GYNECOLOGY AND OBSTETRICS RESEARCH
Openventio PUBLISHERS
ISSN 2377-1542 Open Journal http://dx.doi.org/10.17140/GOROJ-3-131

Mini Review

*Corresponding author
Parveen Parasar, BVSc, MVSc, PhD
Boston Center for Endometriosis
Brigham and Women's Hospital
Boston Children's Hospital
Harvard Medical School
Boston, MA 02115, USA
E-mail: suprovet@gmail.com

Volume 3 : Issue 1
Article Ref. #: 1000GOROJ3131

Article History
Received: June 23rd, 2016
Accepted: July 27th, 2016
Published: July 28th, 2016

Citation
Parasar P. High-throughput next generation sequencing: Applications in reproductive diagnosis and research. *Gynecol Obstet Res Open J*. 2016; 3(1): 19-22. doi: 10.17140/GOROJ-3-131

High-Throughput Next Generation Sequencing: Applications in Reproductive Diagnosis and Research

Parveen Parasar, BVSc, MVSc, PhD*

Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA

INTRODUCTION

The genetic information contained within a cell is transferred through the process of transcription of genes within a genome to produce messenger RNAs (mRNAs) and translation of mRNAs to synthesize proteins. The central dogma pathway represents all three stages of replication, transcription, and translation in the pathway:

DNA → RNA → Protein

Genes in a genome can be identified by creating a complementary DNA (cDNA) library from the pool of ribonucleic acid (RNA) transcripts. To generate a cDNA library, the RNA transcripts from a tissue or from cells are copied into more stable cDNA molecules, which are then stored into an appropriate vector to generate a collection of cDNA clones. The single pass, short 300-500 nucleotide sequences obtained from sequencing either end of the cDNA insert are called expressed sequence tags (ESTs). ESTs can be generated from the cDNA libraries obtained from the patient tissue/samples. Those ESTs can be used to determine the genes that express them and to determine whether they possess any nucleotide or single nucleotide polymorphisms (SNPs) in comparison to normal individual. The sites where deoxyribonucleic acid (DNA) sequences are different at a single nucleotide are called SNPs. Similarly, capability to detect and identify mutations in genes has been utilized by high-throughput sequencing methods.

Predominantly, Sanger method of sequencing is used for DNA sequencing and is the method routinely used for the past 38 years. Sequencing methodologies have been rapidly improved creating very powerful tools for detection and identification of SNPs and various infectious agents. Sanger sequencing process is carried out in a liquid phase where a predefined gene can be targeted in one run allowing extremely low rates of false positive and negative errors and producing high specificity. However, the process is very time consuming, especially when multiple genes are to be studied in a sample to detect variants underlying a disease.

Attempts to sequence larger genomes such as the whole genomes of various animal species, using multicapillary sequencing faced considerable caveats in scalability, speed and resolution. Subsequent major advances in technologies such as cyclic-array sequencing gave rise to 'second-generation or next-generation sequencing (NGS)'. In these technologies, repeated cycles occur during which DNA sequences, immobilized on a solid substrate, and are determined one base position at a time with the use of enzymatic manipulation and imaging-based data acquisition.

NEXT GENERATION SEQUENCING AS A DIAGNOSTIC TOOL

Second generation sequencing also called as NGS sequencing was first reported in 2005. It allows for great increase in throughput and potential cost reduction. With the introduction of next generation sequencing it is now possible to concomitantly sequence multiple genomes in less than two weeks all in single run. There are three overall processes that have been developed in NGS: whole genome sequencing (WGS), whole exome sequencing (WES) and targeted gene

ISSN 2377-1542 Open Journal http://dx.doi.org/10.17140/GOROJ-3-131

sequencing (TGS).[1] WGS is capable of sequencing the entire genome in a single run while WES and TGS are more focused. WES focuses on only the protein coding regions, and TGS is the most focused and examines specific genes or regions of interest making it of key interest to clinical investigations (Figure 1). These advancements in biotechnology have been applied to the study of genetic diseases in animals and thus revolutionized the study of biological and evolutionary processes at the molecular level.

