

Review on Next Generation DNA Sequencing: Historical Emergence, Advantages and Applications

¹Yosef Deneke, ¹Tefera Getahun and ²Garoma Desa

¹Jimma University, College of Agriculture and Veterinary Medicine,
P.O. Box: 307, Jimma, Ethiopia

²National Institute for Control and Eradication of Tsetse Fly and Trypanosomosis,
Kaliti Tsetse Fly Mass Rearing and Irradiation Center, Addis Ababa, Ethiopia

Abstract: The development of Next generation sequencing (Massively Parallel Signature Sequencing) took place in late 20th and early 21st century. The first Next generation sequencing technology was started by a USA based Lynx Therapeutics Company in 2000 and it represented the next step in the evolution of DNA sequencing, through the generation of thousands to millions of DNA sequences in a short time. Innovative fields and applications in biology and medicine are becoming a reality, beyond the genomic sequencing which was original development aim and application. The traditional medicine model of diagnosis has changed to one precision medicine model, leading to a more accurate diagnosis of human diseases and allowing the selection of molecular target drugs for individual treatment. A new generation of non-Sanger-based sequencing technologies has delivered on its promise of sequencing DNA at unique speed, thereby enabling impressive scientific achievements and new biological applications. The potential use of Next generation sequencing in precision medicine is vast and a better awareness of this technique is necessary for a successful execution in the clinical workplace, personal genomics with full analysis of individual genome stretches; exact analysis of RNA transcripts for gene expression and replacing in several analysis by various microarray platforms. An integrated part describing the main Next generation sequencing aspects in the clinic could help beginners, scientists, researchers and health care professionals, as they will be responsible for translating genomic data into genomic medicine.

Key words: Generation • Next • Sequencing • Innovative

INTRODUCTION

Next generation sequencing techniques are high-throughput technologies with proficiencies of sequencing huge numbers of diverse DNA (massively parallel) sequences at a time. This technology monitors the sequential addition of nucleotides to immobilized DNA templates made from target tissue [1]. Unfortunately, the increased throughput of NGS (next-generation sequencing) reactions comes at the cost of shorter sequences, as most sequencing platforms (Illumina, Roche, SoLiD) offer shorter read lengths (30–400 bp) than the conventional Sanger-based method. These shorter sequences are then collected into longer sequences such as complete genomes [2]. In recent times, Sanger

sequencing, referred to as a ‘first-generation’ sequencing method, has partly been replaced by NGS methods. NGS allows identifying biomarkers for early diagnosis as well as for personalized treatments. The emergence of NGS has changed the way clinical research, basic and applied sciences are done. The NGS allows producing millions of data with a smaller investment [3].

NGS technology is gradually used in many fields. Its possibility to gain large amount of data, discover new and essential information about the human genome. This feature opened many contexts of effective applications. The first was whole-genome sequencing (WGS), a method intended for entire genome sequencing which offers the most complete landscape of genomic information and possible biological consequences [4].

Next generation sequencing technology is presently the hottest topic in the field of human and animal's genomics researches [5]. It uses parallel sequencing of multiple small fragments of DNA to determine sequence [6]. In opposite to Sanger sequencing, the rapidity of sequencing and amounts of DNA sequence data generated with NGS, which is reflected a "high-throughput technology", are exponentially better and are made at significantly low costs [7].

Until recently, the Sanger sequencing method was the most widely used sequencing technique and resulted in the merely complete human genome sequence [8]. This technology depends on incorporation of chain-terminating dideoxynucleotides during DNA replication [9]. Fluorescently labeled terminators, capillary electrophoresis separation and laser signal detection have improved the throughput of Sanger sequencing [10]. However, it remains labor-intensive, time-consuming and expensive soon done in large scale. Consequently, the request for quicker, more precise and more cost-effective genomic information has directed to the development of NGS methods [1]. Another potential application of NGS is to advance the diagnosis of cancer. Poor tissue sampling and processing can often make a histological diagnosis tough. NGS-based analysis of tissue can be performed on small amounts of viable tissue and is accurate when sufficient information regarding causative mutations is known, as the genomics of different tumors become apparent, it can be used to identify different molecular subtypes, which is already becoming common with sarcoma fusion proteins [10].

