Screening of G6PD Mutation and polymorphism using high resolution melt curve analysis

A thesis submitted to the BRAC University in

partial fulfillment of requirements for the degree of Master of Science(M.S.) in

Biotechnology

Submitted By:

Shezote Talukder Shithi

ID: 15176001

Department of Mathematics and Natural Science

BRAC University

Mohakhali, Dhaka

Declaration of Authenticity

I, the undersigned, Shezote Talukder Shithi, declare that the research work embodying the results reported in this thesis entitled "Screening of G6PD Mutation and polymorphism using high resolution melt curve analysis", is my original work, gathered and utilized especially to fulfill the purposes and objectives of this study, and has not been previously submitted to any other institution for a higher degree or diploma. It is further declare that this work has been carried out under joint supervision of Professor Firdausi Qadri, Senior Scientist and Executive Director, Institute for developing Science and Health Initiatives (ideSHi), and Professor Naiyyum Choudhury, Biotechnology Program, Department of Mathematics and Natural Sciences (MNS), BRAC University. I also declare that the publications cited in this work have been personally consulted.

Shezote Talukder Shithi

Certified

Professor Dr. Firdausi Qadri

Supervisor

Senior Scientist and Executive Director, Institute for Developing Science and Health Initiatives

Director, Center for Vaccine Sciences
Senior Scientists and Head,
Immunology Laboratory
icddr,b, Dhaka

Professor Dr. Naiyyum Choudhury

Supervisor

Chairman Bangladesh Atomic Energy Regulation Authority

Professor, Department of Mathematics and Natural Sciences
BRAC University

Acknowledgement

First of all, I would like to thank the Almighty Allah for blessing me with the fortitude to carry out my research work successfully.

I would like to express my immense gratitude towards my supervisors Dr. Firdausi Qadri and Dr. Naiyyum Choudhury for providing me the opportunity to conduct my thesis project at the Institute for Developing Science and Health Initiatives (ideSHi) laboratory of the Institute of Public Health. I am also thankful to Professor A.A Ziaudddin Ahmed, Chairperson, Department of Mathematics and Natural Sciences for encouraging me with my studies. I am truly grateful towards my thesis advisor, Dr. Aparna Islam from Department of Mathematics and Natural Sciences of BRAC University. She made herself available whenever I faced any difficulty, or had queries about my research work or thesis; lending a helping hand and a guiding voice throughout the span of my course.

This dissertation would not have been crowned without the contiguity of my eminent instructor Dr. Kaiissar Mannoor who, as the leading scientist and lab head of ideSHi, gave me permission to conduct my research work with him and his team. He consistently motivated me to own the direction of my work, while supporting me in times of need. It has been both an honor and a comfort to have had his consistent support throughout this journey.

It is my honor to applaud my preceptor Suprovath Kumar Sarker, research fellow at ideSHi. Without his thoughtful encouragement and careful supervision the thesis would never have taken shape. I owe him for supporting me every moment during my whole working period at ideSHi.

I cannot fail to appreciate the careful input and rigorous mentorship of MD. Tarikul Islam, who stood with me through every step of the way. I thank him for diligently guiding my transition from an unsure graduate to becoming a confident, capable and independent researcher in my own right.

My utmost gratitude goes out to all the senior supervisors and Research Officers and all the staff of Institute for Developing Science and Health Initiatives (ideSHi), in particular Dr. Nusrat Sultana, Dr. Sazzadul Islam, Sarower Bhuyan and Assifuzzaman Rahat. They have readily made themselves available at every step of this project; aiding me in troubleshooting

problems, guiding the experimental process when necessary and serving as ever-present advisers.

I would also like to thank all my peers Dr. Aparna Biswas, FarhanaManzoor, Pritha Biswas and Shagoofa Rakhshanda at the Institute for Developing Science and Health Initiatives. They have made my placement highly enjoyable, all the while increasing my scientific understanding. Most importantly I would like to thank Saad Hassan Hasib for always being by my side, supporting me mentally and helping me in every way as much as he can.

Lastly, I am forever indebted to my family and friends for being patient with me throughout this journey, providing strength in hard times, and believing in me when I did not believe even in myself. Without their love and support this would not have been possible.

Author,

Shezote Talukder Shithi

Abstract

Glucose-6-phosphate dehydrogenase (G6PD) deficiency is one of the most common X-linked recessive disorders resulting from the defect of a single gene encoding an enzyme called G6PD. G6PD enzymes catalyzes the conversion of glucose-6 phosphate into 6-phosphoglucono-δ-lactone, which is a rate limiting step in pentose phosphate pathway (PPP).PPP supplies reducing energy to red blood cells (RBCs) by maintaining NADPH levels which in turn maintains reduced glutathione levels. Reduced glutathione protects RBCs against oxidative damage. This disorder can cause potentially life threatening hemolytic anemia under conditions of oxidative stress due to consumption of certain foods, drugs, or infections. In 2013, 4100 deaths occurred globally due to G6PD deficiency. The prevalence rate and the cut off values of G6PD in Bangladesh are not yet established. Our unpublished data generated through G6PD gene sequencing of G6PD deficient patients demonstrate that three different mutations, namely G6PD Mahidol (c.G487), Kalyan-Kerala (c.G949) and Orissa (c.C131G) are prevalent in Bangladesh. However, gene sequencing is a time consuming and costly procedure to identify the genetic basis of G6PD deficiency. To overcome the difficulties, a new method using real time PCR-based HRM analysis could successfully identify all mutant alleles based on different melting curve shapes. Six G6PD deficient samples were subjected to HRM analysis using four pairs of G6PD gene-specific primers. Among six mutant alleles, five were found to have a c.G487A (Mahidol) mutation and one had the c.G949A (Kalyan-Kerala) mutation and the results were consistent with sequencing data, confirming the validity of the procedure. The sensitivity and specificity of mutation detection was 100% for c.G487A and c.G949Amutations. For other six G6PD deficient specimens that had c.C131G mutation, PCR-HRM approach could not be applied due to lack of specific primers for that region. However, it is expected that the primer – specific real time PCR-HRM approach would definitely work for c.C131G mutation. The results also demonstrate the usefulness of real time PCR-HRM analysis for detection of heterozygous alleles for c.1311T/C and IVS-11_93t/c polymorphisms of G6PD gene. The findings elucidate a promising potential for use of G6PD primer-specific real time PCR-HRM analysis as a faster and cost effective approach to identify known G6PD mutations and polymorphisms that are prevalent in Bangladeshi population.

Contents

Title **Page**

List of Figures

List of Tables

List of Abbreviation

6PG	**6-phosphoglucono-δ-lactone**
AHA	**Acute hemolysis anemia**
Ala	**Alanine**
BIRDEM	**Bangladesh Institute of Research and Rehabilitation for Diabetes, Endocrine and Metabolic Disorders**
BLAST	**Basic Local Alignment Search Tool**
BMRC	**Bangladesh Medical Research Council**
BMD	**Becker muscular dystrophy**
bp	**Base pair**
BSMMU	**Bangabandhu Sheikh Mujib Medical University**
CNSHA	**Chronic non-spherocytic hemolytic anemia**
CO2	**Carbon dioxide**
C_T	**Threshold cycle**
Da	**Dalton**
DBS	**Dried blood spot**
DMSO	**Dimethyl sulfoxide**
DNA	**Deoxyribonucleic acid**
dNTP	**Deoxynucleoside triphosphate**
dsDNA	**Double stranded DNA**
EDTA	**Ethylenediaminetetraacetic acid**
FD	**Fabry disease**
Fe2+	**Ferrous state of iron**
Fe3+	**Ferric state of iron**
FST	**Fluorescent spot test**
G6PD	**Glucose-6-phosphate dehydrogenase**
gDNA	**Genomic DNA**
Glu	**Glutamic acid**
Gly	**Glycine**
GSH	**Reduced Glutathione**
GSSG	**Glutathone disulfide**
H2O2	**Hydrogen peroxide**
Hb	**Hemoglobin**
HBA	**Hemoglobin subunit alfa(α)**
HBB	**Hemoglobin subunit beta(β)**
HRM	**High resolution melt curve**
IVS	**Intervening sequence**
Kb	**Kilobase**
Lys	**Lysine**
MEC	**Maleria Endemic Country**
MetHb	**Methemoglobin**
MgCl2	**Magnesium chloride**
miRNA	**MicroRNA**
M	**Molar**
mL	**Milliliter**
mM	**Millimolar**
µL	**Microliter**
µmol	**Micromolar**
mRNA	**Messenger RNA**

NADP+	**Nicotinamide adenine dinucleotide phosphate**
NADPH	**Reduced nicotinamide adenine dinucleotide phosphate**
NCBI	**National Center for Biotechnology Information**
ng	**Nanogram**
nm	**Nanomolar**
NO	**Nitric oxide**
O2	**Oxygen**
O2-	**Superoxide anion**
PCR	**Polymerase chain reaction**
PPP	**Pentose phosphate pathway**
RBC	**Red blood cell**
RNA	**Ribonucleic acid**
ROS	**Reactive oxygen species**
Ser	**Serine**
SOD	**Superoxide dismutase**
SMA	**Spinal muscular atrophy**
SMN	**Survival motor neuron**
SNP	**Single nucleotide polymorphism**
Tm	**Melting temperature**
UV	**Ultraviolet**
WHO	**World Health Organization**

Chapter 1: Introduction

1.1 Background of the study

Glucose-6-phosphate dehydrogenase (G6PD) deficiency is the most common enzyme deficiency in humans affecting 400 million people worldwide (Nkhoma, Poole et al. 2009). It has a high prevalence in persons of African, Asian, and Mediterranean descent. It is inherited as an X-linked recessive disorder. G6PD deficiency is polymorphic, with more than 300 variants (Glucose-6-Phosphate Dehydrogenase (G6PD) Deficiency, Paul Schick). Most of the variants appear sporadically as single amino acid defects within a protein of 515 amino acids.

G6PD enzyme is involved in first step of Glucose-6-phosphate catabolism in pentose phosphate shunt. It plays a vital role in various physiological processes such as erythropoiesis, antioxidant defense, and vascularization (Bonilla, Sánchez et al. 2007). The G6PD enzyme oxidizes glucose-6-phosphate and reduces NADP+ to NADPH, which is important to maintain the cellular reduced glutathione level to scavenge dangerous oxidative metabolites and to retain proper 3D structure of Hb during oxidative stress. The pentose monophosphate shunt is the only source of supplying NADPH to red blood cells. In person with G6PD deficiency, reactive oxygen species production during oxidative stress causes denaturation of hemoglobin (Hb) and subsequent Heinz bodies' formation of denatured Hb leads to intravascular hemolysis.

G6PD deficiency is generally asymptomatic in adult life and only manifest during conditions that lead to oxidative stress such as consumption of fava beans, intake of primaquine for malaria treatment and infection with microorganism. This enzyme deficiency confers partial protection against malaria , which probably accounts for high prevalence of G6PD deficiency in malaria endemic areas. However, G6PD deficiency can present itself as neonatal hyperbilirubinemia and elevated bilirubin level for short period of time in deficient neonates can lead to kernicterus and irreversible brain damage. In addition, individuals with this disorder can experience episodes of hemolysis during oxidative stress. Severe form of G6PD deficiency also makes deficient susceptible to chronic granulomatous disease (CGD). Therefore G6PD deficiency disorder should be detected early phase of life so that the deficient people may consult to a genetic counselor to avoid the life threatening complications.

However in Bangladesh, the prevalence rate is not yet known and there is no established cut-off value. Also underlying genetic causes were not investigated prior to this study. G6PD deficient samples were identified by conventional qualitative fluorescent spot test and quantitative spectrophotometric enzyme assay. Underlying gene mutations for G6PD deficiency were identified by G6PD gene sequencing. As conventional methods can't detect G6PD deficiency in certain cases; such as after blood transfusion, during hemolytic episodes; it was the target to show the usefulness of molecular technique (High Resolution Melt Curve Analysis) for detection of common G6PD mutations in Bangladesh even in such conditions. Also it was a target to predict heterozygous mutations using HRM which is not predictable by conventional methods.

1.2 Structure and chromosomal location of G6PD gene (Genetics of G6PD gene)

The human G6PD gene is located in the sub-telomeric region of the long arm of the X chromosome (band Xq28), where several other genes have been mapped, constituting the so-called G6PD cluster (Human Gene Mapping, 1985). This region is of special interest in genetic terms because of a relatively high rate of crossing over at or near an adjacent fragile site (Martini, Toniolo et al. 1986) . G6PD-enzyme is encoded by a human X-linked gene (Xq2.8) consisting of 13 exons which expans 18kb and 12 introns, spanning nearly 20 kb in total (Martini, Toniolo et al. 1986) ; the first exon is non coding, while the remaining 12 range from 120 to 236 bp. An intron is present in the 5-prime untranslated region.

