDORI A.M., POZZA C., GIANFRILLI D., ISIDORI A. 2006. Medical treatment to improve sperm quality. *Reprod. Biomed. Online* **12**: 704-714. [https://doi.org/10.1016/S1472-6483\(10\)61082-6](https://doi.org/10.1016/S1472-6483(10)61082-6)
- KITZMAN J.O., SNYDER M.W., VENTURA M., LEWIS A.P., QIU R., SIMMONS L.E., GAMMILL H.S., RUBENS C.E., SANTILLAN D.A., MURRAY J.C., TABOR H.K., BAMSHAD M.J., EICHLER E.E., SHENDURE J. 2012. Noninvasive whole-genome sequencing of a human fetus. *Sci. Transl. Med.* **4**: 137-176. <https://doi.org/10.1126/scitranslmed.3004323>
- KOKCU A. 2010. Premature ovarian failure from current perspective. *Gynecol. Endocrinol.* **26**: 555-562. <https://doi.org/10.3109/09513590.2010.488773>
- KONG A., FRIGGE M.L., MASSON G., BESENBACHER S., SULEM P., MAGNUSON G., GUDJONSSON S.A., SIGURDSSON A., JONASDOTTIR A., JONASDOTTIR A., WONG W.S., SIGURDSSON G., WALTERS G.B., STEINBERG S., HELGASON H., THORLEIFSSON G., GUDBJARTSSON D.F., HELGASON A., MAGNUSON O.T., THORSTEINSDOTTIR U., STEFANSSON K. 2012. Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* **488**: 471-475. <https://doi.org/10.1038/nature11396>
- LAISSUE P., CHRISTIN-MAITRE S., TOURAINE P., KUTTENN F., RITVOS O., AITTO MAKI K., BOURCIGAUX N., JACQUESSON L., BOUCHARD P., FRYDMAN R., DEWAILLY D., REYSS A.C., JEFFERY L., BACHELOT A., MASSIN N., FELLOUS M., VEITIA R.A. 2006. Mutations and sequence variants in *GDF9* and *BMP15* in patients with premature ovarian failure. *Eur. J. Endocrinol.* **154**: 739-744. <https://doi.org/10.1530/eje.1.02135>
- LI N., WANG L., WANG H., MA M., WANG X., LI Y., ZHANG W., ZHANG J., CRAM D.S., YAO Y. 2015. The performance of whole genome amplification methods and next-generation sequencing for pre-implantation genetic diagnosis of chromosomal abnormalities. *J. Genet. Genomics* **42**: 151-159. <https://doi.org/10.1016/j.jgg.2015.03.001>
- LU S., ZONG C., FAN W., YANG M., LI J., CHAPMAN A.R., ZHU P., HU X., XU L., YAN L., BAI F., QIAO J., TANG F., LI R., XIE X.S. 2012. Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science* **338**: 1627-1630. <https://doi.org/10.1126/science.1229112>
- LUN F.M., CHIU R.W., CHAN K.C., LEUNG T.Y., LAU T.K., LO Y.M.D. 2008. Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin. Chem.* **54**: 1664-1672. <https://doi.org/10.1373/clinchem.2008.111385>
- MARDIS E.R. 2011. A decade's perspective on DNA sequencing technology. *Nature* **470**: 198-203. <https://doi.org/10.1038/nature09796>
- MARTIN J., CERVERO A., MIR P., MARTINEZ-CONEJERO J.A., PELLICER A., SIMON C. 2013. The impact of next-generation sequencing technology on preimplantation genetic diagnosis and screening. *Fertil. Steril.* **99**: 1054-1061.E3. <https://doi.org/10.1016/j.fertnstert.2013.02.001>
- MASCARENHAS M.N., FLAXMAN S.R., BOERMA T., VANDERPOEL S., STEVENS G.A. 2012. National, regional, and global trends in infertility prevalence since 1990: A systematic analysis of 277 health surveys. *PLoS Medicine* **9**: e1001356. <https://doi.org/10.1371/journal.pmed.1001356>
- MAXAM A.M., GILBERT W. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**: 560-564. <https://doi.org/10.1073/pnas.74.2.560>
- METZKER M.L. 2010. Sequencing technologies – the next generation. *Nat. Rev. Genet.* **11**: 31-46. <https://doi.org/10.1038/nrg2626>
- PALINI S., GALLUZZI L., DE STEFANI S., BIANCHI M., WELLS D., MAGNANI M., BULLETTI C. 2013. Genomic DNA in human blastocoele fluid. *Reprod. Biomed. Online* **26**: 603-610. <https://doi.org/10.1016/j.rbmo.2013.02.012>