Next generation sequencing technology is progressively used in many fields. Its power owns the ability to gain huge amount of data and discover new and essential information about the human genome. This feature opened many contexts of successful applications. The first was whole-genome sequencing (WGS), an approach intended for entire genome sequencing. It offers the most complete landscape of genomic information and possible biological consequences [4]. The NGS is basically diverse approach to sequencing that lead to several ground-breaking discoveries and brought a new revolution in genomic research by revealing limitless insight related to genome, transcriptome and epigenome of any species. Hence, NGS technology has brought a new revolution in the welfare of human society. NGS allows sequencing of large stretch of DNA base pairs, producing hundreds of gigabases of data in a single sequential run. More recently, third

generation sequencing (TGS) has evolved; though, it is still in its beginning [8]. There is scarce information on advantages and applications of NGS technique. Therefore the aim of this seminar is to review recent information of historical emergence, advantages and applications of next generation DNA sequencing.

Literature Review

Historical Emergence of next Generation Sequencing

Technique: In 1908, Garrod announced the idea of 'inborn mistake of metabolism' that altered the parts of biochemistry, genetics and medicine. His major contribution was the understanding about the link between gene-enzyme and the molecular basis of genetic diseases. Even though nowadays this concept is considered outdated because of findings like RNA splicing, RNAi and others, his contribution permitted the researchers to know how changes in DNA sequence probably will cause genetic disease. This finding enhances the interest of scientists to know about human DNA sequence and mutations. The quest to understand the nucleotide sequence of DNA began in the 1960s [11].

Scientific improvements in whole genome sequencing progressed through three major technological revolutions: first generation sequencing (whole genome shotgun sequencing), next generation sequencing (NGS high throughput sequencing) and the third generation of sequencing (single molecule long read sequencing) [12]. In 1977 Sanger established the technique termed 'Chain-termination' that became the extremely used method (first generation) to sequencing DNA at a time. The technique involved the use of dideoxynucleotides (ddNTPs), which are deoxynucleotide analogs (dNTPs) that interrupt DNA production and the separation of the diverse DNA fragments in a gel [15, 12] and the development of the polymerase chain reaction and the parting of DNA fragments by capillary electrophoresis continued [14].

The next (second) generation of DNA sequencing is the period of the parallel massive sequencing on a micro scale. The technique was established by Nyrén and colleagues in 1996. It varied considerably from previous ones as it did not use radio or fluorescence-labeled nucleotides and there was no need of electrophoretic run. The technique is based on the action of two enzymes: ATP sulfurylase and luciferase. This method releases light signal in proportion to the amount of nucleotides incorporated and the sequence can be determined according to the serial addition of nucleotides [15].

In future, this technology was promoted and licensed producing the first 'second-generation' equipment, known as 454 (Roche). These amendments and the use of microplates that grouped the process and high-definition detection systems completely increased the amount of DNA sequenced and defined the second generation [16], since they advance the accuracy of mapping, mainly in repetitive regions or where DNA rearrangements or gene fusions occur. The method uses 'reversible terminator chemistry' which is a modified fluorescent dNTP that reversibly blocks DNA synthesis, so the addition of each nucleotide can be harmonized and monitored by a charge-coupled device (CCD) sensor [17]. This is one of the most precise and with lowest error rate of sequencing methodologies used currently; though, it commonly requires advanced DNA concentration. Alternative methodology is based on oligonucleotide ligation sequencing known as SOLiD and developed by Applied Biosystems. The method does not do sequencing by synthesis but by ligation of oligonucleotides fluorescence-labelled. The template from primer (n) is removed and the second round of sequencing is performed with a primer complementary to the (n-1) position [18]. Third-generation is sequencing uses parallel sequencing similar to NGS, but unlike NGS, third-generation sequencing uses single DNA molecules rather than amplified DNA as a template. Thus, third generation sequencing potentially eliminates errors in DNA sequence introduced in the laboratory during the DNA amplification process [19].

