Farasani A, J Genet Genomic Sci 2017, 2: 006
DOI: 10.24966/GGS-2485/100006

HSOA Journal of Genetics & Genomic Sciences

Review Article

# Importance of Exome Sequencing in the Human Diseases and Medical Genetics

**Abdullah Farasani***

*College of Applied Medical Sciences, Jazan University, Jazan, Saudi Arabia*

## Abstract

Exome sequencing has rapidly squeezed the speed for various newly identified diseases genes/variants and build very important role in the human genetics. Till now several studies have confirmed the whole exome sequencing as a diagnostic tool for the documented molecular defects with and without deferred genetic diseases and disorders. The accurate methodology with the exome and next generation sequencings can be identified the complete exons (exome sequencing) and next generation sequencing to sequence DNA at unique speed, permitting formerly unimaginable scientific achievements and novel biological applications. This review describes about the important role of exome and next generation sequencings in human genes in the global population. Maximum studies have documented affecting role of variants in the specific population but this may not vary in the global population. This could be due to the variability in the ethnicity. Novel identified variants from similar diseases may or may not document in the global population. Hence this study concludes exome and next generation sequencings can be implemented in all the human diseases which is and not affected with hereditary and medical genetics. This review suggests applying the techniques exome sequencing and next generation sequencing in the human diseases with all over the global population to confirm the disease biomarker for the clinical practices.

**Keywords:** Exome-sequencing; Human diseases; Medical genetics; Next-generation-sequencing

Exposing the genetic and inherited disparities on human health and diseases has become the major scientific challenges in 21st century. Human genome sequencing and Celera Genomics has published the different versions on human genome. Human Genome Project (HGP), an international project sequenced the human DNA and nucleotide sequences which recognizes and maps all the genes of the human genome. Till now 2.91 billion base pair consensus sequences of human data were generated with whole genome short gun sequencing method [1,2]. There are many genetic diseases are appeared as hereditary and non-hereditary diseases. Hereditary disease is defining as diseases produced by disturbances in storage, transformation and production of genetic information. Recent improvement in the human genetic diseases such as medical genetics, hereditary nature of many diseases and genetic syndromes of undetermined etiology has been blown. Hereditary disease is mainly occurred by abnormalities in chromosomal disorder and gene mutations. The hereditary diseases are broadly based on the involvement of metabolic disorders such as (i) amino acid metabolism (ex; phenylketonuria), (ii) lipid metabolism (ex; Gaucher's disease), (iii) Carbohydrate metabolism (ex; Fructosuria), (iv) mineral metabolism (ex; Wilsons hepatoreticular degeneration) and (v) bilirubin metabolism (ex; Crigler-Najjor syndrome) [3]. The number of human disease genes could be estimated. However, the number of different mutations that could potentially cause human inherited diseases are clearly almost limitless (ex; one was to include all possible frameshift microdeletions and micro insertions), the number of mutations actually in existence and available to be identified and characterized is a complex function of the mutability of each inherited disease gene, the prevalence and ease of ascertainment of the consequent clinical phenotypes, the demographic history of the human population, at our disposal to locate and identify the pathological mutations [4]. Genetic variation is a key biological determinant underpinning evolution and defining the heritable basis of phenotype. Gene function cannot be fully understood without awareness of the potential variability within the gene. Genetic variation takes many forms, but all forms originate from just two types of mutation event. The simplest type of variation results from a simple base substitution. This type of mutation event accounts for the commonest form of variation, the Single Nucleotide Polymorphism (SNP), but also rare mutations which may show Mendelian inheritance in families. Most of the other types of variation result directly or indirectly from the insertion or deletion of a section of DNA. At the simplest level, this can result in the insertion or deletion of one or more nucleotides, the so called Insertion/deletion (Indel) polymorphisms. The most common insertion/deletion events occur in repetitive sequence elements, where repeated nucleotide patterns, so called "Variable Number Tandem Repeat Polymorphisms" (VNTRs), expand or contract as a result of insertion or deletion events. VNTRs are further sub-divided on the basis of the size of the repeating unit; minisatellites are composed of repeat units ranging from ten to several hundred base pairs [5]. Simple tandem repeats (STRs or microsatellites) are composed of 2-6 bp repeat units. Insertion/ deletion events involving large regions ranging from a few kilobases to several megabases are known as Copy Number Variations (CNVs). These events may occur as a result of recombination between flanking repetitive elements. CNVs were once thought to be very rare, restricted to severe genomic syndromes; however, the initial sequencing of the human genome and more recently, studies of samples ascertained for the HapMap project have provided evidence that CNVs are actually commoner than previously expected with most individuals carrying substantial deletion or duplications of DNA possibly with little phenotypic impact [6]. Structural variants are all variants which are not single nucleotide variants, which

***Corresponding author:** Abdullah Farasani, College of Applied Medical Sciences, Jazan University, Jazan, Saudi Arabia, Tel: +966 504790292; E-mail: farasani@hotmail.co.uk

**Citation:** Farasani A (2017) Importance of Exome Sequencing in the Human Diseases and Medical Genetics. J Genet Genomic Sci 2: 006.