Contrary to earlier techniques, NGS reports all nucleotide sequences present in the original sample by enabling deeper sequencing at a faster and economically affordable rate thus improving the likelihood of identifying novel mutations and genetic variations. Furthermore, sampling, sample preparation and enrichment protocols have significant effects on the outcome of NGS-based diagnostics (Figure 2).

Sample Preparation

Sample preparation is composed of individual steps *viz.* homogenization, filtration, and nuclease treatment including nucleic acid extraction and purification followed by amplification. Homogenates and fluids must be centrifuged and microfiltered to get rid of free and released fine particles from larger particles. Enrichment is performed with the use of nuclease treatment. Deoxyribonuclease (DNase) and ribonuclease (RNase) are used in combination or alone to eliminate host contaminants.

Library Construction and Sequencing

Genomic sequences in samples are converted into sequencing libraries apt for cluster generation and sequencing. This process involves following steps:

a. fragmentation of DNA (mechanical or enzymatic shearing).
b. end-repair, modification and ligation of adapters to amplify sheared DNA by adapter- specific primers.
c. size-selection of DNA molecules with an optimal length for the current application of instrument, and
d. enrichment of adapter-ligated DNA by PCR (if the amount of source material is limited).

The constructed libraries are immobilized on a solid

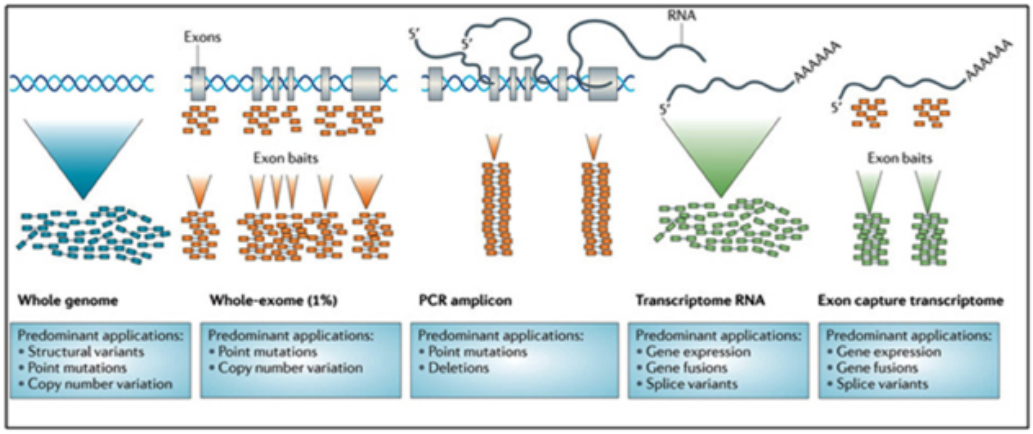

Figure 1: Applications of next-generation sequencing (NGS).[2]

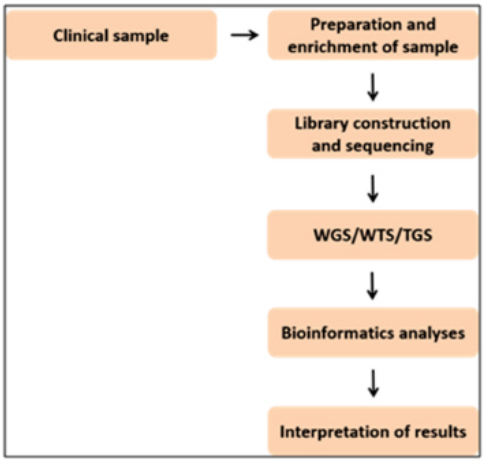

Figure 2: Next-generation sequencing workflow.

