Applications

DEBORAH JOSKO

LEARNING OBJECTIVES

1. Discuss the various applications available on select next generation sequencing platforms.
2. Describe targeted, exome, transcriptome, microbial, *de novo* sequencing, etc.
3. List the assays that are FDA cleared for diagnostic use.
4. Examine the ACMG clinical laboratory standards for next generation sequencing.

ABBREVIATIONS: ACMG - American College of Medical Genetics and Genomics; ATP - adenosine triphosphate; CLIA - Clinical Laboratory Improvement Amendments; CTFR - cystic fibrosis transmembrane conductance regulator; DNA - deoxyribonucleic acid; FDA - Food and Drug Administration; HLA - human leukocyte antigen; HUGO - human genome organization; ID - identification; NGS - next generation sequencing; PGM - Personal Genome Machine; PT - proficiency testing; RNA - ribonucleic acid; rRNA - ribosomal ribonucleic acid; SMRT - single molecule real time; SNP - single nucleotide polymorphism; ZMW - zero mode waveguides.

INDEX TERMS: Next Generation Sequencing; Ion Torrent PCM™; Illumina MiSeq; SMRT Technology; Sequencing Applications; Guidelines; Standards.

Clin Lab Sci 2014;27(3):179

Deborah Josko, Ph.D. MLT(ASCP)M, SM, *Associate Professor and Program Director, Department of Clinical Laboratory Sciences, Medical Laboratory Science Program, Rutgers, The State University of New Jersey – School of Health Related Professions, Scotch Plains, NJ.*

Address for correspondence: Deborah Josko, Ph.D. MLT(ASCP)M, SM, *Associate Professor, Department of Clinical Laboratory Sciences, Medical Laboratory Science Program, Rutgers, The State University of New Jersey – School of Health Related Professions 1776 Raritan Road, Scotch Plains, NJ 07076. (908) 889-2422. daj100@shrp.rutgers.edu.*

Introduction

The evolution of deoxyribonucleic acid (DNA) sequencing techniques along with the advent of various sequencing platforms has revolutionized the field of genomics, research and medicine. Each manufacturer markets its own platform utilizing unique methodologies from the Ion Torrent PGM™ which sequences DNA using a semi-conductor chip[1] to single molecule real time (SMRT) technology by Pacific Biosciences for carrying out sequencing using zero mode waveguides (ZMW).[2] Both principles were discussed in detail in the article entitled "Next Generation Sequencing - Platforms."

Most sequencers have the ability to sequence the entire genome or small/targeted regions for research and clinical purposes. All sequencers perform an array of assays depending on the platform and applications available for that system. This article provides a brief overview of some of the applications available for research and clinical use and discusses the newly published clinical laboratory standards for next generation sequencing.

Next Generation Sequencing Applications: Ion Torrent PGM™

The Ion Torrent PGM™ by Life Technologies is able to perform targeted sequencing, exome sequencing, transcriptome sequencing and genome sequencing.[3] For targeted sequencing, a variety of panels are available that are "ready to use." However, one can also custom design their own panel based on the genomic sequences of interest.

According to the manufacturer's website, there are four "ready to use" panels available that include the "Ion AmpliSeq™ Cancer Hotspot Panel v2 which targets "hot spot" regions of 50 oncogenes and tumor suppressor genes; the Ion AmpliSeq™ Comprehensive Cancer Panel which targets greater than 400 oncogenes

and tumor suppressor genes; the Ion AmpliSeq™ Inherited Disease Panel targets exons of over 300 genes associated with over 700 inherited diseases including neuromuscular, cardiovascular, developmental, and metabolic diseases; and the Ion AmpliSeq™ Sample ID Panel which is a human single nucleotide polymorphism (SNP) genotyping panel that allows generation of unique identification codes for research samples in sequencing runs providing increased confidence during data analysis."[4] Table 1 summarizes some of the features of the Ion AmpliSeq™ Cancer Hotspot Panel v2,[5] the Ion AmpliSeq™ Comprehensive Cancer Panel,[6] and the Ion AmpliSeq™ Inherited Disease Panel.[7] In addition, consumers can custom make their own panel by using a free online design tool. All one needs to know is the human genome organization (HUGO) symbol or the gene regions based on human genome coordinates.[8]