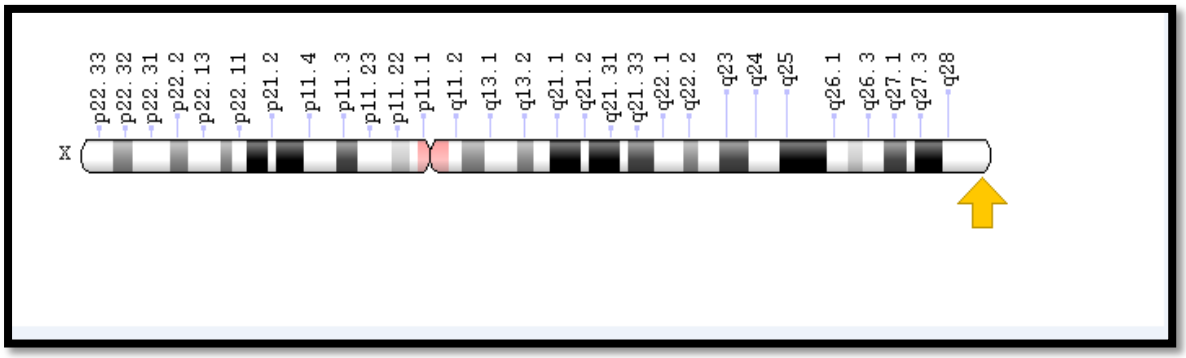

Fig 1.1: Schematic diagram of X-chromosome showing Xq28 region and molecular location of G6PD gene. The G6PD gene located at the end of the long arm (q) of X-chromosome at 28th position and is denoted by the yellow arrow.

The major 5-prime end of mature G6PD mRNA in several cell lines is located 177 bp upstream of the translation-initiating codon. Comparison of the promoter region of G6PD and 10 other housekeeping enzyme genes confirmed the presence of common features. In

particular, in 8 cases in which a 'TATA' box was present, a conserved sequence of 25 bp was seen immediately downstream. Its genomic co-ordination is (GRCH38): X: 154, 531, 389-154,547,585 (NCBI). The G6PD gene is part of the 20,114 bp nucleotide sequences that also contain prominent a CpG island which starting about 680 nucleotides upstream of the transcription initiation site, extending about 1,050 nucleotides downstream of the initiation site, and ending at the start of the first intron. The transcribed region from the initiation site to the poly(A) addition site covered 15,860 bp (Chen, Cheng et al. 1991).

Table 1.1: Cytogenic location and molecular characterization of G6PD gene

DNA	Localization	Xq2.8
	Gene size (in kilobasis)	18.5
	Number of exons	13
	Number of introns	12
mRNA	Size(in nucleotides)	2269
	Transcript variant-1 size	2406 bp
	Transcript variant-2 size	2295 bp
Protein	Number of amino acids	515
	Molecular weight (in Daltons)	59,265
	Subunits by molecule of active enzyme	2 or 4

From study of radiation-induced segregants (irradiated human cells 'rescued' by fusion with hamster cells), showed that the order of 4 loci on the X chromosome is PGK: alpha-GAL: HPRT: G6PD and that the 3 intervals between these 4 loci are, in relative terms, 0.33, 0.30, and 0.23(Goss and Harris 1977). Studying X-autosome translocations in somatic cell hybrids, showed that a breakpoint at the junction of Xq27-q28 separates HPRT from G6PD (Pai, Sprenkle et al. 1980). The G6PD gene, which is transcribed in the telomeric direction, partly overlaps the IKBKG, which is transcribed in the centromeric direction. The genes

share a conserved promoter region that has bidirectional housekeeping activity. In addition, intron 2 of the G6PD gene contains an alternate promoter for the IKBKG gene only determined that the region containing the G6PD gene and the 5-prime end of the IKBKG gene contains Alu elements (Fusco, Paciolla et al. 2012).

1.3 The G6PD Enzyme

Glucose-6-phosphate dehydrogenase enzyme plays a vital role in human metabolism by maintaining RBC integrating. This protects the cell from oxidative attack by radicals derived from oxygen and organic compounds such as drugs and their metabolites. The G6PD enzyme consists of either dimer or tetramer forms of a protein subunit consisting of 514 amino acids(Kotaka, Gover et al. 2005). Each subunit then binds to an NADP+ molecule for structural ability. The majority of the mutations in G6PD occurs due to the disruption of the enzyme structural stability and thus reduces its overall activity. The effect of each mutation on enzyme structure and function depends on the location of the substituted amino acid. These mutations cause the most severe clinical symptoms for glucose-6-phosphate dehydrogenase deficiency disease. Mutations which do not cause any severe reduction in enzyme activity may have effect on genes coding region (fig 1.2) and sometimes on enzyme structure and have been found to reduce the efficiency of protein folding.

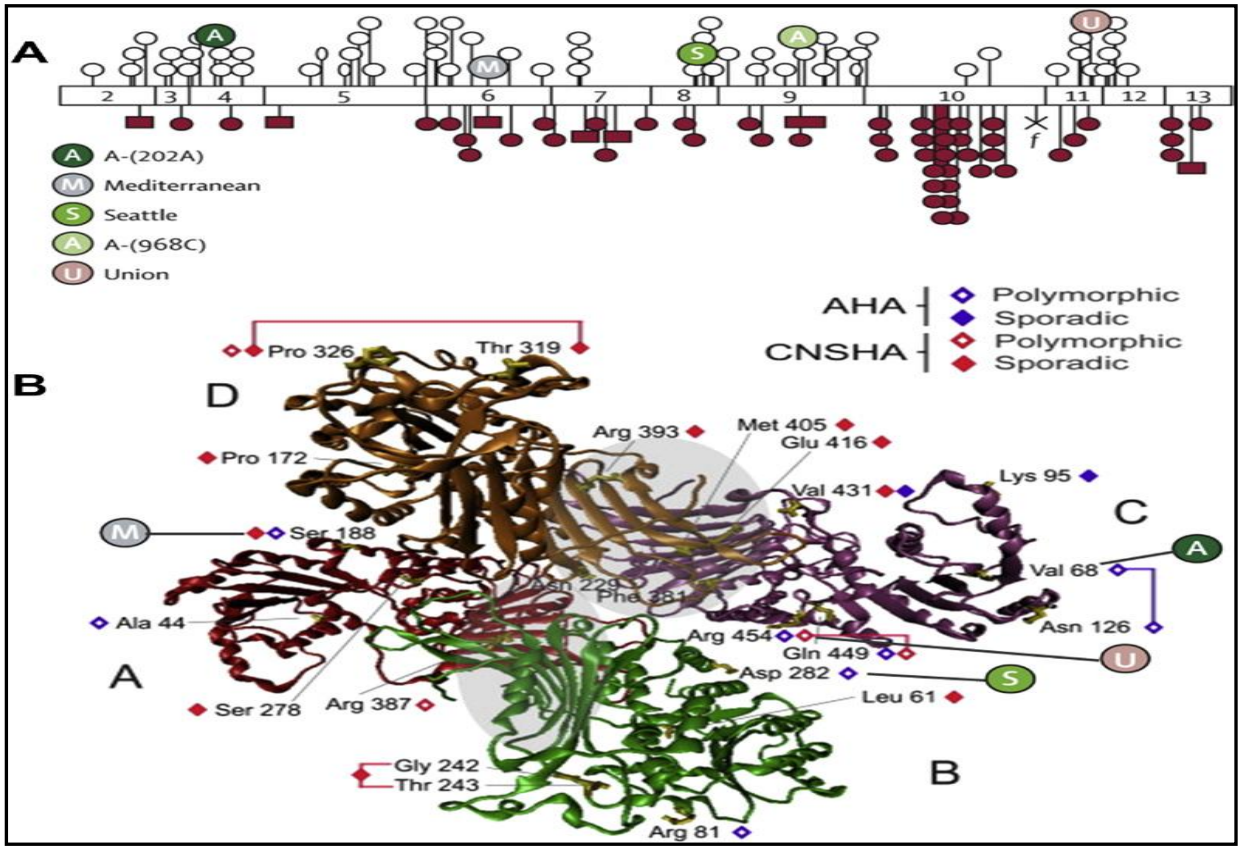

Figure 1.2: Diversity of mutations in the G6PD gene and enzyme. Panel A shows the distribution of common mutations along the G6PD gene coding sequence. Exons are shown as open numbered boxes. Open circles are mutations causing Class II and III variants; filled circles are Class I variants; filled squares are small deletions; the cross represents a nonsense mutation; "f" shows a splice site mutation. Panel B shows the distribution of amino acid substitutions across the enzyme's tetrameric structure (each identical monomer subunit is labelled A–D), numbered according to the affected amino acids. The diamonds indicate polymorphic or sporadic mutations, and their colour shows the associated clinical phenotype. The grey shadowed areas cover the two dimer interfaces. Across this region, a molecule of structural NADP per monomer is buried which stabilises the monomers and the associations between them (Figure reprinted from Luzzatto and Notaro (2001))

Like other enzymes G6PD enzyme decrease with the cell age and it is estimated that in normal blood cells, reticulocytes have five times higher activity level than the oldest 10% of RBCs. (Luzzatto 2006). Therefore, the oldest cells are more vulnerable to oxidative stress. In individuals with reduced enzyme activity duo to genetic mutation the ageing process of RBCs become faster with larger proportion of cells having lower enzyme level and thus these individuals are at risk of oxidative damage. The properties of these enzymes variants correspond to a wide spectrum of enzyme biochemical phenotypes.

1.4 Clinical importance of G6PD

1.4.1 Elimination of reactive oxygen species in RBC

Glucose-6-Phosphate dehydrogenase enzyme has several functions in human physiology. The first and most important is the elimination of reactive oxygen species in RBC. One of the reactive oxygen species is hydrogen peroxide which is produced in stressful condition. Hydrogen peroxide destroy the RBCs (fig 1.5) by affecting the elasticity of the membrane protein skeleton resulting oxidative stress on biophysical properties (Hale, Winlove et al. 2011). Hence G6PD enzyme is important to keep red blood cells flawless by producing glutathione peroxidase through pentose phosphate pathway. Glutathione peroxidase using reduced glutathione eliminates hydrogen peroxide which is achieved by the function of glutathione reductase enzyme. Glutathione reductase activity requires NADPH produced by the G6PD activity and pentose phosphate pathway (Beutler, Duparc et al. 2007).The pentose phosphate pathway and the role of G6PD in this pathway is highlighted in figure 1.3.

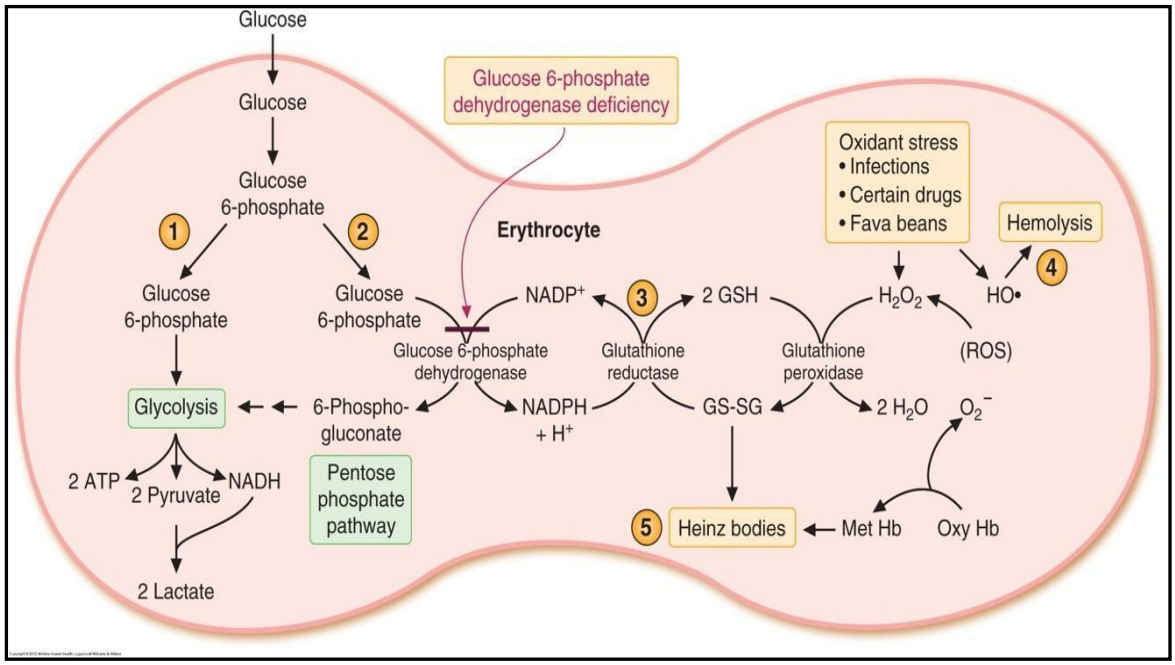

Figure 1.3: Pentose phosphate pathway, redox metabolism in RBC and glutathione metabolism in the elimination of ROS in an individual with normal G6PD activity. 6PG = 6-phosphoglucono-δ-lactone, NADP = Nicotinamide adenine dinucleotide phosphate, NADPH = Reduced form of NADP, Hb = Hemoglobin, MetHb = Methemoglobin, CO2 = Carbon dioxide, Fe3+ and Fe2+ is the ferric and ferrous state of iron.