ISSN 2377-1542 Open Journal http://dx.doi.org/10.17140/GOROJ-3-131

substrate for clonal amplification to generate distinct clusters of DNA copies. The commonly used solid platform are beads inside aqueous reaction bubbles (emulsion PCR or emPCR such as in 454 Pyrosequencing machine from Roche), glass flow cell (solid such as in Illumina machine from Illumina Inc.), ion-based sequencing chips (such as Ion Torrent from Life Technologies) etc.

Bioinformatics

The most common type of the data output from the machine is in the Fast Q format. Post-acquisition of data, researcher performs a quality control on the dataset to establish that the experiment or sample was processed according to standard parameters. Subsequently, data are screened for irrelevant sequence information, such as non-specific contaminants, and tag sequences. DNA or RNA sequence data are mapped towards the host genome or the known homologue. In the absence of homologue, data are processed through an algorithm to form larger sets of continuous reads or contigs. Contigs are them mapped towards the closest relative homologue to form a draft genome.[3,4] Bioinformatics tools enable targeted sequencing of RNA (RNA-seq) which not only maps the available mRNA sequence but also non-coding RNAs such as micro RNA (miRNA), small interfering RNA (siRNA, long non-coding RNA (lncRNA), and ribosomal RNA (rRNA). Consequently, next generation sequencing and associated bioinformatics is an efficient tool for visualizing RNA content within a sample.

APPLICATIONS OF NEXT GENERATION SEQUENCING IN OBSTETRICS AND GYNECOLOGY RESEARCH

Next-generation sequencing has been pivotal in diagnosis of known and emerging new mutations and SNPs that correlate with the susceptibility to disease occurrence. This technology enables researchers to identify gene-polymorphisms. NGS has been employed in many areas in medicine including cancer research, personalized-precision medicine and reproductive medicine and obstetrics and gynecology research.[5] Below are some examples of applications of NGS in reproductive diagnostics and gynecologic research.

Preimplantation Genetic Screening (Pgs)/Pre-Implantation Genetic Diagnosis (PGD)

Preimplantation genetic diagnosis eliminates the risk of passing genetic disorders to new born even prior to the woman becoming pregnant. This can be accomplished by examination and identification of potential genetic defects in developing embryo. With the increased age the aneuploidy of embryos increases which results in miscarriage, congenital abnormalities an implantation failure. Patients of recurrent implantation failure and recurrent miscarriages would benefit by Preimplantation genetic screening as it will improve clinical outcomes by increasing implantation rates and reducing miscarriage rates. Million of DNA fragments from embryos can be analyzed using NGS with greater sensitivity.[6-8]

Non-Invasive Prenatal Testing (NIPT)

Cell-free DNA (cfDNA) is now-a-days used as an advanced screening tool to determine fetal aneuploidy. Amount of fetal DNA from each chromosome is proportional to that of the mother in a normal pregnancy which is deviated slightly in an aneuploid pregnancy. For example, in Down syndrome, chromosome 21 has 3 copies of fetal DNA instead of normal 2 fetal copies. NGS allows detection and identification of such minute variations by sequencing and enumerating millions of cfDNA fragements from maternal plasma.[9,10]

Prenatal Diagnosis of Rare Genetic Diseases

NGS is now commonly used to sequence the protein coding genes collectively known as exome (whole exome sequencing (WES)) and is very powerful tool to detect many unknown and new genetic disease.[11]

Endometrial Receptivity Analysis (ERA)

Endometrial receptivity and implantation are critical features for a successful pregnancy outcome. ER is a characteristic phenotype which allows impantation and adhesion of embryo with the endometrium. With the advancements in transcriptomic analyses the genetic signatures of viable and potentially successful phenotype of endometrium can be predicted which leads to improved assisted reproductive technologies (ART) with better outcomes.[12] NGS-transcriptomics allows characterization of genes even at the miRNA level to establish sample-specific molecular profile in embryonic cells. This profile could be used as biomarkers defining a successful biological process or a disease. Consequently, this concept is useful in determining a window for successful transfer of embryo in ART and *in vitro* fertilization (IVF) techniques.[13]