Exome sequencing involves the sequencing of exomes (the gene regions that code for proteins) as opposed to sequencing the entire genome. It is estimated 1-2% of the genome actually codes for protein and approximately 85% of the disease-causing variants can be found in this coding region making exome sequencing a valuable tool in personalized medicine.[9,10] Transcriptome sequencing or ribonucleic acid (RNA) sequencing allows scientists to study all the RNA in a cell (transcriptome) in order to determine gene expression (which genes are turned on or off at any given time as well as gene function).[11] It is also used in identifying genetic variants that are responsible for causing disease such as cancer and metabolic disorders.[11] Although the coding regions of DNA have been of interest in the determination of disease states since the completion of the Human Genome Project, it has been found that the non-coding regions (regions that do not code for a protein and constitute 98% of the genome) are just as important and significant and may play a role in mutagenesis resulting in cancer and other disorders.[12,13,14] For a comprehensive overview of non-coding RNAs and their significance in disease conditions, refer to the article by Hwang et. al. entitled "Non-coding RNAs and Diseases."[15]

Microbial sequencing is an extremely important tool for the microbiologist with 16s rRNA sequencing directing taxonomic reorganization of microorganisms.[16] In addition, microbial sequencing is useful in surveillance outbreaks and determining etiology of the disease.[17] Another application of the Ion Torrent PGM™ is its ability to perform microbial sequencing and determine the strain associated with a particular outbreak. For example, Rhode et. al. published an article in the New England Journal of Medicine regarding the use of the Ion Torrent PGM™ and its ability to identify the strain belonging to the entero-aggregative *E. coli* lineage in an outbreak in Germany in 2011.[18] For more information regarding the global collaborative to identify this strain refer to the Application Note Shiga toxin-producing *Es-*

Table 1. Comparison of the "Ion AmpliSeq™ Cancer Hotspot Panel v2 (CHPv2),[5] the Ion AmpliSeq™ Comprehensive Cancer Panel (CCP),[6] and the Ion AmpliSeq™ Inherited Disease Panel (IDP).[7]

Parameters	CHPv2	CCP	IDP
Targets	Hotspot regions, including~2800 COSMIC mutations of 50 oncogenes and tumor suppressor genes, with wide coverage of the KRAS, BRAF, and EGFR genes	Exons within ~400 oncogenes and tumor suppressor genes	Exons of over 300 genes associated with >700 inherited diseases, including neuromuscular, metabolic, cardiovascular, and developmental diseases
Amplicon Length	111-187 bp average 154 bp	125-175 bp average 155 bp	125-225 bp average 197 bp
Primer Pool size	207 primer pairs in 1 tube	~16,000 primers in 4 tubes	>10,000 primer pairs in 3 tubes
Input DNA	10 ng/DNA sample	10 ng per pool 40 ng/DNA sample	10 ng per pool 30 ng/DNA sample
Required Time To results	10 hours (DNA to annotated variants)	16 hours	12 hours

cherichia coli by Life Technologies.[19]

In summary, the Ion Torrent PGM™ offers an array of applications that can be found on the manufacturer's website (www.lifetechnologies.com).[3] Although this instrument and the panels are not FDA cleared, it is being used in the clinical setting especially in identifying mutations on the cystic fibrosis transmembrane conductance regulator (CTFR) gene (ATP-binding cassette sub-family C, member 7).[20] In a recent study performed by Abou Tayoun et. al., the Ion Torrent PGM™ was used to identify mutations and single nucleotide variants on the CTFR gene. Based on their results the authors state, "with continued optimization and system refinements, PGM sequencing promises to be a powerful, rapid, and scalable means of clinical diagnostic sequencing."[20]

Illumina MiSeq

The Illumina MiSeq platform, a fully automated system, is designed to sequence small genomes and perform targeted gene sequencing. The following applications are available which can be found on the manufacturer's website: targeted gene sequencing, metagenomics, small genome sequencing, targeted gene expression, amplicon sequencing, and HLA typing.[21] According to the National Institute of Health - Genetics Home Reference Glossary, metagenomics is defined as "the study of a collection of genetic material (genomes) from a mixed community of organisms. Metagenomics usually refers to the study of microbial communities."[22] By studying and sequencing the 16S rRNA gene and analyzing targeted areas in this gene, scientists have the ability to identify species of a multitude of microorganisms in complex microbial populations; determine mutations and antimicrobial resistance patterns; and reclassify microorganisms taxonomically based on their genotypic not phenotypic characteristics as done in the past.[23,24]