G6PD is a housekeeping gene, expressed in all cell of the body that catalyzes the oxidation of glucose-6-phosphate (G6P) to 6-phosphoglucono-δ-lactone, which is then hydrolyzed to 6-phosphoglucono-v-lactone this, in turn, through the action of the enzyme phosphogluconate

dehydrogenase (6PGD), is further oxidized and decarboxylated to the pentose sugar ribulose 5-phosphate. Both G6PD and 6PGD have NADP as coenzyme, and therefore 2 molecules of NADPH are formed per molecule of G6P oxidized by G6PD (shown in figure 1.3).

G6PD functions to reduce NADP while oxidizing glucose-6-phosphate. Glucose-6-phosphate dehydrogenase plays a key role in the production of ribose 5-phosphate and the generation of NADPH in the hexose monophosphate pathway. Because this pathway is the only NADPH-generation process in mature red cells, which lack the citric acid cycle, a genetic deficiency of G6PD is often associated with adverse physiologic effects (Takizawa, Huang et al. 1986).

In the way the enzyme provides a source of reducing power that maintains sulfhydryl groups and thus aid in the detoxification free radicals and peroxides. It was said before that G6PD provides the only means of generating NADPH, and in its absence the erythrocyte is particularly vulnerable to oxidative damage (fig 1.4). Although the same G6PD gene is expressed in all tissues, the effect of a deficiency of the enzyme is most sever in erythrocytes, possibly because of their long non nucleated life span, and perhaps because they contain proteases that are more likely to degrade the mutant enzyme than are the proteases of other tissues. Anemia is therefore the most frequent manifestation G6PD deficiency. Impaired leukocyte function has also been documented but is rare.((Mamlok, Mamiok et al. 1987). Enzyme activity in the liver seems to be reduced considerably, at least in some patients. In patients with neonatal jaundice or viral hepatitis, the enzyme activity of the liver may also have a major role in causing hyperbilirubinemia.

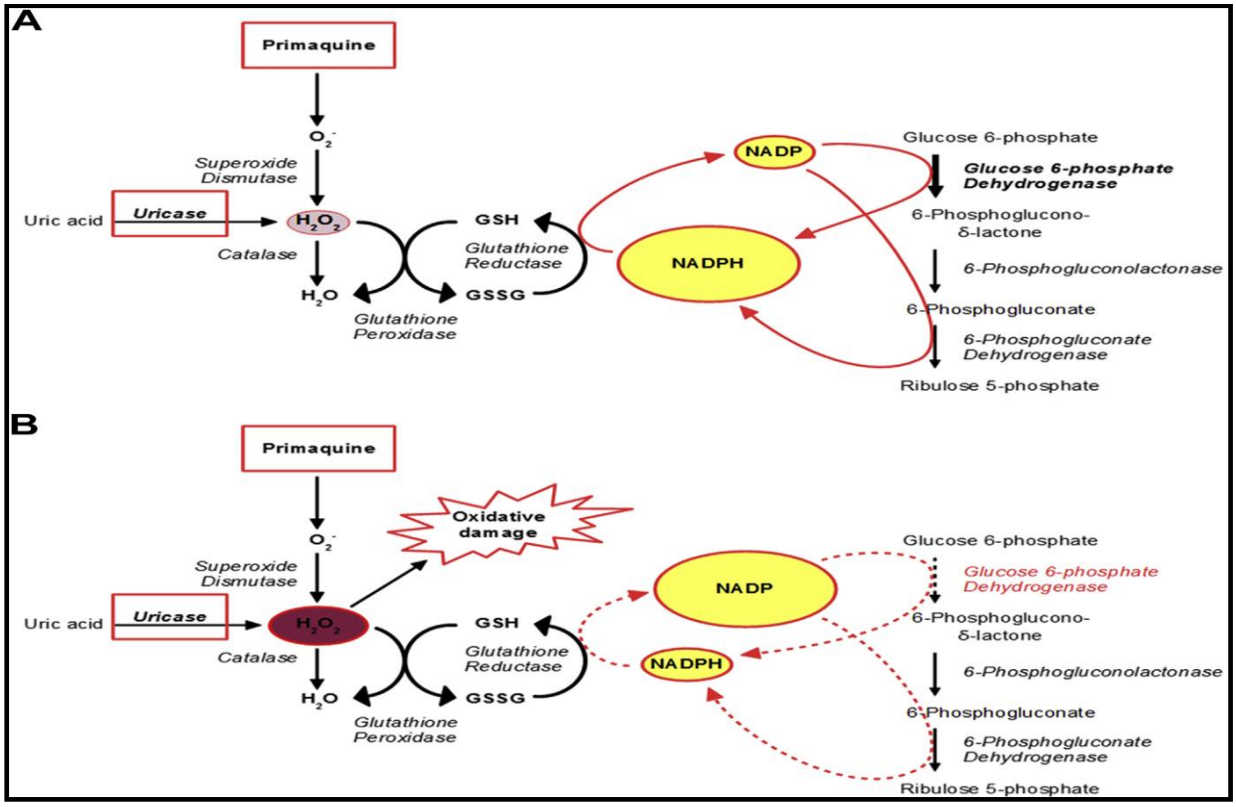

Figure 1.4: Role of G6PD in protection against oxidative damage. (A) In G6PD-normal red cells, pentose phosphate pathway—provide ample supply of NADPH, which in turn regenerates GSH when this is oxidized by reactive oxygen species (e.g, O_2^- andH_2O_2). (B) In G6PD-deficient red cells, where the enzyme activity is reduced, NADPH production is limited and it may not be sufficient to cope with the excess of reactive oxygen species generated in the presence of pro oxidant compounds.

1.4.2 Binding of hemoglobin to oxygen

Methemoglobin (MetHb) is a form of the hemoglobin (Hb) which is an oxygen-carrying metalloprotein and the iron in the heme group of MetHb is in the ferric (Fe3+) state, not the ferrous (Fe2+) of normal hemoglobin. In contrast to oxyhemoglobin, MetHb cannot bind oxygen. However, NADPH is needed for the activity of the methemoglobin reductase (diaphorase I). Methemoglobin reductase uses NADPH to reduce the Fe3+ in methemoglobin to Fe2+, thereby allowing hemoglobin to bind and then, transport oxygen.

1.4.3 Cellular growth

Many scientific journals stated the importance of G6PD in cellular growth by providing NADPH for redox regulation**.** The reason is that increased G6PD activity under normal and neoplastic cell growth is to provide ribose-5-phosphate for nucleic acid synthesis. Rribose-5-phosphate is the extreme critical product of G6PD activation and it was seen that using a

fibroblast cell lines in which G6PD activity had been inhibited and found that addition of RNA (a well-utilized physiological source of ribose-5-phosphate), did not restore cell growth (Stanton 2012). It also demonstrated that both ribose-5-phosphate and NADPH are required for cell growth and pentose phosphate pathway is the only means of providing them to cells.

1.4.4 Regulation of the activity of the KU protein

KU protein is involved in the repairing of DNA double strand breaks after the damage caused by radiation or exposure to thiol-specific oxidant, hydroxyethyldisulfide. The intervention of the G6PD is affected through the pentose cycle and consists of facilitating the union of KU homodimer with reduced cysteine residues to the DNA in repairing process (Ayene, Stamato et al. 2002)

1.5 Pathophysiology of G6PD deficiency

A variety of other pathologic effects have been attributed to G6PD deficiency, including changes in blood pressure and in increased incidence of cancer, seizure disorders, and mental diseases and diabetes. It has also been suggested that G6PD deficiency may ameliorate the severity of sickle cell disease, but this finding may well have been due to population stratification. Recent studies have large numbers of patients have not revealed any effect of G6PD deficiency on the course of sickle cell disease (ERNEST BEUTLER, MD).The involvement of human G6PD deficiency with hemolysis has been the main concerned for years. Hemolysis may be caused due to the exposure to some foods and administered of some drugs or infections. In G6PD deficient people, the apparition of Heinz bodies and their inefficiency to protect the GSH which is used to convert hydrogen peroxide and organic hydroperoxides (very reactive compounds) into stable compounds. Lack of GSH means GSH cannot protect the cell from the reactive oxygen species, which damage the cell membrane by oxidizing the protein and lipid constituents, as a result peroxide build up. As peroxides build up, cell membrane integrity and flexibility is reduced hemolysis. NADPH also protects proteins such as hemoglobin from oxidation by protecting the sulphydryl goups. Heinz bodies (intracellular hemoglobin precipitates) form from oxidized, denatured hemoglobin which is no longer soluble.

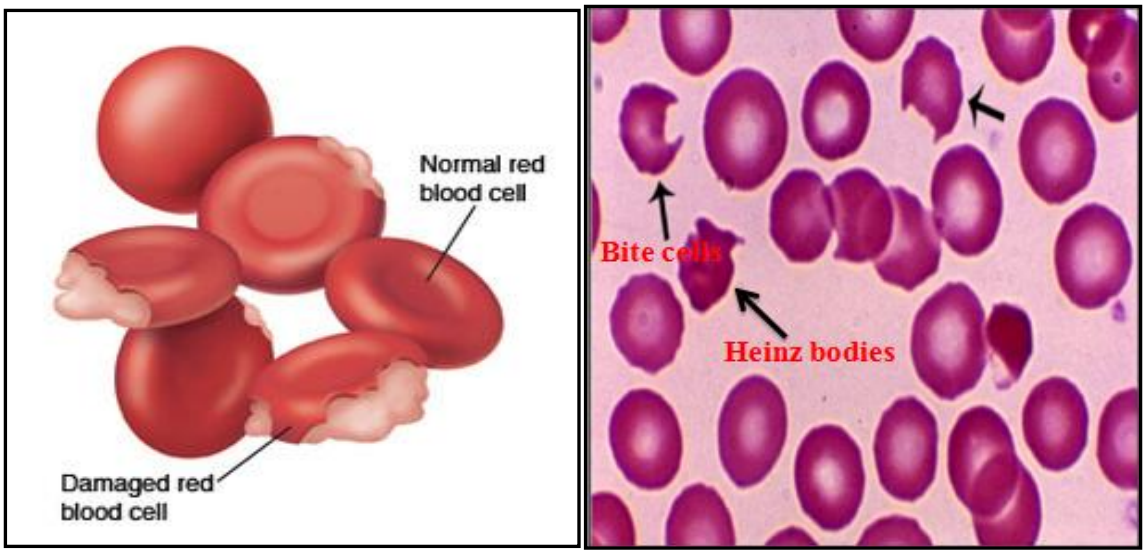

Figure 1.5: Hemolysis in G6PD deficiency due to exposure to drugs and compounds that result in production of reactive oxygen species and formation of Heinz bodies within RBC. Black arrows showing cells with Heinz bodies and red arrows showing RBCs that have undergone membrane damage due to Heinz bodies attachment to the cell membrane and consequent loss of portion of membrane in the spleen and liver to phagocytic cells.

These particles of denatured protein structures remain adhered to the cell membrane by disulfide bonds, cross-linking of sulfhydryl groups on globin chains and disrupting the normal red cell membrane structure. An ensuing loss of a portion of the membrane occurs as the damaged red cells that contain Heinz bodies are phagocytosed and sequestered by macrophages in the spleen and taken out of the circulation (Allahverdiyev et al., 2012) (Figure 1.5).