Application and Advantages of next Generation Sequencing: Next-generation sequencing technology considered as the future of high-throughput data analysis and genomic sequencing, providing a way to gain high-throughput data with sensitivity and specificity. It offers the possibility to have massive parallel multigene sequencing in few hours, with essential time and cost reduction, by using very small amount of nucleic acids [20]. The clear potential of these technologies is to improve sequencing ability, leading to more complete definition of the genomic background. This is vital, particularly for the study of complicated diseases, such as cancer, since it permits to obtain a wider view of the genotype. NGS techniques can provide valuable data about mutational status, copy number differences, transcriptomics and epigenetics with the occasion to combine current existing single genetic tests into an exceptional test able to detect multiple variants. Great care should be focused on the probability to generate

databases where the sequencing information of single patients can be stored, resulting accessible for future use by clinicians in terms of retrospective analysis and, possibly, therapeutic conclusions [21].

Diverse bioinformatics tools are supplied with the NGS platforms, but additional measures are crucial to encourage good data generation and interpretation [22]. High-throughput sequencing of the human genome helps us to discover genes and regulatory pathways related with disease. It helps in the faster diagnosis and outcome of disease-targeted sequencing may help in better therapeutic decision-making for several genetic diseases, including many cancers. RNA-Seq (NGS of RNA) provides complete transcriptomic information of a sample without any need of previous knowledge associated to genetic sequence of an organism. RNA-Seq provides a strong alternative approach to Microarrays for gene expression studies and let the researchers imagine RNA expression in the form of sequence [23]. Different study is common in medical genetic, where DNA sequence and data are compared with a reference sequence to list the differences. These differences may range from single nucleotide polymorphisms (SNPs) to complex chromosomal rearrangement [24].

Next Generation Sequencing Cancer Genomics Identification: Currently, NGS has been used to characterize genomic changes such as mutations, insertions/deletions and copy number changes and the frequency with which they occur in various tumor types. Efforts such as the International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA) aim to set such genomic alterations across many tumor types [25]. Comparison of sequenced genomes to reference genomes permits for the identification of genome changes that may be important in disease development and progression [26]. However, such comparison depends on the launch of extensive and precise reference genomes, which is a bulky task. Additionally, the complexity of genomic deviations in cancer marks it difficult to depend on standard reference genomes [2].

Consequently, simplified methods of identifying driver mutations are necessary. Numerous theories exist for the potential identification of driver mutations. One of such hypothesis is that mutations that happen with higher frequency are more likely to give to tumor development and growth. Genome-wide association studies (GWAS) aim to compare the incidence of commonly known single nucleotide polymorphisms

(SNPs) in genomes from patients with and without a specific disease. SNPs that happen at a higher frequency in the diseased population are identified as potentially causative. If a specific mutation is not found in high frequency, but the same molecular pathway contains frequent genomic changes, those changes may also be significant. Another theory is that changes present in both germ line and tumor tissue of the same patient is likely to be essential to tumor development [27].

Next Generation Sequencing Aid in Clinical Decision:

Nowadays NGS has made possible a better understanding of genetic diseases and became an important technological advance in the practice of diagnostic and clinical medicine. NGS allows the analysis of multiple regions of the genome in one single reaction and has been shown to be a cost-effective and an efficient tool in investigating patients with genetic diseases. Genetic data produced via NGS provides significant benefits to medical practice including accurate identification of biomarkers of disease, detecting inherited disorders and identifying genetic factors that can help predict responses to therapies.

However, recommendations on clinical implementation of NGS that are still in discussion and that hamper its use in the genetic clinic. A variety of molecular diagnostic test use sequencing technology, such as single- and multi-gene panel tests, cell-free DNA for non-invasive prenatal testing, whole-exome sequencing (WES), whole-genome sequencing (WGS). Considering that the use of NGS as a diagnostic tool is recent, there are challenges including when to order, on whom to order and how to interpret and communicate the results to the patient and family [10]. Once driver mutations have been identified in a tumor, the next step is to assess whether those mutations are “actionable.” Actionable alterations affect the function of a cancer-related gene and can be targeted with approved or investigational therapies. Assessing functionality is a difficult task and requires predictive knowledge of genome alterations. Often, early-phase studies are used to assess the role various mutations based on rates of response to targeted therapies. However, enrollment in such studies requires that physicians be aware of genome alterations and potential trials for each patient [28].