**Received:** August 13, 2017; **Accepted:** October 02, 2017; **Published:** October 17, 2017

include insertion-deletion, block substitutions, inversions of DNA sequences and copy number differences. Large-scale genotyping of common variants led to an avalanche of discoveries of variants associating with common and complex diseases. Current studies based on whole-genome and Exome Sequencing (ES) are beginning to yield rare variants associating with common diseases. They also provide unprecedented information about human sequence diversity and insights into the structure and history of human populations [7].

The systematic identification of rare alleles is associated with common traits typically requires resequencing instead of genotyping and has therefore been challenging. Such studies have largely been limited to assessing rare variants that have been found by the targeted sequencing of candidate genes or of genomic regions identified by linkage or Genome Wide Association Study (GWAS), assumption is that rare variants that influence a trait co localize with common variants that influence risk. Although, whole-genome sequencing will be performed on all subjects with in a disease cohort, this approach is currently costly. In the meantime, exome sequencing provides an opportunity to capture nearly all of the rare and very rare alleles in the protein-coding genes that are present in a sample, although the contribution of exome sequencing to our understanding of complex diseases has been much smaller than its contribution to our understanding of Mendelian traits. Exome sequencing is often used in conjunction with two sampling strategies: family-based phenotypes and extreme phenotypes. In families in which multiple individuals are affected with a common trait, one approach is to sequence the most distally related individuals: the more distantly related the individuals, the fewer genetic variants they share. However, even distantly related individuals share many variants that require further stratification to identify a potentially causal allele. An alternative, family-based approach, which is used to identify *de novo* variants, involves sequencing parent-offspring trios in which only the offspring is affected. This strategy has been used to identify candidate genes for several complex traits [8,9].

GWAS is defining as an investigation of the association between common genetic variation and disease. This type of analysis requires a dense set of markers that capture a substantial proportion of common variation across the genome, and large numbers of study subjects. Presently, DNA sequencing has become a major concerns of medical research. Most direct way to examine the human diseases identified in the genes are through genetic heterogeneity, compare the DNA sequences among specific diseases. Genomic DNA analysis was based on Sanger sequencing, which was found to be effort in identifying known variants existent in 10-20% of specific DNA. Watson and Crick, fundamental discovery for double structure of DNA and pioneering the expansion of method to detect DNA sequences (ATGC) through sanger sequencing, develops in capacity, capability and applications Sanger sequencing is defined, an approach in which dye-labelled Normal Deoxynucleotides (dNTPs) and dideoxy-modified dNTPs are mixed. Standard PCR reaction is carried out and, as elongation occurs, few strands incorporate a dideoxy-dNTP, thus terminating elongation. The strands are then separated on a gel and the terminal base label of each strand is identified by laser excitation and spectral emission analysis.

Benefit of DNA sequencing for genetic heterogeneity was melodramatically enhanced by approval of Next-Generation Sequencing (NGS) for large scale disease analysis [10,11]. NGS is also termed as high throughput sequencing/second generation sequencing and has been introduced as effective tool for genetic screening and identifying new disease-causing variants discovered by whole exome sequencing. This technique is cost effective and generates large number of DNA sequence variants for finding candidates for disease cause through large-scale analysis. The NGS technique is similar to capillary electrophoresis in which DNA polymerase catalyses the incorporation of fluorescently labeled dNTPs into the DNA template strands during sequencing cycles of DNA synthesis. However, NGS-based variant detections are prone to erroneous calls and generate low-interest variants in the form of genotype false-positives [12]. Whole Genome Sequencing (WGS) is termed as process of determining the complete analysis of DNA sequence of an organisms' genome at a single time. NGS allows scientists/researchers to explore genetic variations from human genome project, which was initially completed in 2003. Currently, any sequencing technique is much more efficient, and any human genome can be sequenced in a matter of days for under $10,000. The first human genome cost $2.7 billion. Today, most genetic testing focuses on one or a few genes, rather than the entire genome. However, with the falling cost of genome sequencing, more individuals are pursuing this option. Physicians can look at an entire genome to see how specific treatments for a disease will be affected by an individual's unique genetics. For example, the physician may opt to look at genes involved in drug metabolism when deciding dosage. In the future, WGS may enable everyone to develop a personalized treatment plan [13]. Sequencing of the exome, rather than the entire human genome, is well justified as an efficient strategy to search for alleles underlying rare Mendelian disorders. Earlier study from the group Ng et al., has confirmed affected gene for monogenetic disorder in the unrelated individuals [14]. Bamshad et al., has concluded as exome sequencing as a tool for the discovery of Mendelian disease gene [15]. Mendelian disorder is defined as phenotypes caused by a mutation/s in a single gene and inherited in a dominant, recessive or X-linked pattern. NGS has been widely using in the clinics for the diagnosis purposes [16]. Kim et al., has discovered the stop-gain mutation in TP53 gene in Langerhans cell sarcoma through WES. Both the NGS and ES identified promoter, coding and non-coding variations in maximum all the diseases/disorders [17]. Several studies have confirmed both the NGS and ES has high successful rate in the clinical and medical genetics. The ES and NGS has already proven remarkably successful in confirming the cause of Mendelian diseases. Consequently, there is growing interest in the introduction of NGS into the clinic to aid in the diagnosis of conditions for which no genetic cause can be found with targeted testing or chromosomal arrays. In clinical setting, patients with undiagnosed genetic conditions tend to present with a wide range of clinical features, and is often necessary to consider each patient's genome individually, rather than looking for common disrupted genes in multiple cases with a similar phenotype [18].