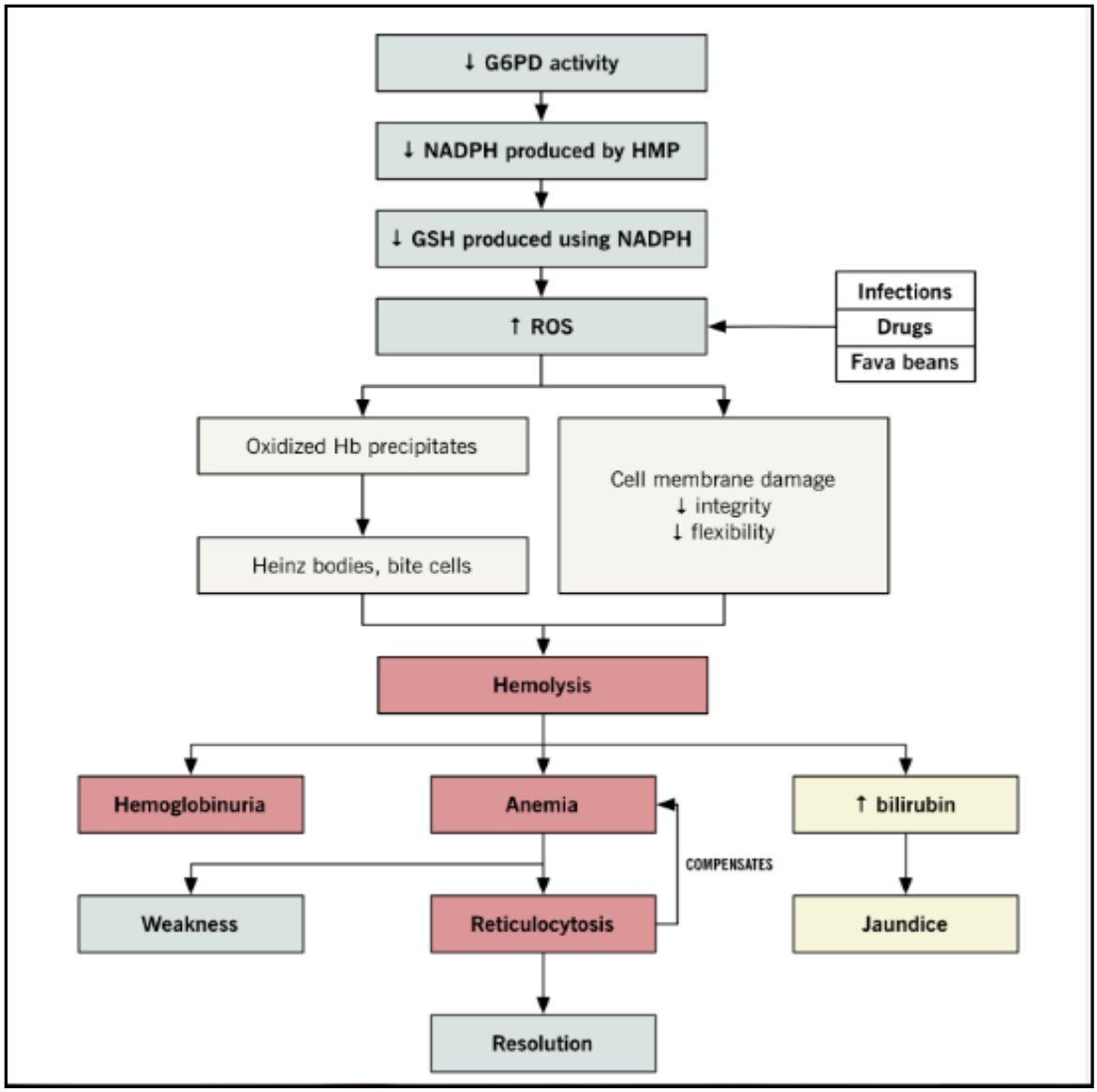

Figure 1.6: Pathophysiology and clinical features of G6PD deficiency. G6PD = Glucose-6-phosphate dehydrogenase, HMP = hexose monophosphate**,** NADPH = reduced form of Nicotinamide adenine dinucleotide phosphate, GSH = reduced glutathione, ROS = Reactive oxygen species (Figure reprinted from http://www.pathophys.org/g6pd/)

The hemoglobin is then metabolized to bilirubin which is further lead to hyperbilirubinemia and jaundice. The hemoglobin is rarely excreted through urine as well as kidney, but in severe case it can be happened.

1.6 Inheritance pattern of G6PD

G6PD is an X-linked recessive disorder, with an inheritance pattern similar to that hemophilia and color blindness: men usually manifest the abnormality and females are carriers. Females may be symptomatic if they are homozygous or inactivation of their normal X chromosome occurs. It requires both the X-chromosome to be defective to make a girl G6PD deficient. It cannot be passed from one person to another in any other way. Males can either be G6PD deficient or unaffected. Female can be G6PD deficient, partially deficient or unaffected. The condition is transmitted by mutations in the G6PD gene located on the X-chromosome and follows an X-linked recessive inheritance pattern which is shown in table 1.2

Table 1.2: Inheritance pattern of G6PD

Mother	Father Unaffected	Father Affected
Unaffected	All children unaffected	All boys unaffected All girls partially deficient
Partially deficient	Boys 50%chance affected Girls 50% chance partially deficient	Boys 50% chance affected Girls 50% chance of being partially deficient or fully deficient
Fully deficient	Boys all deficient Girls all partially deficient	Boys all deficient Girls all fully deficient

If the father is carrying the defective X-chromosome, he is not going to pass it to his son. The mother delivers X-chromosome to her offspring no matter the child is either male or female. So it can be said that if a child inherits defective X-chromosome, it is from his mother.

When the mother is normal and the father is affected, contributing the non-functioning X-chromosome to his daughter, daughters will always be the carrier of X-linked disease. All the sons remain unaffected in this case and the possible conditions are shown in figure1.8.

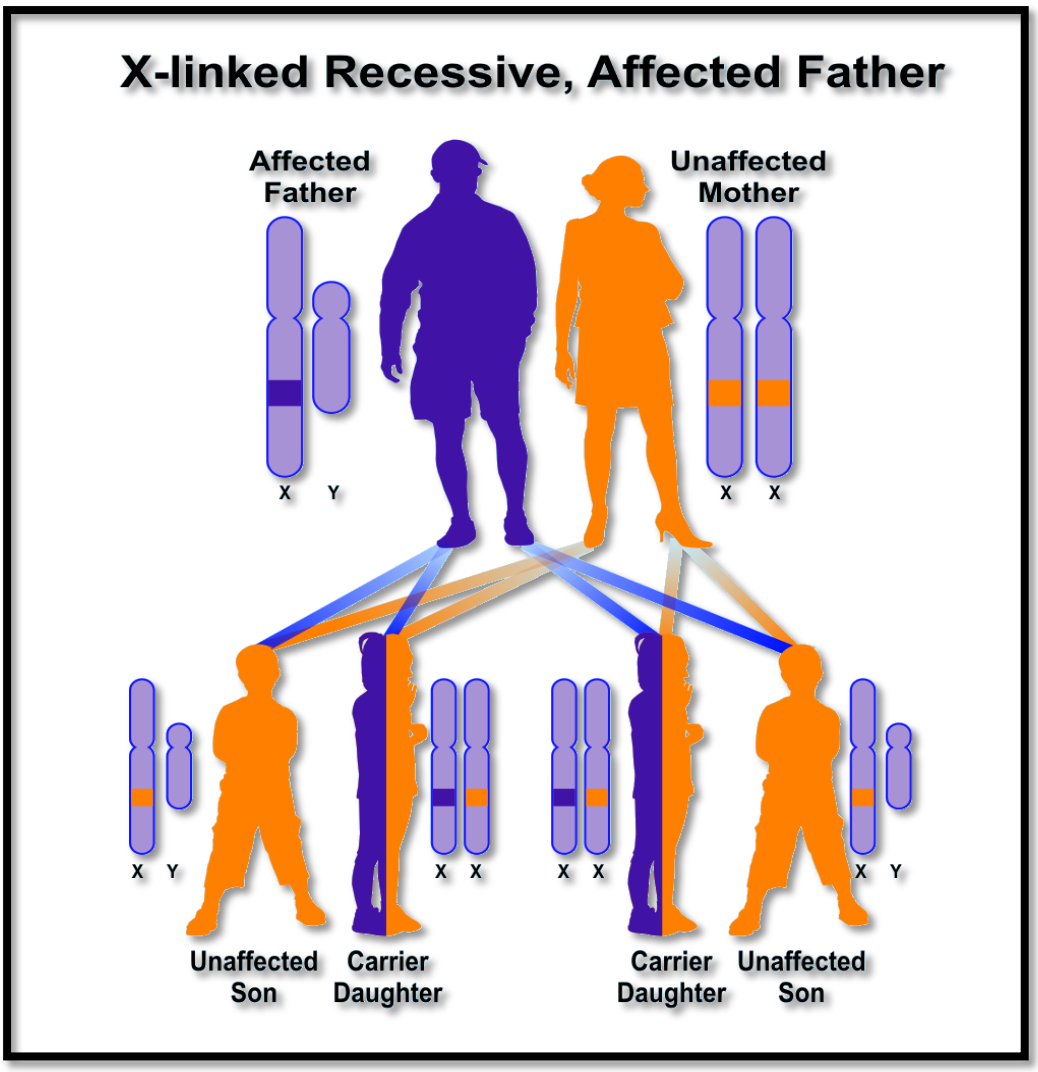

Figure 1.7: Schematic representation of inheritance pattern of X-linked disease when the father is affected and mother is normal. A pattern of G6PD (X-linked) inheritance illuminates that if the father is affected and the mother is normal, the father contributes to pass his defective X-chromosome to his daughter, as a result the daughter becomes carrier and may remain asymptomatic throughout her entire lifetime. But in this case the boys will never inherit the X-chromosome from his father and sons remain unaffected.

A girl may also get affected when mother is carrier or both parents contribute defective X chromosome to their daughter. In this case father is affected with the disease and mother is carrier or affected and both deliver defective x-chromosome to the daughter.

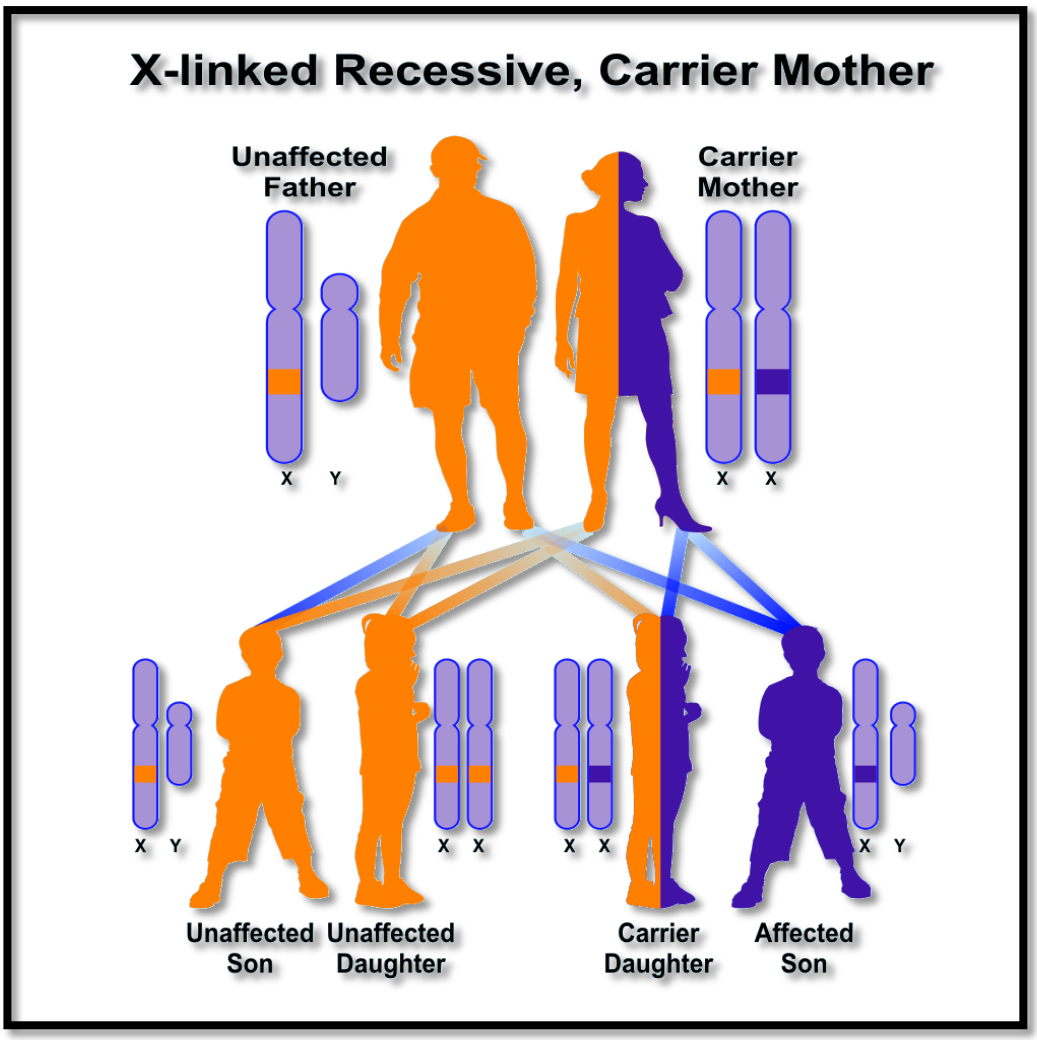

Figure 1.8: Schematic representation of inheritance pattern of X-linked disease when the mother is a carrier and father is normal. This model illustrates that when the carrier mother contributes the defective X-chromosome to her son; the son will become affected and suffer from the disease. But when the carrier mother passes the X-chromosome to a daughter, she becomes either a carrier or normal depending on the X-chromosome from mother.