Although the various platforms discussed have the ability to sequence a variety of genes and determine mutations and variants, only the Illumina MiSeqDx Cystic Fibrosis Clinical Sequencing Assay and the Illumina MiSeqDx Cystic Fibrosis 139-Variant Assay have been FDA cleared for diagnostic use.[25] There are no set guidelines for next generation sequencing; however, the American College of Medical Genetics and Genomics (ACMG) recently published practice guidelines for clinical standards for next generation sequencing in 2013 and are discussed in this article.[26]

SMRT Technology by Pacific Biosciences

Sequencers marketed by Pacific Biosciences utilize SMRT technology to perform a variety of applications. The principle for this platform was discussed in detail in the article entitled, "Next Generation Sequencing - Platforms." According to the manufacturer, "this system has the ability to observe structural and cell type variation in real time which is not available or offered by the other sequencing platforms making this system quite unique."[27] Applications include: *de novo* genome assembly, targeted sequencing, base modification detection and for use in microbial research.[27]

In order to fully comprehend the significance of a microbial genome and its implication in disease, the entire sequence must be known. However, this can be difficult with some of the limitations seen in current NGS platforms (for example the inability to resolve large structural variations).[28] In order to overcome some of these limitations, *de novo* genome sequencing can be used. *De novo* sequencing involves the sequencing of a genome for the first time without the use of a reference. The DNA fragments are pieced together/assembled by a computer program/algorithm to determine the genome of a particular microbe.[29] This process is quite complex; therefore, refer to the article published in Nature Methods (2012) by Baker entitled, "*De novo* genome assembly: what every biologist should know" for a comprehensive overview.[29]

Another application this platform offers is detection of DNA base modifications through SMRT sequencing. Detecting modifications such as methylation is a key component in understanding biological functions when studying gene expression, communication between host and pathogen, and DNA damage and repair.[30] This system can detect damaged DNA bases in the template affected by chemical modifications such as oxidation, radiation, hydrolysis, etc. which can aid in understanding of the effects of DNA damage in conditions such as neurodegenerative disorders, cancer and the aging process.[31] To understand the significance of identifying these base modifications and their effect on biological systems, refer to the article by Clark et. al. entitled, "Direct Detection and Sequencing of Damaged DNA Bases."[31]

In summary, NGS platforms have a variety of capabilities and applications whether sequencing the entire genome or targeted regions as described. Although the Illumina MiSeqDx Cystic Fibrosis Clinical Sequencing Assay and the Illumina MiSeqDx Cystic Fibrosis 139-Variant Assay are the only assays that have been FDA cleared for diagnostic use to date[25] concern exists over the standardization and validation process when using this technology in the clinical laboratory. Therefore, the American College of Medical Genetics and Genomics published practice guidelines for clinical standards for next generation sequencing in an attempt to standardize the use of sequencing in the clinical setting.[26] An overview and summary of these guidelines are presented in the subsequent section.

Practice Guidelines for Clinical Standards for Next Generation Sequencing

Along with rapidly developing NGS platforms and the applications offered comes the challenge of proper validation and standardization when using this technology in the clinical setting. NGS has revolutionized the way clinicians are treating patients and determining at-risk individuals for a particular disease. To date these platforms and panels are not FDA cleared for diagnostic use except for the Illumina MiSeqDx Cystic Fibrosis Clinical Sequencing Assay and the Illumina MiSeqDx Cystic Fibrosis 139-Variant Assay.[25] Therefore, the American College of Medical Genetics and Genomics (ACMG) has developed guidelines and standards to assist laboratorians with validation, interpretation and reporting of variants.[26] This section provides a brief overview of the guidelines published. For detailed information, the document can be found online at http://www.nature.com/gim/journal/v15/n9/full/gim201392a.html.[32]

The ACMG examined three main areas in NGS: disease-targeted gene panels, exome sequencing, and genome sequencing. The authors compared and contrasted the analytical and clinical sensitivity for all three applications and described what each type of sequencing covers.[26] Procedures in sample preparation, library generation, barcoding, target enrichment, sequencing platforms and data analysis in next generation sequencing are provided to familiarize the reader with the various steps NGS entails.[26] Test ordering is discussed along with a notation of the importance of ordering specific panels that limit the content of the test and focus on the regions where mutations are present for a particular disease. Targeted sequencing panels produce higher analytical sensitivity and specificity as compared to exome and genome sequencing due to the small area that is being sequenced. The authors also describe factors that must be considered when a laboratory is considering development of NGS services such as "costs, analytical sensitivity and specificity, and analysis complexity." All of these elements must be taken into consideration when investigating whether to use disease-targeted gene panels, exome sequencing or whole genome sequencing."[26]

