Chapter 27
Impact of Human Exome Sequencing on Clinical Research

Anu Acharya
Mapmygenome India, India

Jasmine Khurana
Mapmygenome India, India

Shibichakravarthy Kannan
Ocimum Biosolutions, India

Sushma Patil
Mapmygenome India, India

Brajendra Kumar
Ocimum Biosolutions, India

Geethanjali Tanikella
Mapmygenome India, India

ABSTRACT

Recent advances in human exome sequencing and the associated advantages have made it a technology of choice in various domains. The savings in time, cost and data storage compared with whole genome sequencing make this technology a potential game changer in clinical research settings. Recent advances in NGS have made it feasible to use exome sequencing in clinical research for identifying novel and rare variants that can lead to change in protein structure and function which may finally culminate into a totally different phenotype. If whole exome is not desired the same technology can be used for studying target exonic regions to investigate causative genes for a specific phenotype associated with disease. Exome sequencing has emerged as an effective and efficient tool for the translational and clinical research. There is a demand for systematically storing variant information in large databanks. Meaningful information from the exome-seq data can be combined with other data. This can be correlated with clinical findings within a clinical trial setting for a better study outcome.

DOI: 10.4018/978-1-5225-2237-9.ch027

INTRODUCTION

A Brief History of NGS

The major breakthrough in the field of genetics was due to the Human Genome Project. Completion of this project was possible because of the revolutionary Sanger's sequencing method. Sanger's chain termination method and Maxim-Gilbert's sequencing method have laid the foundation for advanced sequencing techniques such as next generation sequencing, leading to progress in genomics and proteomics. Sanger's sequencing method utilized the basic principle of DNA replication, which dictates that for template elongation, 3' -OH of the previous nucleotide should be free. Therefore, in the reaction mixture scientists introduced dideoxynucleotide, which lacked free 3' -OH group. Thus, the chain is terminated whenever dideoxynucleotide is incorporated, giving rise to templates of various basepair lengths. At that time, there was no procedure available to differentiate between all four nucleotides. Therefore, four different mixtures had to be prepared containing one of ddATP, ddGTP, ddCTP, ddTTP and other three remaining nucleotides, polymerase enzyme and radio-labeled primer. After PCR reaction, all the four reactions were loaded in different wells of agarose gel and depending on the size, strands were separated (ABI, 2007). Agarose gel had its limitation with the formation of ssDNA loops. Therefore, denaturing polyacrylamide gel electrophoresis was used instead. Use of radioactively labeled primer and four different reactions made it laborious. Thus, non-radioactive based first generation sequencing came into picture. Leroy Hood's laboratory made some modification of Sanger's method in mid 1980s by introducing fluorescent dyes instead of radioactively labeled primers (ABI, 2007). This made it possible for all the reactions to take place in one reaction mix. Further, the method was simplified by introduction of capillary electrophoresis, thus, making the system more flexible for increasing the sample size and significantly sped up the analysis (Zhou et al., 2010a). This resulted in higher instrument throughput with 96 samples on ABI 3730 platform and 384 samples on Amersham MegaBACE in one run (Zhou et al., 2010a).These advances have helped in completing the Human Genome Project in record time, way ahead of schedule. These first -generation sequencers can achieve sequencing length up to 1000 bp, with raw accuracy as high as 99.999%, at a cost as little as $0.50/Kilobase and throughput close to 600000 bp/day (Zhou et al., 2010a). Even with this advancement, it became necessary to invent new methods to reduce cost and increase speed of analysis. Next Generation Sequencing technology was developed to address these issues (Zhou et al., 2010b).

Exome Capture Kits

Selection of appropriate target capture kit is important for qualitative and quantitative identification of variants in the target regions in human exome (protein coding regions). The quality analysis and association of correct variants to the human diseases is a crucial objective of exome sequencing in clinical research.

Depending on sequencing platform and specific application, it is important to opt for appropriate exome capture kit based on the coverage of different exome database, target coverage efficiency, GC bias, sensitivity in single nucleotide variant detection, sensitivity in small indel detection, and technical reproducibility (Chilamakuri et al., 2014; Parla et al., 2011).

Some popular commercially available exome capture kits in use by researchers and clinically certified labs are: