

Next generation sequencing (NGS) in glucose-6-phosphate dehydrogenase (G6PD) deficiency studies

Neda M. Bogari

Faculty of Medicine, Department of Medical Genetics, Umm Al-Qura University, Makkah-24382, Kingdom of Saudi Arabia; Neda M Bogari - Email: NMbogari@uqu.edu.sa; Phone: +966555500150; *Corresponding author

Received March 05, 2016; Revised March 07, 2016; Accepted March 07, 2016; Published April 10, 2016

Abstract:

Glucose-6-phosphate dehydrogenase (G6PD) deficiency is commonly observed in human males. It is a genetic disorder affecting the red blood cells. The diagnosis of G6PD is usually based on blood analysis and there is no specific molecular or genetic test. The complete gene sequence of G6PD is known for different ethnicities. Known single nucleotide polymorphism (SNP) associated with G6PD is available in the public databases. Hence, robust, fast and efficient sequencing of G6PD is critical in disease diagnosis. The application of next generation sequencing (NGS) with its high reliability is useful in G6PD diagnosis.

Keywords G6PD, genetic disorder, mutation, polymorphism, sequencing analysis.

Background:

Glucose-6-phosphate dehydrogenase (G6PD) deficiency is an X-linked genetic disorder according to World Health Organization [1]. The G6PD enzyme in erythrocytes regenerates nicotinamide adenine dinucleotide phosphate (NADPH) and subsequently produces glutathione (GSH). This protects red cells against oxidative attacks. G6PD-deficient erythrocytes are particularly susceptible to hemolysis upon exposure to oxidants such as fava beans or primaquine. G6PD deficiency has been associated with hemolytic crises as a major clinical manifestation [2]. G6PD deficiency was first observed in people sensitive to the haemolytic effect of the anti-malarial drug primaquine [3, 4]. The individuals with G6PD susceptibility to hemolysis induced with anti-malarial drug have shown geographic correlation between G6PD deficiency prevalence and endemic malaria. However, the testing and distribution of anti-malarial drugs in endemic areas were influenced by estimates of the prevalence of G6PD deficiency [5]. The enzyme G6PD is known to protect red blood cells (RBCs) from the harmful effects of reactive oxygen species (ROS). G6PD gene mutations reduces enzyme levels and or alter its structure causing hemolytic anemia, as a result of ROS accumulation of [6]. Data for about 400 variants believed to be unique have been characterized on the basis of their biochemical behavior and are grouped into 5 classes according to the level of residual enzyme activity and clinical manifestations [7, 8]. The role of enzyme G6PD in RBC

involved in the oxidation of glucose and production of NADPH, the co-enzyme that protects RBC from oxidative stress is well documented [9].

Genetics and molecular diagnosis

G6PD has been mapped to sub-telomeric region of the human X-chromosome *i.e.* on Xq28 by cell hybrid studies [10] and it consists of 13 exons and 12 introns [11]. Mental retardation and fragile x syndromes are located at Xq27-28 regions and is responsible for an X-linked form. Thus, this *genetic loci* is of interest for studying G6PD deficiency. Moreover, blood coagulation factor IX is also assigned at the sub-telomeric region of the X chromosome long arm *i.e.* Xq27.1-2 [12]. There are many other mutations linked with the Xq28 region of the X chromosomes. G6PD deficiency is the most common enzymatic disorder of RBC in humans affecting people worldwide [13, 14]. The human X chromosome has a unique biology in evolution as the sex chromosome shared by both male and female [15]. X-linked genetic conditions are more likely to affect males than females because of its single copy in males. G6PD deficiency will manifest in females when both copies of the gene are defective in the genome. A recessive trait is observed when a certain heritable trait is expressed [16]. The band size of G6PD gene encoded by mRNA genes is 2,269 bps [17].

Our understanding of the complex genetic basis of human diseases has improved over the years [18]. However, there are several unknown factors in the understanding of many genetic diseases and disorders including G6PD deficiency. Gene identification for successful diagnosis, prevention, and treatment is critical [19]. Genome sequencing to characterize genetic variation in a specific population, determine haplotype structure used to impute alleles and boost the outcome of Genome-wide association studies (GWAS) [20]. G6PD deficiency was studied using mutation, polymorphic variants with and without amino acid substitution, restriction fragment length polymorphism, Alu Sequence, insertion/deletion and DNA sequencing. Sequence-tagged sites (short stretches of DNA) that can be specifically detected by the polymerase chain reaction are used as landmarks [21]. Nonetheless, no such genetic marker is explicitly documented till date. The success and efficiency of biochemical analysis is doubtful. Clinician and physicians depends on hematological tests to identify the disease. Therefore, it is of interest to use next generation sequencing (NGS) and exome sequencing techniques for SNP specific disease diagnosis.

Next generation sequencing

The Next generation sequencing (NGS) is also termed as high throughput sequencing/second generation sequencing and has been introduced as an 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 [22].