The ACMG guidelines depict a flow diagram of the steps involved in test development and the validation process ending in quality management. For example, test development should include optimization of the entire test such as "general assay conditions, coverage, sample pooling and analysis setting/thresholds."[26] Test validation should include all sample types analyzed as well as "sensitivity, specificity, robustness and reproducibility."[26] Lastly, quality management must comprise quality control procedures for every run and periodic proficiency testing (PT) must be performed.[26] Although PT services are not yet available for NGS, laboratories are required to perform PT testing according to CLIA guidelines using national programs if available.[26] For detailed information regarding the testing, validation and quality management of NGS, refer to the ACMG practice guidelines.

An in-depth section is also included that describes reporting standards such as "turnaround times, data interpretation, reporting of incidental findings, the written report, and data reanalysis." The guidelines appendix includes sample medical reports using a pan cardiomyopathy panel (51 genes) and exome sequencing demonstrating what should be incorporated in both a positive and negative report.[26] The reports are very detailed and include a full explanation of: "DNA variants; interpretation summary; recommendation; individual variant interpretations (if positive); variants of unlikely clinical significance; test background; test method; limitations; and references."[26] This resource provides a detailed explanation on reporting standards and what must be included in the written report. Clinical laboratorians who are using NGS for clinical diagnosis should follow these guidelines and download

this document for future reference.

Summary

Next generation sequencing platforms and the applications that are offered have revolutionized the way a physician will treat and monitor a patient based on the individual's own genetic make-up. Whether whole genome sequencing, exome sequencing, or targeted sequencing is performed, the information generated must be analyzed, interpreted, and reported correctly. Since the various platforms and application panels are not FDA cleared (with the exception of the Illumina MiSeqDx Cystic Fibrosis Clinical Sequencing Assay and the Illumina MiSeqDx Cystic Fibrosis 139-Variant Assay)[25] clinical laboratorians are faced with the challenge of standardizing and validating the various panels and platforms for appropriate quality management. Therefore, the American College of Medical Genetics and Genomics published guidelines for ordering, test development, validation and reporting of genetic information.[26] These guidelines should be followed by all laboratorians performing NGS to ensure quality results and to provide proper interpretation of all genomic variants identified.

NOTE: The author is not endorsing any particular company or product and has no financial gain or otherwise interest in the products presented.

REFERENCES

1. Ion Torrent™ Next-Generation Sequencing Technology. Available from http://www.lifetechnologies.com/us/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-technology.html. Accessed 2014 March 31.
2. SMRT Cells. Available from http://www.pacificbiosciences.com/products/consumables/SMRT-cells/. Accessed 2014 March 31.
3. Ion Torrent™ Next-Generation Sequencing Applications. Available from http://www.lifetechnologies.com/us/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-applications.html. Accessed 2014 March 31.
4. Ion AmpliSeq™ Ready-to-Use Panels. Available from http://www.lifetechnologies.com/us/en/home/life-science/sequencing/dna-sequencing/targeted-sequencing/targeted-sequencing-ion-torrent-next-generation-sequencing.html. 2014 March 31.
5. Ion AmpliSeq™ Cancer Hotspot Panel v2. Available from http://tools.lifetechnologies.com/content/sfs/brochures/Ion-AmpliSeq-Cancer-Hotspot-Panel-Flyer.pdf. Accessed 2014 March 31.
6. Ion Ampliseq™ Comprehensive Cancer Panel. Available from https://tools.lifetechnologies.com/content/sfs/brochures/Ion_CompCancerPanel_Flyer.pdf. Accessed 2014 March 31.
7. Ion AmpliSeq™ Inherited Disease Panel. Available from http://tools.lifetechnologies.com/content/sfs/brochures/CO25400_Ion_InheritDisease_Flyer_17July2012.pdf. Accessed 2014 March 31.
8. Ion AmpliSeq™ Custom DNA Panels. Available from http://www.lifetechnologies.com/us/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-workflow/ion-torrent-next-generation-sequencing-select-targets/ampliseq-target-selection/custom-panels.html. Accessed 2014 March 31.
9. Exome Sequencing by Ion Torrent Next-Generation Sequencing. Available from http://www.lifetechnologies.com/us/en/home/life-science/sequencing/dna-sequencing/exome-sequencing/exome-sequencing-ion-torrent-next-generation-sequencing.html#enrich. Accessed 2014 April 1.
10. Clinical Exome Sequencing | UCLA Clinical Genomics Center. Available from http://pathology.ucla.edu/body.cfm?id=393. Accessed 2014 April 1.
11. Transcriptome. Available from http://www.genome.gov/13014330. Accessed 2014 April 4.
12. 'Junk' DNA Reveals New Disease-causing Mutations. Available from http://www.biosciencetechnology.com/news/2013/10/junk-dna-reveals-new-disease-causing-mutations. Accessed 2014 April 4.
13. Annotating non-coding regions of the genome. Available from http://www.geboc.org/genetic_progress/reference/Annotation_non-cod.pdf. Accessed 2014 April 4.
14. Sequencing and Non-coding RNA Program. Available from http://www.mdanderson.org/education-and-research/resources-for-professionals/scientific-resources/core-facilities-and-services/ncrna-program/index.html. Accessed 2014 April 4.
15. Huang Y, Wang JP, Yu XL, Wang ZB, Xu TS, Cheng XC. Non-coding RNAs and Diseases. Mol Bio. 2013;47:465-75.
16. DNA Sequencing. Available from http://www.microbialid.com/services/dna.html. Accessed 2014 April 4.
17. Bacterial typing. Available from http://www.lifetechnologies.com/us/en/home/life-science/sequencing/dna-sequencing/microbial-sequencing/microbial-identification-ion-torrent-next-generation-sequencing/bacterial-typing-ion-torrent-next-generation-sequencing.html#pubs. Accessed 2014 April 4.
18. Rohde H, Qin J, Cui Y, Li D, Loman NJ, Hentschke M, et. al. Open-Source Genomic Analysis of Shiga-Toxin–Producing *E. coli* O104:H4. N Engl J Med 2011;365:718-24.
19. Application Note – Shiga-toxin producing *Escherichia coli*. Available from https://tools.lifetechnologies.com/content/sfs/brochures/Small-Genome-Rapid-Characterization-Ecoli.pdf. Accessed 2014 April 4.
20. Abou Tayoun AN, Tunkey CD, Pugh TJ, Ross T, Shah M, Lee CC, et. al. A Comprehensive Assay for CFTR Mutational Analysis Using Next-Generation Sequencing. Clin Chem 2013;59:1481-8.
21. MiSeq Desktop Sequencer. Available from http://www.illumina.com/systems/miseq.ilmn. Accessed 2014 April 9.
22. Metagenomics. Available from http://ghr.nlm.nih.gov/glossary=metagenomics. Accessed 2014 April 9.
23. High-Speed, Multiplexed 16S Amplicon Sequencing on the MiSeq ® System. Available from http://res.illumina.com/documents/products/appnotes/appnote_16s_sequencing.pdf. Accessed 2014 April 9.

24. Liu Z, DeSantis TZ, Andersen GL, Knight R. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. Nucl Acids Res. 2008:36:18.
25. List of Human Genetic Tests. Available from http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm330711.htm. Accessed 2014 April 9.
26. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS. Brown KK, Deignan JL, et. al. ACMG clinical laboratory standards for next-generation sequencing. Genet Med. 2013;15:733-47.
27. Pacific Biosciences Overview. Available from http://www.pacificbiosciences.com/applications/overview/index.html. Accessed 2014 April 13.
28. De Novo Assembly. Available from http://www.pacificbiosciences.com/applications/denovo/index.html. Accessed 2014 April 13.
29. Baker M. *De novo* genome assembly: what every biologist should know. Nat Methods 2012;9:333-7.
30. Base Modification Detection. Available from http://www.pacificbiosciences.com/applications/base_modification/index.html. Accessed 2014 April 13.
31. Clark TA, Spittle KE, Turner SW, Korlach J. Direct Detection and Sequencing of Damaged DNA Bases. Genome Integrity 2011;2:10.
32. ACMG clinical laboratory standards for next-generation sequencing. Available from http://www.nature.com/gim/journal/v15/n9/full/gim201392a.html. Accessed 2014 March 31.