Another way a girl may also be affected with G6PD deficiency even though she is heterozygote in terms of G6PD gene. A girl may acquire a defective X-chromosome from either of the parent and one normal X-chromosome from one of the parent. In this situation she should have normal G6PD activity as G6PD deficiency is an X-linked recessive disease and thus requires two defective X-chromosome to make a girl G6PD deficient. But genomic imprinting or X-chromosome inactivation may lead to G6PD deficiency in a heterozygous female due to inactivation of normal G6PD gene containing X-chromosome and leaving the defective G6PD gene containing X-chromosome in active state.

1.7 Prevalence of G6PD Deficiency

Glucose-6-phophate dehydrogenase (G6PD) deficiency is the most prevalent enzyme deficiency in the world, affecting 10% of the world's population (Acero 2013). It affects an estimated 400 million people worldwide and is most prevalent in Africa, Southeast Asia and the Middle East and lowest or absent in the Americas (Liwayway H. Acero et al). Howes and co-workers have identified 1,734 community G6PD surveys globally of which 1,289 (74%) were conducted in malaria-endemic countries and used this evidence-base to model a continuous prevalence map of the deficiency (Lorenz von Seidlein et al). Highest median prevalence (peaking at 32.5%) was predicted across sub-Saharan Africa and the Arabian Peninsula. Although G6PDd prevalence was generally lower across central and south East Asia, rarely exceeding 20%, the majority of G6PDd individuals (67.5% median estimate) were from Asian countries. They estimated a G6PDd allele frequency of 8.0% (interquartile range: 7.4–8.8) across MECs, and 5.3% (4.4–6.7) within malaria-eliminating countries.

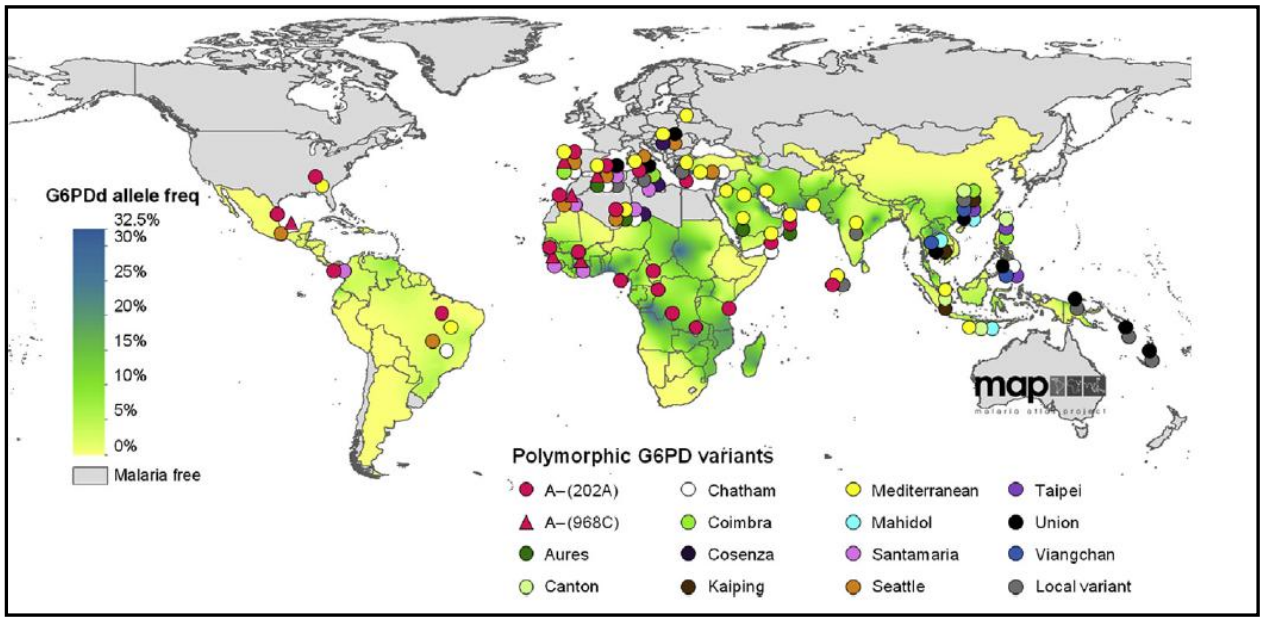

Fig 1.9: Worldwide distribution of G6PD deficiency. Color shades on the map indicate the median predicted allele frequency of G6PD deficiency in malaria-endemic and malaria eliminating countries, according to the geostatistical model designed by Howes and coworkers. Each colored circle illustrates the geographic distribution of 1 polymorphic G6PD allele present in more than 1 population. (Source: Howes et al, 2012)

The highest prevalence rate of G6PDd across the entire MEC region was 32.5% in the Eastern Province of Saudi Arabia. Further east, predicted prevalence remained high into southern Pakistan. This region had the highest uncertainty of the entire map (IQR exceeding

30%). No surveys were available from the south of Pakistan, and the closest neighbouring surveys in southern Iran, Oman. Prediction uncertainty dropped across central and southeast Asia, and predicted prevalence remained largely <10%, with three notable G6PDd prevalence hotspots in the central and southeast Asia regions peaking to >20%: (i) among the tribal, endogamous groups of Orissa province in east India, (ii) a patch along the northern Lao/Thai border, and (iii) much of the Solomon Islands archipelago. Underlying the broadly smooth continental-level variation, some areas were predicted to have highly heterogeneous sub-national G6PDd prevalence. Across Lao People's Democratic Republic (PDR), for instance, frequencies were predicted to range from 1% to 23%; predictions in Indonesia were from 0% to 15% in Nusa Tenggara; in Papua New Guinea, frequencies ranged from 1% along the southern coast to 15% along the East Sepik northern coast (Howes RE, Piel FB, Patil AP, Nyangiri OA, Gething PW, Dewi M, et al. (2012).

National frequency estimates ranged from 0.1% in Cape Verde and the Democratic People's Republic of Korea 22.3% in the Solomon Islands, 22.5% in the Congo and 23.0% in Benin. However, if the attention is shifted away from Africa towards the highly populous Asian countries, notably China and India where 41.3% of G6PDd males within MECs were predicted to be. Overall, the Americas contributed only 4.5% of the G6PDd male population, sub-Saharan Africa 28.0%, and Asia an estimated 67.5%.

1.8WHO Classification of G6PD

G6PD gene is probably the most polymorphic locus in humans, with over 400 allelic variants known. These variants have been biochemically characterized based on: the different residual enzyme activities, electrophoretic mobility patterns, and physicochemical (such as thermostability and chromatographic behavior) or kinetic (km for glucose-6-phosphate or NADPH, pH dependence, and utilization of substrate analogues) properties. In the following table (1.3) the classification of G6PD deficiencies shows different variants based on the criteria mentioned above.

Table1.3: Classification of G6PD deficiencies following the WHO recommendation (Vivek Chowdhry et al, 2011)

Class	Level of deficiency	Enzyme activity	Protein location	Clinical symptomps	Prevalance	Type of mutations
I	Severe	<1% or not detectable	Dimer interface, Structural NADP	Chronic hemolytic anemia	Uncommon ;occurs across populations	G6PD Buenos Aires, G6PD Durham
II	Severe	<10%	Dimer interface, NADP binding site	Acute hemolytic anemia fava beans And drug dependent	Varies; more common in Asia and Mediterran ean populations	G6PD-Mediterranean G6PDCassano ,G6PD-Santamaria
III	Moderate	10–60%	Scattered throughout the whole enzyme	Occasionally acute hemolytic anemia	10% of black males in the United States	G6PD-A$^-$, G6PD-Seattle, G6PD-Canton, G6PD-Rignano
IV	Mild to none	60–90% normal activity	Neutral protein site	Asymptomatic	Rare	G6PD-Montalbano, G6PD Orissa
V	None	>110% increased activity	Neutral protein site or promoter mutations	Asymptomatic	Rare	Not reported

1.9 Signs & Symptoms of Glucose-6-Phosphate Dehydrogenase Deficiency

Most individuals with G6PD deficiency are asymptomatic. The severity of symptoms associated with (G6PDD) varies greatly from case to case, depending upon the form of the disorder that is present, and some people have no symptoms at all. When symptoms are present, they may include fatigue, pale color, shortness of breath, rapid hearbeat, jaundice or yellow skin color, dark urine and enlarged spleen (splenomegaly). In the relatively rare, severe, potentially life-threatening cases, symptoms include, in addition to those listed above, others such as: blood in the urine (hemoglobinuria), shock, kidney (renal) failure and congestive heart failure in which the heart is unable to pump blood effectively throughout the body.

1.10 Diagnosis of G6PD

The available tests to assess an individual's G6PD activity status can be broadly divided into three categories. Figure 1.10 shows the subdivisions of phenotypic tests for screening of G6PD deficiency.

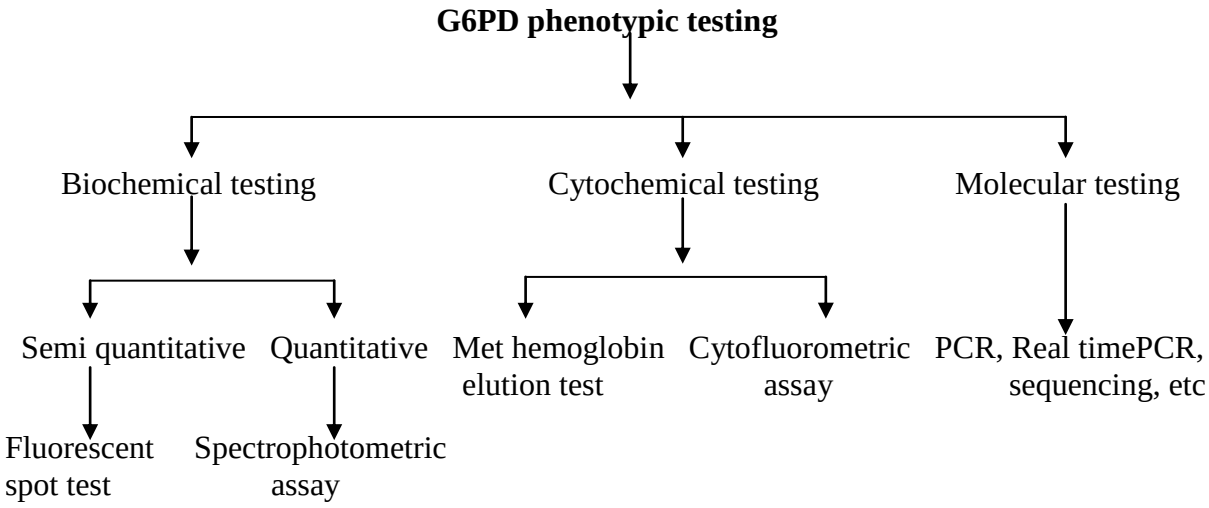

Figure1.10: Summary of phenotypic tests that can be used for screening of G6PD deficiency.

1.11 Treatment of G6PD deficiency

Treatment is applied to G6PD patients only when symptoms arise and this may include blood transfusions, maintaining adequate urine output and splenectomy (Rosse and Bunn 1998). If the anemia is severe, a blood transfusion may be required. In the newborn, a measure to reduce the bilirubin is needed as this is harmful to the baby. This can be done by shining a `blue' light to the baby to destroy the bilirubin through the skin. This is called `phototherapy' (Liwayway H. Acero, 2012). If the level is higher, then an exchange transfusion in which the blood with high bilirubin and low hemoglobin is removed and replaced by blood with low bilirubin and high hemoglobin thus reducing the bilirubin and normalizing the hemoglobin. The patient who suffered a hemolytic episode needs to take plenty of water. However prevention is the most important measure (Vandaveer, Woodward et al. 2003). It is best to avoid the proscribed drugs and foods that may cause hemolysis. Some of the inhibited diets and drugs for G6PD patients are given in table 1.4.

Table 1.4: List of some Foods, Chemicals and Drugs that should be avoided by G6PD deficient (Lee, Lai et al. 2016)

Food/drinks	Drugs/chemicals
Fava Beans -Dingdong nuts, Mr. Bean Red wine **Legumes** -Abitsuela, Garbansos, Kadyos, Munggo, Blueberry **Soya foods** -Taho, Tokwa, Soy sauce Tonic water Bittermelon/ amplaya Henna Herbs Vitamin K and its derivatives **Artificial food coloring substances** Methylene blue Toluidine blue **Preservatives** Sulfites and foods containing them	**Anti-malarials** Primaquine, Chloroquine, Pamaquine **Sulphonamides** Sulphanilamide, Sulphapyridine, Sulphamethoxazole **Sulphones** Dapsone, Thiazolesulfone, Glucosulphone sodium, Glyburide/Glibenclamide, Mafenide acetate, Salicyazosuphapyridine/Sulfasalazine,Sulphadimidine ,Sulphafurazone, Sulphamethazole **Other sulfur containing** Glibenclamide **Nitrofurans** Nitrofurantoin, Nitrofuracin, Furazolidone **Miscellaneous** Naphthalene (mothballs), Niridazole, Methylene blue, Urate oxidase, Phenylhydrazine, Phenazopyridine, Isobytylnitrite, Acetanilid, Spiramycin, Menthol

1.12 Real-Time PCR

A real-time polymerase chain reaction is a laboratory technique of molecular biology based on the polymerase chain reaction (PCR). It monitors the amplification of a targeted DNA molecule during the PCR. Real-time PCR can be used quantitatively (Quantitative real-time PCR- numbers of copies of DNA), semi-quantitatively, or qualitatively (Qualitative real-time PCR- presence or absence of sequence).Two common methods for the detection of PCR products in real-time PCR are:

(1) non-specific fluorescent dyes that intercalate with any double-stranded DNA, and

(2) sequence-specific DNA probes consisting of oligonucleotides that are labelled with a fluorescent reporter which permits detection only after hybridization of the probe with its complementary sequence.

In conventional PCR, the amplified product, or amplicon is detected by and end point analysis, by running DNA on an agarose gel after the reaction has finished. In contrast, real time PCR allows the accumulation of amplified product to be detected and measured as the reaction progresses. Real time detection of PCR products is made possible by including in the reaction a fluorescent molecule that reports an increase in the amount of DNA with a proportional increase in fluorescent signal. The fluorescent chemistries employed for this purpose include DNA-binding dyes and fluorescently labeled sequence specific primers or probes. Specialized thermal cyclers equipped with fluorescence detection modules are used to monitor the fluorescence as amplification occurs. The measured fluorescence reflects the amount of amplified product in each cycle.

The main advantage of real time PCR over conventional PCR is real-time PCR allows determining the starting template copy number with accuracy and high sensitivity over a wide dynamic range. Real-time PCR is quantitatively known as qPCR. In contrast, conventional PCR is at best semi quantitative. Additionally, real-time PCR data can be evaluated without gel electrophoresis, resulting in reduced experiment time and increased throughput. Finally, because reactions are run and data are evaluated in a closed-tube system, opportunities of contamination are reduced and the need for post amplification manipulation is eliminated.

1.13 How Real-Time PCR Works?

To understand how real-time PCR works, we have to examine sample amplification plot from figure 1.11. In this plot, the PCR cycle number is shown on the X-axis, and the fluorescence from the amplification reaction, which is proportional to the amount of amplified product in the tube, is shown on the Y-axis.

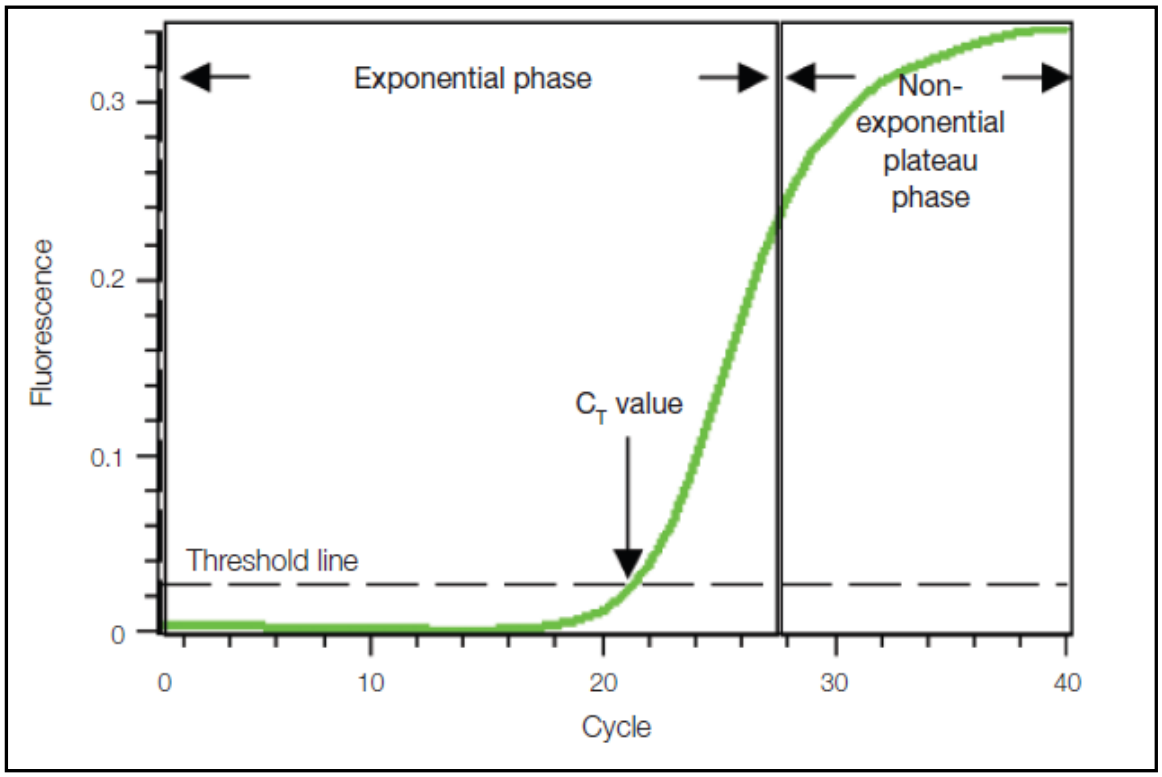

Figure 1.11: Amplification plot. Baseline-subtracted fluorescence is shown. (Source: Bio-rad, real time PCR).

The amplification plot shows two phases, an exponential phase followed by non-exponential plateau phase. During the exponential phase, the amount of PCR product approximately doubles in each cycle. As the reaction proceeds, however, components are consumed, and one or more of the components becomes limiting. At this point the reaction slows and enters the plateau phase (cycles 28-40 in figure 1.11).

Initially, fluorescence remains at background levels, and increases in fluorescence are not detectable (cycle 1-18 in figure 1.11) even though product accumulates exponentially. Eventually, enough amplified product accumulates to yield a detectable fluorescence signal. The cycle number at which it occurs is called the threshold cycle or $C_{T.}$ Science the C_T value

is measured in the exponential phase, when reagents are not limited, real-time PCR can be used to reliably and accurately calculate the initial amount of template present in the reaction.

The C_T of a reaction is determined mainly by the amount of template present at the start of the amplification of the reaction. If a large amount of template is present at the start of the reaction, relatively few amplification cycles will be required to accumulate enough products to give a fluorescent signal above background. Thus, the reaction will have a low C_T value. In contrast, if a small amount of template is present at the start of the reaction; more amplification cycle will be required for the fluorescent signal to rise above background. Thus, the reaction will have a high or late C_T. This relationship forms the basis for the quantitative aspect of real-time PCR.

1.14 High resolution melt (HRM) curve analysis

High Resolution Melting Analysis (HRM or HRMA) is a recently developed technique for fast, high-throughput post-PCR analysis of genetic mutations or variance in nucleic acid sequences. It enables researchers to rapidly detect and categorize genetic mutations (e.g. single nucleotide polymorphisms (SNPs)), identify new genetic variants without sequencing (gene scanning) or determine the genetic variation in a population (e.g. viral diversity) prior to sequencing.

The first step of the HRM protocol is the amplification of the region of interest, using standard PCR techniques, in the presence of a specialized double-stranded DNA (dsDNA) binding dye. This specialized dye is highly fluorescent when bound to dsDNA and poorly fluorescent in the unbound state. This change allows the user to monitor the DNA amplification during PCR (as in quantitative PCR).

After completion of the PCR step, the amplified target is gradually denatured by increasing the temperature in small increments, in order to produce a characteristic melting profile; this is termed melting analysis. The amplified target denatures gradually, releasing the dye, which results in a drop in fluorescence. When set up correctly, HRM is sensitive enough to allow the detection of a single base change between otherwise identical nucleotide sequences. HRM uses low-cost dyes and requires less optimization than similar systems based on TaqManR and fluorescence resonance energy transfer (FRET) probes. Compared to these methods HRM is a simpler and more cost-effective way to characterize multiple samples.

1.15 Melting Profile Principle

Amplification product melting analysis is not a novel concept. Post-PCR melt curve analysis using SYBRR Green I dsDNA intercalating dye to detect primer-dimers or other non-specific products has been performed for over a decade. Today this procedure is called Low Resolution Melting. A Low Resolution Melt curve is produced by increasing the temperature, typically in 0.5 °C increments, thereby gradually denaturing an amplified DNA target. Since SYBRR Green I is only fluorescent when bound to dsDNA, fluorescence decreases as duplex DNA is denatured. The melting profile depends on the length, GC content, sequence and heterozygosity of the amplified target. The highest rate of fluorescence decrease is generally at the melting temperature of the DNA sample (Tm). The Tm is defined as the temperature at which 50% of the DNA sample is double stranded and 50% is single-stranded. The Tm is typically higher for DNA fragments that are longer and/or have a high GC content. The fluorescence data from low resolution melting curves can easily be used to derive the Tm by plotting the derivative of fluorescence vs. temperature (-dF/dT against T) as shown in Figure 1.12.

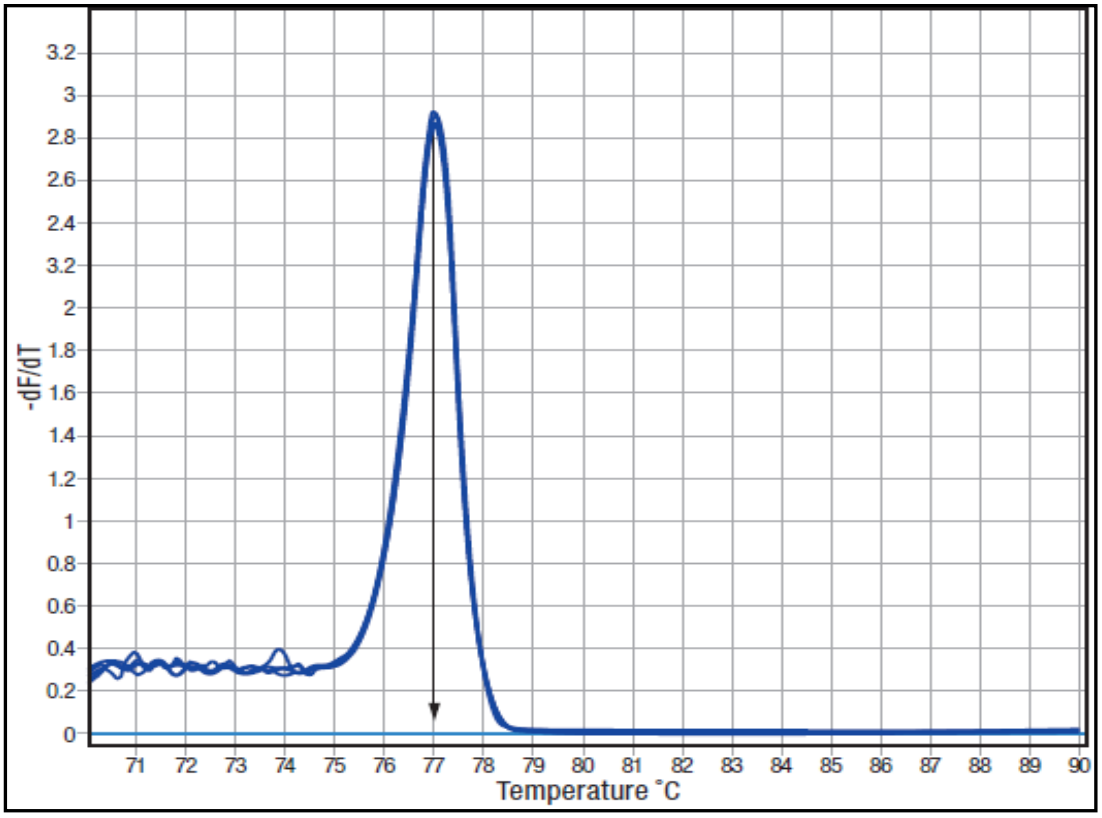

Figure 1.12: Low Resolution Melt profile derivative plot (-dF/dT against T). The steepest slope is easily visualized as a melt peak. In this example the Tm of the amplicon is 77 °C (Sourec: Kapa Biosystem)