From this review, conclusion can be confirmed as both exome and next generation sequencings can be implemented in all the human diseases which is and not affected with hereditary and medical genetics. The accurate variant can be confirmed in promoter regions such as 3′UTR regions, coding and non-coding regions of the human genome. The results may not vary from population to population due to the ethnicity. The racism will definitely play a major role in the hereditary and non-hereditary diseases of the different global population. The second generation sequencing techniques will enable us to identify all the variants in an individual's personal genome and, in particular, clinically relevant alleles. Apart from this, whole genome sequencing is also expected to bring a major shift in clinical practice

in terms of diagnosis and understanding of diseases, ultimately enabling personalized medicine based on one's genome. This review suggests to apply the exome sequencing and next generation sequencing in the human diseases with all over the global population to confirm the disease biomarker for the clinical practices.

## References

1. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. (2001) The sequence of the human genome. science 291: 1304-1351.
2. Frazer KA, Murray SS, Schork NJ, Topol EJ (2009) Human genetic variation and its contribution to complex traits. Nat Rev Genet 10: 241-251.
3. Francois J (1981) [Metabolic disorders and corneal changes (author's transl)]. Klin Monbl Augenheilkd 178: 419-423.
4. Cooper DN, Chen JM, Ball EV, Howells K, Mort M, et al. (2010) Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. Hum Mutat 31: 631-655.
5. Barnes MR (2003) Human genetic variation: databases and concepts. In: Barnes MR, Gray AC (eds.). Bioinformatics for Geneticists. Wiley Online Library, New York, USA.
6. Ng PC, Kirkness EF (2010) Whole genome sequencing. Methods Mol Biol 628: 215-226.
7. Gudbjartsson DF, Helgason H, Gudjonsson SA, Zink F, Oddson A, et al. (2015) Large-scale whole-genome sequencing of the Icelandic population. Nat Genet 47: 435-444.
8. Schork NJ, Murray SS, Frazer KA, Topol EJ (2009) Common vs. rare allele hypotheses for complex diseases. Curr Opin Genet Dev 19: 212-219.
9. Alkan C, Coe BP, Eichler EE (2011) Genome structural variation discovery and genotyping. Nat Rev Genet 12: 363-376.
10. Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17: 333-351.
11. Levy SE, Myers RM (2016) Advancements in next-generation sequencing. Annu Rev Genomics Hum Genet 17: 95-115.
12. Bogari NM (2016) Next Generation Sequencing (NGS) in Glucose-6-Phosphate Dehydrogenase (G6PD) deficiency studies. Bioinformation 12: 41-43.
13. Berglund EC, Kiialainen A, Syvänen AC (2011) Next-generation sequencing technologies and applications for human genetic history and forensics. Investig Genet 2: 23.
14. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, et al. (2009) Targeted capture and massively parallel sequencing of 12 human exomes. Nature 461: 272-276.
15. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, et al. (2011) Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 12: 745-755.
16. Blackburn PR, Selcen D, Gass JM, Jackson JL, Macklin S, et al. (2017) Whole exome sequencing of a patient with suspected mitochondrial myopathy reveals novel compound heterozygous variants in RYR1. Mol Genet Genomic Med 5: 295-302.
17. Kim SW, Choi MK, Han HS, Song H, Koh Y, et al. (2017) A Case of Pulmonary Langerhans Cell Sarcoma Simultaneously Diagnosed with Cutaneous Langerhans Cell Histiocytosis Studied by Whole-Exome Sequencing. Acta Haematol 138: 24-30.
18. Need AC, Shashi V, Hitomi Y, Schoch K, Shianna KV, et al. (2012) Clinical application of exome sequencing in undiagnosed genetic conditions. J Med Genet 49: 353-361.