Common Applications of Next Generation Sequencing Techniques into Clinical Practice

Multi-Gene Panels: Single-gene testing is selected when the clinical features for a patient are distinctive for a

particular disorder and the link between the disorder and the specific gene is well established and has the minimal locus heterogeneity [29]. It is more appropriate. For instance in diagnosis of cancer Tothill and colleagues [30] demonstrated the application of this multi-gene panel by analyzing samples of patients with cancers of unknown primary. The clinical supervision of patients with CUP is hindered by the lack of a definitive site of origin and this kind of NGS analysis could help to define new therapeutic options. In multi-gene panel tests, many genes associated with a specific phenotype are sequenced and analyzed concurrently, decreasing cost and improving efficiency of genetic diagnosis.

The conversion from single-gene to multi-gene testing should not compromise the sensitivity of the test to identify variants, principally at genes that are responsible for a significant proportion of the defects (core genes). The sensitivity of NGS does not depend only on horizontal coverage but the vertical coverage is important as well. Additional genes will increase the chance of the diagnostic, but this should not be at cost of missing mutations that would previously have been detected by single-gene testing. Sanger sequencing or other available techniques can help to solve this problem for filling in low-coverage and no-coverage regions [31].

Whole-Genome and Exome Sequencing: Whole-genome sequencing (WGS) also recognized as, full-genome sequencing, complete genome sequencing or entire genome sequencing is the process of determining the complete DNA sequence of an organism's genome at a single time. The major benefit of WGS is completed coverage of the genome, including promoters and regulatory regions. In whole-exome sequencing (WES), all coding regions are sequenced with a relatively deeper depth. Compared to WGS, the major benefit of WES is a significant cost reduction [32]. NGS tools are still prone to sequencing artifacts and Sanger sequencing is recommended to confirm the variants detected before returning the results to the patient [33].

RNA-Sequencing: RNA-seq permits detection of new genes and isoforms, gene fusions, splice and chimeric variants, genomic alters and gene expression quantification. Though, RNA-seq outperforms microarray in transcriptomic analysis [34]. RNA-seq is considered a complementary method depending on the needs and resources available, assisting clinicians in making decisions. In clinical practice, RNA measurement has applications across different areas in human health such

as therapeutic selection, disease diagnostic and treatment [35]. Clinical diagnosis of infectious disease through RNA-seq is still rare, since quantitative PCR (RT-qPCR) assays are still the most common technique used for viral detection and genotyping. Applications of NGS in virology diagnostic can be used for analysis of patients with mysterious illness, mainly during outbreaks and epidemics [36]. NGS is a more powerful tool for ctRNA detection; however, RT-qPCR remains more usable for clinical diagnostic applications [37].

Epigenetics: An emerging arena that has a great impact on medicine and clinical diagnostic is epigenetics. Epigenetics mechanisms embody another layer of gene regulation and NGS permitted to recognize the epigenetics status on a large scale and at a single base-resolution, with mainly DNA methylation, histone modification and non-coding RNA (ncRNA)-associated silencing. DNA methylation was the first epigenetic mechanism identified and is the best known and the most frequent in human cancer. It involves covalent modification of cytosine through the addition of a methyl group to cytosines of CpG (cytosine/guanine) islands. This methylation is maintained by DNA methyltransferase (DNMTs) and plays roles for gene transcriptional repression, transposable elements silencing and viral defence. Unmethylated DNA is found in active regions of chromatin and methylated DNA is found in inactive regions [38].