The principle of HRM is the same as a Low Resolution Melt, except that the temperature difference between each fluorescence reading is reduced. During a Low Resolution Melt curve analysis, the temperature increases are typically in 0.5 °C steps, but for HRM this is reduced to 0.008 - 0.2 °C increments. This allows a much more detailed analysis of the melting behavior. HRM sensitivity and reliability has been improved with the use of a variety of new dsDNA intercalating dyes; this is covered in more detail in the following section. The introduction of HRM has revived the use of DNA melting for a wide range of applications, including:

- SNP genotyping
- Mutation discovery (gene scanning)
- Heterozygosity screening
- DNA fingerprinting
- Haplotype blocks characterization
- DNA methylation analysis
- DNA mapping
- Species identification
- Viral/bacterial population diversity investigation
- HLA compatibility typing
- Association study (case/control)

1.16 The Intercalating Dye

Chemical compounds with a high affinity for DNA whose molecules intercalate between the planar base pairs of the DNA helix. They are used for visualizing DNA, changing its density and to induce breakage.There are various types of dsDNA intercalating dyes, which have distinctly different properties. The dyes used for HRM are different from dyes typically used for standard quantitative PCR (qPCR) assays. Signal-to-noise ratio and amplification efficiency are not essential requirements for HRM. Instead, the dye must provide detailed information on the melting behavior of an amplified target. Ideally the dye should not bind preferentially to pyrimidines or purines, change the Tm of the amplicon, or inhibit DNA amplification. The three main classes of dsDNA binding dyes are-

Non-saturating dyes: SYBR® Green I is the most commonnon-saturating dsDNA intercalating dye. It is generally unsuitable for most HRM applications. At high concentrations, SYBR® Green I stabilizes the duplex DNA and inhibits the DNA polymerase. To allow reliable amplification, low concentrations of SYBR® Green I must therefore be used. At lower concentrations, SYBR® Green I is able to redistribute from the

melted regions of single-stranded DNA back to the regions of dsDNA, which results in poor base-difference discrimination, as detailed in Figure 1.13. To overcome this limitation, saturating dyes have been developed.

Saturating dyes: A new class of dsDNA intercalating dyes that do not inhibit DNA polymerases, or alter the Tm of the product, have recently been developed; these dyes can be added at higher concentrations than non-saturating dyes, ensuring more complete intercalation of the amplicon. The dye is not able to redistribute during melting because the dsDNA is saturated. More precise examination of the melting behavior is therefore possible, as indicated in Figure 1.13. Dyes such as SYTO9® and LC Green® can be used at the saturating concentrations required for HRM.

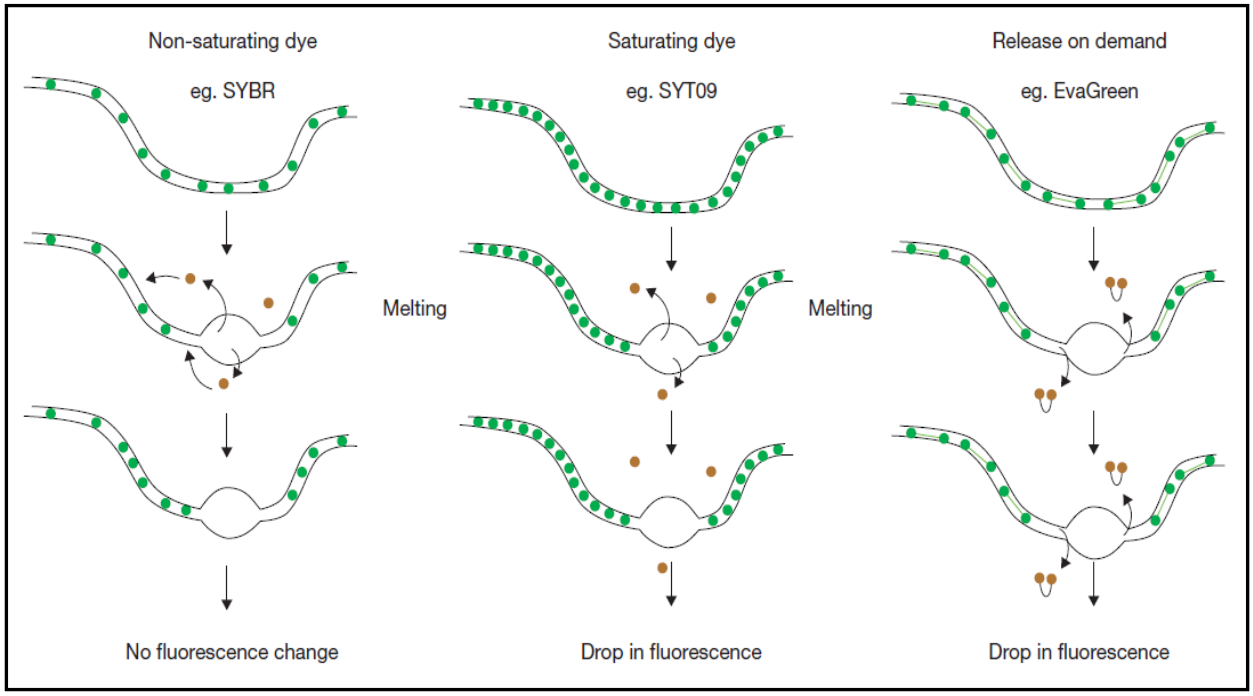

Figure 1.13:Non-saturating, saturating and 'release-on-demand' dsDNA intercalating dyes. Melting of the duplex as the temperature increases releases the intercalated dyes. At non-saturating concentrations the dye rapidly rebinds to regions that remain double stranded; consequently there is no drop in fluorescence. Saturating and 'releaseon-demand' dyes do not redistribute from the melted regions of single-stranded DNA back to dsDNA, resulting in a reduction of fluorescence. This difference gives dyes suchas EvaGreenR the high sensitivity required for HRM analysis (source: Kapa Biosystem)

Release-on-demand dyes: The "release-on-demand" class of dyes, which include EvaGreen®, can be added at non-saturating concentrations. This is due to the novel method of fluorescence emission, where the fluorescent signal is quenched when the dye is free in solution. Upon binding to duplex DNA, the quenching factor is released and the dye emits high fluorescent signal. This allows non-saturating concentrations of the dye to be used,

ensuring that there is no PCR inhibition, whilst the unique dye chemistry provides highly sensitive HRM analysis.

1.17. Aims and objectives of the study

The main goals of this study are-

1. To make use of a rapid, reliable, sensitive and more cost effective molecular method for detection of common mutations of Glucose-6-phosphate dehydrogenase gene in Bangladeshi population.
2. To employ a molecular technique to detect heterozygous G6PD mutations in deficient samples, which is not possible using conventional qualitative and quantitative tests.

Chapter 2: Materials and methods

2.1 Study Place

This study was conducted in the Genetics and Genomics laboratory of institute for developing science and health initiative (ideSHi). All participants in this study were in the age range of 0-15 years. Ethical approval for this study was obtained from the ethical review board of Bangladesh Medical Research Council (BMRC).

2.2 Study Participants

A total of 121 participants suspected for G6PD deficiency based on signs and symptoms like hyperbilirubinemia, hemolytic anemia, and dark urine were enrolled into the study. Samples were collected from Bangabandhu Sheikh Mujib Medical University (BSMMU), Bangladesh Institute of Research and Rehabilitation for Diabetes, Endocrine and Metabolic Disorder (BIRDEM) and Shishu Hospital. Informed consent was taken from each participant along with their case history and clinical information including sex, age, blood group, bilirubin level, hemoglobin level etc.

2.3 Study design

2.3.1. Sample collection

Blood specimens from suspected patients were collected in anticoagulant coated vacutainers like heparin or EDTA tube for G6PD enzyme assay and genetic analysis. For genetic analysis from RNA, 250µl of blood was taken in a micro-centrifuge tube containing 750µl Trizol ® LS (Ambion, Life Technology, Cat.No.10296-028). In case of newborn, few drops of blood were collected on a dried blood spots card (DBS) through heel pricking. For patients within 5-15 years of age, blood were collected from vein by venipuncture method and spotted on the DBS card.

2.3.2 Sample storage

The sample collected in the anticoagulant tube (EDTA) was stored at 2-8 degree centigrade for the first day, and then transferred to -70 degree on the second day. The sample collected in Trizol® LS was stored at -70 degree immediately after transferring the blood into Trizol® from the very first day. The DBS card was stored at -70 for long time preservation in an airtight bag.

2.3.3 Sample processing

Blood collected in EDTA tube was used for fluorescent spot test (FST), enzyme activity measurement and mutational analysis from DNA. Blood stored in Trizol LS was used for DNA isolation, RNA isolation, PCR (polymerase chain reaction), reverse transcription, real

time PCR and further mutational analysis. The dried blood sample on DBS card was stored for future usage.

2.3.4 Overview of the study plan

The overall plan of this study is given below-

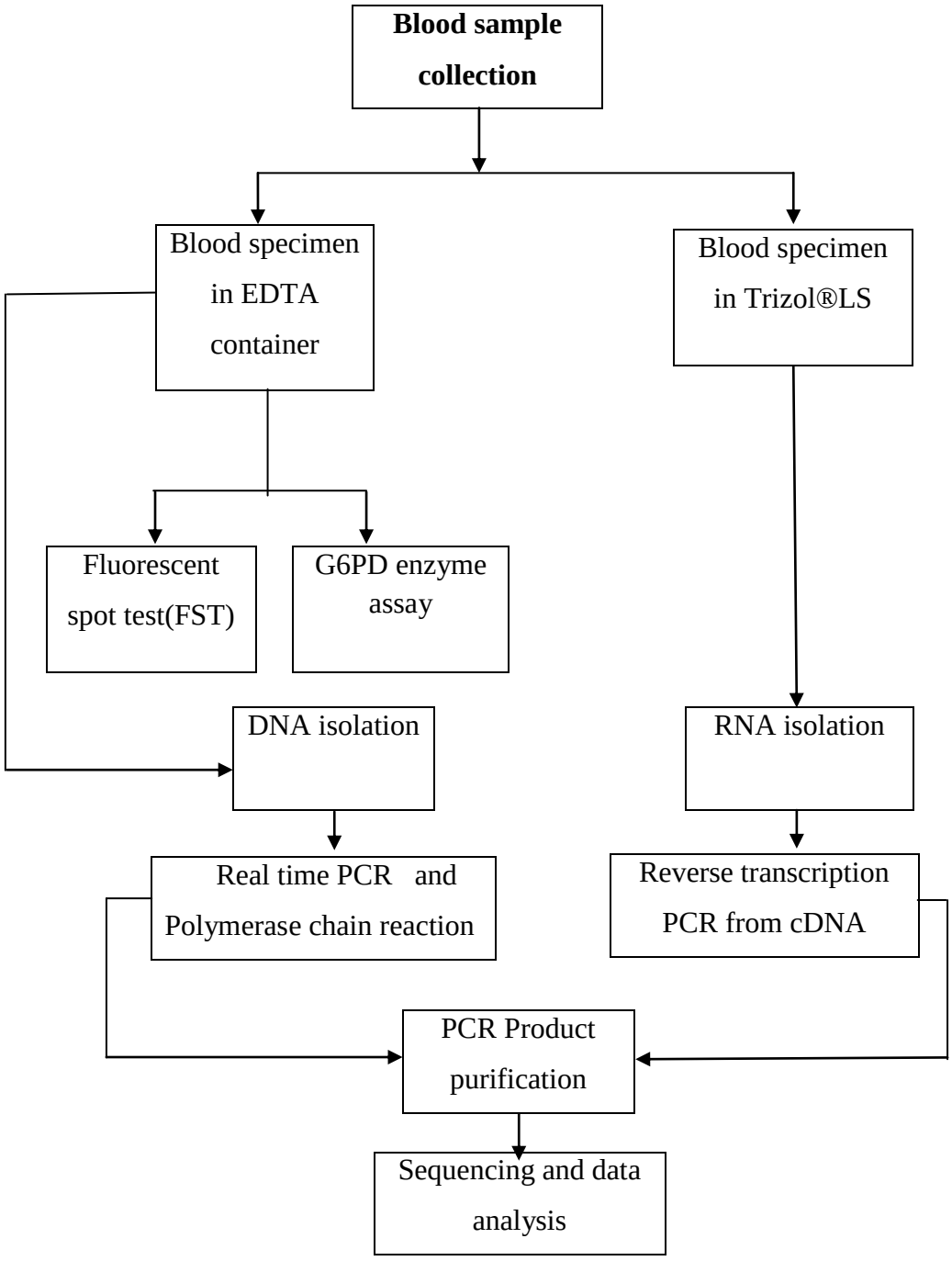

Figure 2.1: Flow chart of study plan

2.4 Fluorescent spot test

The principle of the fluorescent spot test is based upon the fluorescence of NADPH under long-wave UV light. The reduction of $NADP^+$ to NADPH occurs in the presence of G6PD and the rate of NADPH formation is proportional to G6PD activity. Results were classified in three groups: "bright fluorescence" (BF), "weak fluorescence" (WF), and "no fluorescence" (NF). The FST for G6PD deficiency involves this reaction:

$$\text{Glucose-6-Phosphate} + \text{NADP} \xrightarrow{\text{G6PDH}} \text{6-Phosphogluconate} + \text{NADPH}$$