- PALOMAKI G.E., DECIU C., KLOZA E.M., LAMBERT-MESSERLIAN G.M., HADDOW J.E., NEVEUX L.M., EHRLICH M., VAN DEN BOOM D., BOMBARD A.T., GRODY W.W., NELSON S.F., CANICK J.A. 2012. DNA sequencing of maternal plasma reliably identifies trisomy 18 and trisomy 13 as well as Down syndrome: An international collaborative study. *Genet. Med.* **14**: 296-305. <https://doi.org/10.1038/gim.2011.73>
- PAPAGEORGIOU E.A., PATSALIS P.C. 2013. Maternal plasma sequencing: a powerful tool towards fetal whole genome recovery. *BMC Medicine* **11**: 56. <https://doi.org/10.1186/1741-7015-11-56>
- PERSANI L., ROSSETTI R., CACCIATORE C. 2010. Genes involved in human premature ovarian failure. *J. Mol. Endocrinol.* **45**: 257-279. <https://doi.org/10.1677/JME-10-0070>
- PETERS B.A., KERMANI B.G., ALFEROV O., AGARWAL M.R., MCELWAIN M.A., GULBAHCE N., HAYDEN D.M., TANG Y.T., ZHANG R.Y., TEARLE R., CRAIN B., PRATES R., BERKELEY A., MUNNÉ S., DRMANAC R. 2015. Detection and phasing of single base *de novo* mutations in biopsies from human in vitro fertilized embryos by advanced whole-genome sequencing. *Genome Res.* **25**: 426-434. <http://www.genome.org/cgi/doi/10.1101/gr.181255.114>
- QIN Y., ZHAO H., KOVANCI E., SIMPSON J.L., CHEN Z.J., RAJKOVIC A. 2007. Mutation analysis of NANOS3 in 80 Chinese and 88 Caucasian women with premature ovarian failure. *Fertil. Steril.* **88**: 1465-1467. <https://doi.org/10.1016/j.fertnstert.2007.01.020>
- QUAIL M.A., SMITH M., COUPLAND P., OTTO T.D., HARRIS S.R., CONNOR T.R., BERTONI A., SWERDLOW H.P., GU Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina Miseq sequencers. *BMC Genomics* **13**: 341. <https://doi.org/10.1186/1471-2164-13-341>
- REDWINE D.B. 1999. Ovarian endometriosis: a marker for more extensive pelvic and intestinal disease. *Fertil. Steril.* **72**: 310-315. [https://doi.org/10.1016/S0015-0282\(99\)00211-3](https://doi.org/10.1016/S0015-0282(99)00211-3)
- SANGER F., BROWNLEE G.G., BARRELL B.G. 1965. A two-dimensional fractionation procedure for radioactive nucleotides. *J. Mol. Biol.* **13**: 373-398. [https://doi.org/10.1016/S0022-2836\(65\)80104-8](https://doi.org/10.1016/S0022-2836(65)80104-8)
- SCHUSTER S.C. 2008. Next-generation sequencing transforms today's biology. *Nat. Methods* **5**: 16-18. <https://doi.org/10.1038/nmeth1156>
- SHENDURE J., JI H. 2008. Next-generation DNA sequencing. *Nat. Biotechnol.* **26**: 1135-1145. <https://doi.org/10.1038/nbt1486>
- SUZUKI M.M., BIRD A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**: 465-476. <https://doi.org/10.1038/nrg2341>
- SUZUMORI N., SEKIZAWA A., EBARA T., SAMURA O., SASAKI A., AKAISHI R., WADA S., HAMANOUE H., HIRAHARA F., IZUMI H. et al. 2018. Fetal cell-free DNA fraction in maternal plasma for the prediction of hypertensive disorders of pregnancy. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **224**: 165-169. <https://doi.org/10.1016/j.ejogrb.2018.03.048>
- THENMOZHI R., LEE J.S., PARK N.Y., CHOI B.O., HONG Y.B. 2020. Gene therapy options as new treatment for inherited peripheral neuropathy. *Exp. Neurobiol.* **29**: 177-188. <https://doi.org/10.5607/en20004>
- THONNEAU P., MARCHAND S., TALLEC A., FERIAL M.L., DUCOT B., LANSAC J., LOPES P., TABASTE J.M., SPIRA A. 1991. Incidence and main causes of infertility in a resident population (1,850,000) of three french regions (1988-1989). *Hum. Reprod.* **6**: 811-816. <https://doi.org/10.1093/oxfordjournals.humrep.a137433>