Next Generation Sequencing in Genomic Evolution, Intertumor and Intratumor Heterogeneity: Complicating the application of genomic medicine is the fact that causal mutations can evolve during the course of cancer. As tumors are treated or as they grow, a variability of acquired genomic changes may occur. For instance, melanoma treated with BRAF or MEK inhibitors has been shown to acquire BRAF amplifications and downstream changes that lead to reactivation of the MAP kinase pathway [39]. Likewise, increased signaling via the phosphatidylinositol 3-kinase/Akt pathway may contribute to resistance in HER2-positive breast cancer [40]. Tumors may also develop intertumor and intratumor heterogeneity. Intertumor heterogeneity refers to differences in changes of tumors at different sites, while intratumor heterogeneity refers to differences in alterations within a tumor[41].

Future Next Generation Sequencing Applications and Directions: Several other applications of NGS are under development. One potential future application of NGS

is the evaluation of circulating tumor cells or free plasma DNA to detect early relapse or residual cancer. Once tumor-specific genome alterations have been identified by NGS, PCR assays could be used to detect circulating tumor cells or free-plasma DNA harboring the same changes. Disease status, drug responsiveness and relapse could be serially evaluated. The monitoring strategy would, yet, require the mutation being tested be present in all tumor cells and remain present in the course of disease [42].

CONCLUSIONS

Next generation sequencing is a new technique for sequencing genomes at high speed and minimal cost. It stands out as one of the most powerful and effective approach for fast DNA/RNA sequencing. Next-generation high-throughput DNA sequencing techniques are opening fascinating opportunities in the life sciences. Novel fields and applications in biology and medicine are becoming a reality, beyond the genomic sequencing which was original development goal and application. The next-generation sequencing technologies offer original and rapid ways for genome-wide characterization and profiling of mRNAs, small RNAs, transcription factor regions, structure of chromatin and DNA methylation patterns, microbiology and metagenomics. Finally, NGS can identify molecular aberrations that render tumors exquisitely sensitive to certain therapies, resulting in exceptional responses. Such extraordinary outcomes can improve our understanding of molecular features that can predict response to certain drugs.

REFERENCES

1. Metzker, M.L., 2010. Sequencing technologies the next generation. *Nature Reviews Genetics*, 11(1): 31.
2. Ulahannan, D., M.B. Kovac, P.J. Mulholland, J.B. Cazier and I. Tomlinson, 2013. Technical and implementation issues in using next-generation sequencing of cancers in clinical practice. *British Journal of Cancer*, 109(4): 827.
3. Rabbani, B., H. Nakaoka, S. Akhondzadeh, M. Tekin and N. Mahdieh, 2016. Next generation sequencing: implications in personalized medicine and pharmacogenomics. *Molecular BioSystems*, 12(6): 1818-1830.
4. Cirulli, E.T. and D.B. Goldstein, 2010. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews Genetics*, 11(6): 415.