One study reported the immensely usefulness of this high-throughput technology in targeting genes and molecular markers in obstetrics gynecological cancers such as ovarian cancer. NGS has been applied to analyze TGFβ/SMAD4 targets in ovarian cancer to identify genome-wide SMAD4 targets in epithelial ovarian cancer and aberrant TGFβ/SMAD4 signaling in ovarian tumorigenesis.[6] Taken together, NGS is robust technique and has great potential for reproductive diagnostics to detect genetic diseases. In addition, it is robust tool to control and monitor the genetic diversity of viral, bacterial and other causative agents that infect reproductive system and identification of potentially genetic variants.

These advantages as screening and characterization tool make next-generation sequencing ideal for vaccine development and quality control. Collectively, these merits demonstrate the new possibilities opened up by the NGS and metagenomics analyses to study known as well as new emerging diseases in domestic and wild animal populations.

The combination of diverse molecular biology and ge-

ISSN 2377-1542 Open Journal http://dx.doi.org/10.17140/GOROJ-3-131

nomics skills in multidisciplinary fashion is very important to enhance and extend our ability to develop effective and accurate diagnostic tools and disease control measures.

REFERENCES

1. Cordero P, Ashley EA. Whole-genome sequencing in personalized therapeutics. *Clin Pharmacol Ther*. 2012; 91(6):1001-1009.

2. Simon R, Roychowdhury S. Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov*. 2013; 12(5): 358-369. doi: 10.1038/nrd3979

3. Yegnasubramaniam S. Explanatory chapter: Next generation sequencing. *Methods Enzymol*. 2013; 529:201-208. doi: 10.1016/B978-0-12-418687-3.00016-1

4. Rehm HL, Bale SJ, Bayrak-Toydemir P, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med*. 2013; 15(9): 733-747. doi: 10.1038/gim.2013.92

5. Bombard Y, Bach PB, Offit K. Translating genomics in cancer care. *J Natl Compr Cancer Netw.* 2013; 11: 1343-1353.

6. Kennedy BA, Deatherage DE, Gu F, et al. ChIP-seq defined genome-wide map of TGFβ/SMAD4 targets: Implications with clinical outcome of ovarian cancer. *PLoS One.* 2011; 6(7): e22606. doi: 10.1371/journal.pone.0022606

7. Hallam S, Nelson H, Greger V, et al. Validation for clinical use of, and initial clinical experience with, a novel approach to population-based carrier screening using high-throughput, next-generation DNA sequencing. *J Mol Diagn.* 2014, 16: 180-189. doi: 10.1016/j.jmoldx.2013.10.006

8. Fiorentino F, Bono S, Biricik A, 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

9. Bianchi DW, Parker RL, Wentworth J, et al. DNA sequencing versus standard prenatal aneuploidy screening. *New Engl J Med.* 2014; 370(9). Web site. http://www.nejm.org/doi/pdf/10.1056/nejmoa1311037. Accessed June 2nd, 2016

10. Norton ME, Jacobsson B, Swamy GK, et al. Cell-free DNA analysis for noninvasive examination of trisomy. *New Engl J Med.* 2015; 372: 1589-1597. doi: 10.1056/NEJMoa1407349

11. Farwell KD, Shahmirzadi L, El-Khechen D, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model–based analysis: Results from 500 unselected families. *Genet Med.* 2014; 17(7): 578-586. doi: 10.1038/gim.2014.154

12. Díaz-Gimeno P, Ruíz-Alonso M, Blesa D, Simón C. Transcriptomics of the human endometrium. *Int J Dev Biol*. 2014; 58(2-4): 127-137. doi: 10.1387/ijdb.130340pd

13. Werner T. Next generation sequencing allows deeper analysis and understanding of genomes and transcriptomes including aspects to fertility. *Reprod Fertil Dev*. 2011; 23(1): 75-80. doi: 10.1071/RD10247