- THUDI M., LI Y., JACKSON S.A., MAY G.D., VARSHNEY R.K. 2012. Current state-of-art of sequencing technologies for plant genomics research. *Brief Funct. Genomics* **11**: 3-11. <https://doi.org/10.1093/bfpg/elr045>
- TURNER K.A., RAMBHATLA A., SCHON S., AGARWAL A., KRAWETZ S.A., DUPREE J.M., AVIDOR-REISS T. 2020. Male infertility is a women's health issue – Research and clinical evaluation of male infertility is needed. *Cells* **9**: 990. <https://doi.org/10.3390/cells9040990>
- WANG K., YUEN S.T., XU J., LEE S.P., YAN H.H., SHI S.T., SIU H.C., DENG S., CHU K.M., LAW S., CHAN K.H., CHAN A.S., TSUI W.Y., HO S.L., CHAN A.K., MAN J.L., FOGLIZZO V., NG M.K., CHAN A.S., CHING Y.P., CHENG G.H., XIE T., FERNANDEZ J., LI V.S., CLEVERS H., REJTO P.A., MAO M., LEUNG S.Y. 2014. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat. Genet.* **46**: 573-582. <https://doi.org/10.1038/ng.2983>
- WANG X., HE Y., FU X., PENG D. 2012. Protective effect of lentivirus-mediated Bcl-2 gene transfection against phosphoramidate mustard-induced apoptosis of human ovarian granulosa cells. *J. South. Med. Univ.* **32**: 932-936. (In Chinese with English abstract)
- WILTSHIRE A., GHIDEI L., BRAYBOY L.M. 2020. Infertility and assisted reproductive technology outcomes in Afro-Caribbean women. *J. Assist. Reprod. Genet.* **37**: 1553-1561. <https://doi.org/10.1007/s10815-020-01826-2>
- WU Y., HALVERSON G., BASIR Z., STRAWN E., YAN P., GUO S.W. 2005. Aberrant methylation at HOXA10 may be responsible for its aberrant expression in the endometrium of patients with endometriosis. *Am. J. Obstet. Gynecol.* **193**: 371-380. <https://doi.org/10.1016/j.ajog.2005.01.034>
- ZAMUDIO N.M., SCOTT H.S., WOLSKI K., LO C.Y., LAW C., LEONG D., KINKEL S.A., CHONG S., JOLLEY D., SMYTH G.K., DE KRETZER D., WHITELAW E., O'BRYAN M.K. 2011. DNMT3L is a regulator of X chromosome compaction and post-meiotic gene transcription. *PLoS One* **6**: e18276. <https://doi.org/10.1371/journal.pone.0018276>
- ZEGERS-HOCHSCHILD F., ADAMSON G.D., DYER S., RACOWSKY C., DE MOUZON J., SOKOL R., RIENZI L., SUNDE A., SCHMIDT L., COOKE I.D., SIMPSON J.L., VAN DER POEL S. 2017. The international glossary on infertility and fertility care, 2017. *Hum. Reprod.* **32**: 1786-1801. <https://doi.org/10.1093/humrep/dex234>
- ZHANG C.Z., ADALSTEINSSON V.A., FRANCIS J., CORNILS H., JUNG J., MAIRE C., LIGON K.L., MEYERSON M., LOVE J.C. 2015. Calibrating genomic and allelic coverage bias in single-cell sequencing. *Nat. Commun.* **6**: 6822. <https://doi.org/10.1038/ncomms7822>
- ZHAO M.T., RIVERA R.M., PRATHER R.S. 2013. Locus-specific DNA methylation reprogramming during early porcine embryogenesis. *Biol. Reprod.* **88**: 48. <https://doi.org/10.1095/biolreprod.112.104471>
- ZHENG Y.M., LI L., ZHOU L.M., LE F., CAI L.Y., YU P., ZHU Y.R., LIU X.Z., WANG L.Y., LI L.J., LOU Y.Y., XU X.R., LOU H.Y., ZHU X.M., SHENG J.Z., HUANG H.F., JIN F. 2013. Alterations in the frequency of trinucleotide repeat dynamic mutations in offspring conceived through assisted reproductive technology. *Hum. Reprod.* **28**: 2570-2580. <https://doi.org/10.1093/humrep/det294>
- ZONG C., LU S., CHAPMAN A.R., XIE X.S. 2012. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* **338**: 1622-1626. <https://doi.org/10.1126/science.1229164>