5. Pareek, C.S., R. Smoczynski and A. Tretyn, 2011. Sequencing technologies and genome sequencing. *Journal of Applied Genetics*, 52(4): 413-435.
6. Rizzo, J.M. and M.J. Buck, 2012. Key principles and clinical applications of “next-generation” DNA sequencing. *Cancer Prevention Research*, 5(7): 887-900.
7. Voelkerding, K.V., S.A. Dames and J.D. Durtschi, 2009. Next-generation sequencing: from basic research to diagnostics. *Clinical Chemistry*, 55(4): 641-658.
8. International Human Genome Sequencing Consortium, 2004. Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011): 931.
9. Sanger, F., 1983. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, 80: 2432-2436 Year Changed 1977 to 1983.
10. Nikiforov, Y.E., S.E. Carty, S.I. Chiosea, C. Coyne, U. Duvvuri, R.L. Ferris, W.E. Gooding, S.P. Hodak, S.O. Le Beau, N.P. Ohori and R.R. Seethala, 2014. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer*, 120(23): 3627-3634.
11. Scriver, C.R., 2008. Garrod's Croonian Lectures (1908) and the charter ‘Inborn Errors of Metabolism’: albinism, alkaptonuria, cystinuria and pentosuria at age 100 in 2008. *Journal of Inherited Metabolic Disease*, 31(5): 580-598.
12. Loman, N.J. and M.J. Pallen, 2015. Twenty years of bacterial genome sequencing. *Nature Reviews Microbiology*, 13(12): 787-794.
13. Smith, L.M., S. Fung, M.W. Hunkapiller, T.J. Hunkapiller and L.E. Hood, 1985. The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. *Nucleic Acids Research*, 13(7): 2399-2412.
14. Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich and N. Arnheim, 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230(4732): 1350-1354.
15. Ronaghi, M., S. Kara Mohamed, B. Pettersson, M. Uhlén and P. Nyren, 1996. Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry*, 242(1): 84-89.
16. Margulies, M., M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemben, J. Berka, M.S. Braver Man, Y.J. Chen, Z. Chen and S.B. Dewell, 2005. Genome sequencing in micro fabricated high-density picolitre reactors. *Nature*, 437(7057): 376.
17. Bentley, D.R., S. Balasubramanian, H.P. Swerdlow, G.P. Smith, J. Milton, C.G. Brown, K.P. Hall, D.J. Evers, C.L. Barnes, H.R. Bignell and J.M. Boutell, 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218): 53.
18. McKernan, K.J., H.E. Peckham, G.L. Costa, S.F. McLaughlin, Y. Fu, E.F. Tsung, C.R. Clouser, C. Duncan, J.K. Ichikawa, C.C. Lee and Z. Zhang, 2009. Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. *Genome Research*, 19(9): 1527-1541.
19. Munroe, D.J. and T.J. Harris, 2010. Third-generation sequencing fireworks at Marco Island. *Nature Biotechnology*, 28(5): 426.
20. Goswami, R.S., R. Luthra, R.R. Singh, K.P. Patel, M.J. Routbort, K.D. Aldape, H. Yao, H.D. Dang, B.A. Barkoh, J. Manekia and L.J. Medeiros, 2016. Identification of factors affecting the success of next-generation sequencing testing in solid tumors. *American Journal of Clinical Pathology*, 145(2): 222-237.
21. Kamalakaran, S., V. Varadan, A. Janevski, N. Banerjee, D. Tuck, W.R. McCombie, N. Dimitrova and L.N. Harris, 2013. Translating next generation sequencing to practice: opportunities and necessary steps. *Molecular Oncology*, 7(4): 743-755.
22. Dong, L., W. Wang, A. Li, R. Kansal, Y. Chen, H. Chen and X. Li, 2015. Clinical next generation sequencing for precision medicine in cancer. *Current Genomics*, 16(4): 253-263.
23. Grada, A. and K. Weinbrecht, 2013. Next-generation sequencing: methodology and application. *The Journal of Investigative Dermatology*, 133(8): e11.
24. Nekrutenko, A. and J. Taylor, 2012. Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. *Nature Reviews Genetics*, 13(9): 667.
25. Weinstein, J.N., E.A. Collisson, G.B. Mills, K.R.M. Shaw, B.A. Ozenberger, K. Ellrott, I. Shmulevich, C. Sander, J.M. Stuart and Cancer Genome Atlas Research Network, 2013. The cancer genome atlas pan-cancer analysis project. *Nature Genetics*, 45(10): 1113.