The huge amount of sequence data is rapidly generated by NGS in a very cost-effective way and research scientists starts to grab the advantage [23]. An enormous numbers of short reads produced by NGS provide opportunities for the development of novel applications that benefit from the particular data format. For example: NGS technologies have been widely applied in contexts whereby sequencing of only a portion of the molecule is sufficient [24]. GWAS using single nucleotide polymorphism (SNP)-arrays can change rapidly and sequencing will become the molecular microscope, although replication, transcription, translation, methylation and nuclear DNA folding are completely different processes and can be studied using sequencing [20]. NGS technique is used to identify promising biomarker candidates for early diagnosis; to detect the causative variants and guide targeted treatment decisions. The phenotypic changes can also be queried with NGS technique, which include a variety of epigenetic modifications of DNA that regulate gene expression [25]. This technique has been used to screen the Mendelian and non-mendelian diseases in the human genome. Earlier studies have confirmed that NGS has improved genetic testing in families with histories of high penetrance cancer genes such as *BRCA1*, *BRCA2*, *APC*, and *TP53*. Several investigators have tested the

Illumina Hi Seq platform in detecting *BRCA1*, *BRCA2*, and *TP53* from a tumor cell line [26]. However, NGS analysis identified all known variants in the tumor cell line with sensitivity and specificity greater than traditional diagnostic methods, demonstrating the effectiveness of NGS as a diagnostic tool [27]. The clinical applications of NGS are (a) genome wide discovery of casual variants (b) targeted sequence (c) RNA-sequencing and (d) epigenetic profiling. The concluding remarks of NGS have transported a paradigm shift in basic and important application technique in both rare and common human genome disorders to establish the diagnostic tool in clinical and forensic sciences to solve the specific cases which could not be possible with Sanger sequencing.

Conclusion:

The clinical usefulness of genomic sequencing requires advancement in our knowledge of the genome and Bioinformatics system to process genetic data advancements that would be built on a foundation of 100% genome sequencing [27]. The lack of known relationship between the G6PD deficiency and molecular studies were due to the non-recognition of exact gene function, which affects the human diseases. The view described here suggests implementing the NGS analysis in G6PD deficiency subjects to diagnose the common, rare and novel variants in rare disorders when there is no particular test to identify using molecular methodologies. Future studies would include single molecular sequencing, diagnostic yield and 100% genome sequencing for studying genetic disorders among human ethnicities and especially in G6PD deficiencies.

References:

- [1] Alharbi KK, *Saudi Med J*. 2015 **36**: 544 [PMID: 25935173]
- [2] Yang HC, *et al. J Inflamm (Lond)*. 2015 **12**: 34 [PMID: 25945076]
- [3] Beutler E, *Blood Rev*. 1996 **10**: 45 [PMID: 8861278]
- [4] Mehta A, *et al. Baillieres Best Pract Res Clin Haematol*. 2000 **13**: 21 [PMID: 10916676]
- [5] Nkhoma ET, *et al. Blood Cells Mol Dis*. 2009 **42**: 267 [PMID: 19233695]
- [6] Al-Musawi BM, *et al. BMC Blood Disord*. 2012 **12**: 4 [PMID: 22452742]
- [7] Alharbi KK & Khan IA, *J Int Med Res*. 2014 **42**: 1161 [PMID: 25169987]
- [8] Frigerio R, *et al. Haematologica*. 1994 **79**: 319 [PMID: 7806085]
- [9] Biscaglia S, *et al. J Atheroscler Thromb*. 2015 **22**: 535 [PMID: 25843116]
- [10] Szabo P, *et al. Proc Natl Acad Sci U S A*. 1984 **81**: 7855 [PMID: 6595664]
- [11] Alharbi KK, *et al. Bioinformation*. 2013 **9**: 49 [PMID: 23390344]
- [12] Alharbi KK, *et al. Bioinformation*. 2012 **8**: 1255 [PMID: 23275730]
- [13] Dogra N, *et al. Perfusion*. 2010 **25**: 417 [PMID: 20705643]
- [14] Noori-Dalooi MR, *et al. Journal of Sciences, Islamic Republic of Iran*. 2006 **17**:313
- [15] Graves JA. *Cell*. 2006 **124**: 901 [PMID: 16530039]
- [16] Barakat TS, *et al. Exp Cell Res*. 2010 **316**:679 [PMID: 20083102]
- [17] Scriver CR, *Pediatr Res*. 1995 **38**: 819 [PMID: 8618780]
- [18] Risch N, *et al. Science*. 1996 **273**: 1516 [PMID: 8801636]

- [19] Khan IA, *et al. Intractable Rare Dis Res.* 2016 **5**: 25 [DOI: 10.5582/irdr.2015.01040]
- [20] Buermans HP & den Dunnen JT, *Biochim Biophys Acta.* 2014 **1842**: 1932 [PMID: 24995601]
- [21] Green ED, *et al. Genomics.* 1991**11**: 548 [PMID: 1837788]
- [22] Park MH *et al. PLoS One.* 2014 **9**: e86664 [PMID: 24489763]
- [23] Ekblom R & Galindo J, *Heredity (Edinb).* 2011 **107**: 1 [PMID: 21139633]
- [24] Morozova O & Marra MA, *Genomics.* 2008 **92**: 255 [PMID: 18703132]
- [25] Kruglyak KM *et al. Adv Exp Med Biol.* 2016 **890**: 123 [PMID: 26703802]
- [26] Meldrum C, *et al, Clin Biochem Rev.* 2011 **32**: 177 [PMID: 22147957]
- [27] Shen T, *et al. Front Genet.* 2015 **6**: 215 [PMID: 26136771]

Edited by P Kanguane

Citation: Bogari, *Bioinformation* 12(2): 41-43 (2016)

License statement: This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License