26. Nielsen, R., J.S. Paul, A. Albrechtsen and Y.S. Song, 2011. Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics*, 12(6): 443.
27. Sjöblom, T., S. Jones, L.D. Wood, D.W. Parsons, J. Lin, T.D. Barber, D. Mandelker, R.J. Leary, J. Ptak, N. Silliman and S. Szabo, 2006. The consensus coding sequences of human breast and colorectal cancers. *Science*, 314(5797): 268-274.
28. Gray, S.W., K. Hicks-Courant, A. Cronin, B.J. Rollins and J.C. Weeks, 2014. Physicians' attitudes about multiplex tumor genomic testing. *Journal of Clinical Oncology*, 32(13): 1317.
29. Xue, Y., A. Ankala, W.R. Wilcox and M.R. Hegde, 2015. Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. *Genetics in Medicine*, 17(6): 444-451.
30. Tothill, R.W., J. Li, L. Mileskin, K. Doig, T. Siganakis, P. Cowin, A. Fellowes, T. Semple, S. Fox, K. Byron and A. Kowalczyk, 2013. Massively-parallel sequencing assists the diagnosis and guided treatment of cancers of unknown primary. *The Journal of Pathology*, 231(4): 413-423.
31. Weiss, M.M., B. Van Der Zwaag, J.D. Jongbloed, M.J. Vogel, H.T. Brüggewirth, R.H. Lekanne Deprez, O. Mook, C.A. Ruivenkamp, M.A. Van Slegtenhorst, A. Van Den Wijngaard and Q. Waisfisz, 2013. Best Practice Guidelines for the Use of Next-Generation Sequencing Applications in Genome Diagnostics: A National Collaborative Study of Dutch Genome Diagnostic Laboratories. *Human Mutation*, 34(10): 1313-1321.
32. Gonzaga-Jauregui, C., J.R. Lupski and R.A. Gibbs, 2012. Human genome sequencing in health and disease. *Annual Review of Medicine*, 63: 35-61.
33. Rehm, H.L., S.J. Bale, P. Bayrak-Toydemir, J.S. Berg, K.K. Brown, J.L. Deignan, M.J. Friez, B.H. Funke, M.R. Hegde and E.Lyon, 2013. ACMG clinical laboratory standards for next-generation sequencing. *Genetics in Medicine*, 15(9): 733.
34. Zhang, W., Y. Yu, F. Hertwig, J. Thierry-Mieg, W. Zhang, D. Thierry-Mieg, J. Wang, C. Furlanello, V. Devanarayan, J. Cheng and Y. Deng, 2015. Comparison of RNA-seq and microarray-based models for clinical endpoint prediction. *Genome Biology*, 16(1): 133.
35. Byron, S.A., K.R. Van Keuren-Jensen, D.M. Engelthaler, J.D. Carpten and D.W. Craig, 2016. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nature Reviews Genetics*, 17(5): 257.
36. Capobianchi, M.R., E. Giombini and G. Rozera, 2013. Next-generation sequencing technology in clinical virology. *Clinical Microbiology and Infection*, 19(1): 15-22.
37. Pimentel, F., P. Bonilla, Y.G. Ravishankar, A. Contag, N. Gopal, S. La Cour, T. Lee and A. Niemz, 2015. Technology in microRNA profiling: circulating microRNAs as noninvasive cancer biomarkers in breast cancer. *Journal of Laboratory Automation*, 20(5): 574-588.
38. Rodenhiser, D. and M. Mann, 2006. Epigenetics and human disease: translating basic biology into clinical applications. *Cmaj*, 174(3): 341-348.
39. Van Allen, E.M., N. Wagle, A. Sucker, D.J. Treacy, C.M. Johannessen, E.M. Goetz, C.S. Place, A. Taylor-Weiner, S. Whittaker, G.V. Kryukov and E. Hodis, 2014. The genetic landscape of clinical resistance to RAF inhibition in metastatic melanoma. *Cancer Discovery*, 4(1): 94-109.
40. Nahta, R., D. Yu, M.C. Hung, G.N. Hortobagyi and F.J. Esteva, 2006. Mechanisms of disease: understanding resistance to HER2-targeted therapy in human breast cancer. *Nature Reviews Clinical Oncology*, 3(5): 269.
41. Meric-Bernstam, F., G.M. Frampton, J. Ferrer-Lozano, R. Yelensky, J.A. Pérez-Fidalgo, Y. Wang, G.A. Palmer, J.S. Ross, V.A. Miller, X. Su and P. Eroles, 2014. Concordance of genomic alterations between primary and recurrent breast cancer. *Molecular Cancer Therapeutics*, 13(5): 1382-1389.
42. Yap, T.A., M. Gerlinger, P.A. Futreal, L. Pusztai and C. Swanton, 2012. Intratumor heterogeneity: seeing the wood for the trees. *Science Translational Medicine*, 4(127): 127ps10-127ps10.