(Not Fluorescent) (Fluorescent)

2.4.1 Equipment and reagents used for FST

2.4.1.1Equipment

- Biosafety cabinet (level II)
- Gel Doc,XR (BioRad, USA)
- Refrigerator

2.4.1.2Materials

- 11 cm filter paper
- Pipettes (10-200)
- Eppendorf

2.4.2. Procedure

2.4.2.1. Preparation of screening solution

1 mL screening solution was prepared in a 1.5 mL micro-centrifuge tube by mixing-

- ✓ 100 µL glucose-6-phosphate (Sodium-salt) (0.001 M, Sigma, Pro. Code. 1001715787),
- ✓ 100 µL $NADP^+$ (0.0075M,Sigma, Pro. Code. 1001774334),
- ✓ 200 µL Saponin (1%, Sigma, Pro. Code. 101396863),
- ✓ 300 µL Tris-HCl buffer pH 7.8 (0.75 M, Sigma, Pro. Code. 1001681493),
- ✓ 100 µL glutathione (0.008 M, Sigma, Pro. Code 1001850842) and
- ✓ 200 µL deionized water.

The reagents that were used for the preparation of screening solution are given along with their storage temperature, concentration and volume in Table 2.1.

Table 2.1: Reagents used for preparation of screening solution

Name of the reagents	Storage temperature	Concentration	Volume (ml)
Glucose-6-phosphate (Sodium-salt)	-20^{o}C	0.001M	0.1
NADP+ solution	-20^{o}C	0.0075M	0.1
Saponin	-20^{o}C	1%	0.2
Tris-HCl buffer (pH-7.8)	Room temperature	0.75M	0.3
Oxidized glutathione (GSSG) solution Filtered deionized	-20^{o}C	0.008M	0.1
H2O	Room temperature		0.2
Total volume			1.0

The mixture was vortexed for 2-3 minutes to homogenize the screening solution. Prepared screening solution was kept at 4^{o}C.

2.4.3 Fluorescent spot test

100 µL of screening solution was taken in a micro-centrifuge tube. 10 µL of whole blood was added to the screening solution. The mixture was then mixed well by pipetting and tapping. 10 µL of the mixture was used to make spots on a filter paper (Double Ring 102, China) at intervals of 0, 5, 10 and 15 minutes. The spots were allowed to dry at room temperature for >2hours. After drying, the reading of the filter paper was taken using Gel DocTM XR+ (BioRad, USA). Each test included a G6PD deficient sample as a deficient control. For a moderately deficient individual, the blood specimen was taken a second time to confirm the first result. Old and freshly drawn blood was used for FST along with deficient control.

2.4.4 Precautions

The general precautions required for handling laboratory reagents were taken. All samples were handled carefully as blood specimen can contain infectious pathogens that can be contracted from unprotected handling of blood. The wastes were disposed-off properly.

2.5 G6PD enzyme activity measurement

2.5.1 Principle

Reduction of $NADP^+$ to NADPH by G6PD causes a gradual increase in absorbance at 340 nm. The G6PD enzyme activity is quantitatively determined by measurement of the rate of change of absorbance at 340 nm due to reduction of $NADP^+$ to NADPH (Minucci, 2009).

Glucose-6-phosphate + $NADP^+$ $\xrightarrow{\text{G6PD}}$ Gluconate-6-phosphate + NADPH + H^+

2.5.2 Equipments and reagents

2.5.2.1 Equipments

- Micro-centrifuge
- Water bath
- Spectrophotometer or EON ELISA reader (Bio Tek, USA)
- Vortex machine
- Biosafety cabinet (Level II)
- Refrigerator (+4°C, -70°C and -20°C)

2.5.2.2 Materials

- BD Vacutainer (EDTA-coated)
- Bio-Cell Cuvette (BioTek, USA)
- Take3 plate (BioTek, USA)
- 1 mL micropipettes
- 20-200 µL micropipettes
- 1-10 µL micropipettes

2.5.2.3 Reagents

Reagents used for G6PD enzyme activity measurement were RANDOX G6PD assay kit (Cat. no. PD410, 100mL, Randox Laboratories Limited, UK). The components along with their storage temperature is given below-

Reagents	Storage temperature
0.9 % NaCl solution	Room temperature
R1 solution: Buffer (Randox)	2-8^{o}C
R2 Solution: NADP (Randox)	2-8^{o}C
R3 Solution: Substrate (Randox)	2-8^{o}C
R4 Solution: Digitonin (Randox)	2-8^{o}C

2.5.3 Enzyme activity measurement

First, 100 µL of whole blood from EDTA-coated vacutainer was added to 900 µL saline (0.9%) and washed by centrifuging at 3,000 rpm for 10 minutes. The supernatant was then discarded, the pellet resuspended in saline, and washed 3-4 times. The pellet of erythrocyte was suspended in 250 µL of R4 solution containing digitonin. The suspension was incubated at 2-8°C for 15 minutes and then centrifuged again at 3000 rpm for 10 minutes. The resultant hemolysate was used for the enzyme activity measurement. One mL of R1 solution was taken in an Eppendorf tube and incubated at 37°C in a water bath for 2-3 minutes. The R1 solution was then taken out of the water bath followed by addition of 30 µL of R2 solution and 15 µL of hemolysate. The mixture was vortexed to mix the components well and then incubated for 5 minutes at 37^{o}C in a water bath. Following incubation, 15 µL of R3 solution was added to the mixture. The mixture was vortexed to ensure proper mixing and spun briefly for 1-2 seconds. The mixture was then pipetted into a Bio-Cell cuvette (BioTek, USA). Care was taken not to leave any bubbles within the cuvette as bubbles affect absorbance and may give falsely high or low value. The cuvette was placed in a Take3 plate (BioTek, USA) and the Take3 plate was then placed in the EON ELISA reader (BioTek, USA). Five absorbances (A) were taken at 340 nm at 1-minute interval, and the change in absorbance (ΔA) was used for measurement of enzyme activity.

2.5.4 Calculation

When the temperature was 37°C, no temperature correction factor (TCF) was required. If results for patient samples were reported at a temperature other than 37°C, a TCF was needed to be used. To calculate G6PD activity as U/g hemoglobin, the following equation was used –

$$\text{G6PD U/g Hb} = \frac{33650 \times \Delta A\ 340\ \text{nm/min} \times 100}{\text{Hb (g/dL)} \times 1000}$$

100 = Factor to convert mL to dL

33650 = Product of molar extinction co-efficient and dilution factor

1000 = Factor to convert mU to U

ΔA 340 nm/min = Change in absorbance at 340 nm per minute

The enzyme activity was measured at an ambient temperature of 25°C. To get the enzyme activity at 37°C, resultant enzyme activity was multiplied by temperature correction factor 2.076 as absorbance was measured at 25°C.

2.5.5 Normal range of enzyme activity

According to kit provider/manufacturer, the normal range of G6PD enzyme activity was 6.97 - 20.5 U/g Hb (at 37^{o}C).

2.5.6 Quality Control

Randox G6PD deficient control (Cat. No. PD2617, 100 mL) and G6PD normal control (Cat. No. PD2618, 100 mL) were used for validation of the process.

2.5.6.1 G6PD deficient control

Preparation

First, the lyophilized red cell hemolysate that came with the kit was reconstituted with 500 µL of deionized water. The vial of the reconstituted hemolysate was allowed to stand at room temperature for 15 minutes so that the hemolysate was dissolved completely. The vial was then inverted several times to ensure saturation of the solution and spun for few seconds to make sure that nothing remained on the wall and cap of the vial.

Stability

The lyophilized red cell hemolysate was stable until expiry date and the reconstituted G6PD deficient control was stable up to 15 days at 2-8oC. According to kit provider/manufacturer, G6PD deficient control activity range was 223 ± 45 U (at 37oC).

2.5.6.2 G6PD normal control

Preparation

The lyophilized red cell hemolysate that came with the kit was reconstituted with 500 µL of deionized water. The vial of the reconstituted hemolysate was allowed to stand at room temperature for 15 minutes so that the hemolysate dissolved completely. The vial was

inverted several times to ensure saturation of the solution and spun briefly to make sure that nothing remained on the wall and cap of the vial.

Stability

The lyophilized red cell hemolysate was stable until expiry date and the reconstituted G6PD deficient control was stable up to 15 days at 2-8oC. According to kit provider/ manufacturer, G6PD normal control activity range was 1359 ± 272 U (at 37^{o}C).

2.5.7 Precautions

The normal precautions required for handling laboratory reagents were exercised

- As R1 Solution contains sodium azide, care was taken so that no ingestion or contact with skin or mucous membranes could occur. In the case of skin contact, affected area was needed to be flushed with profuse amounts of water. In the case of contact with eyes or if ingested, immediate medical care was necessary to be sought.

- As sodium azide can form potentially explosive azides by reacting with lead and copper plumbing, enormous amounts of water was used during the disposal or flushing of such reagents to prevent azide build up. Exposed metal surfaces were needed to be rinsed with 10% sodium hydroxide.

- As reticulocytes have higher G6PD enzyme activity levels than mature RBCs, this assay was recommended not to be performed after a severe hemolytic crisis since G6PD levels might appear falsely elevated. In such cases, testing was done after the level of mature red cells have returned to normal.

2.6 DNA Extraction

2.6.1 Principle

The addition of lysis buffer to sample causes lysis of cells and brings DNA out in the solution. Subsequent centrifugation causes the nucleic acid to precipitate at the bottom of the tube. Resuspension of the pellet in denaturing buffer causes denaturation of protein due to presence of chaotropic agents and protease, leading to the efficient removal of proteins and other contaminants. DNA is precipitated by addition of isopropanol. Precipitated DNA is washed with 70% ethanol, dried and re-dissolved in hydration buffer containing Tris-HCl buffer with pH 8.5.

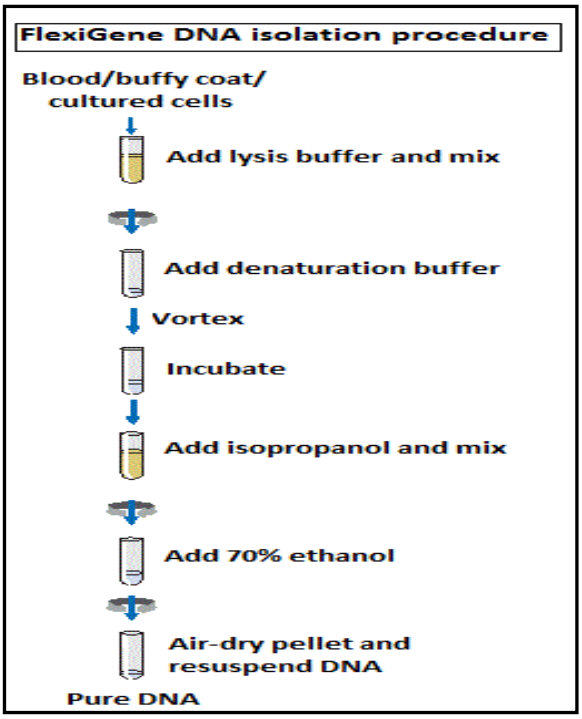

Fig 2.2: DNA extraction from whole blood

2.6.2 Equipment and reagents for DNA extraction

2.6.2.1 Equipment

- 100% isopropanol
- 70% ethanol
- Pipettes
- Sterile, DNase-free pipette tips (with aerosol barrier)
- 1.5 mL and 2.0 mL micro-centrifuge tubes
- Biosafety hood
- Water bath
- Vortex mixture
- Centrifuge machine

2.6.2.2 Reagents

QIAGEN FlexiGene® DNA Kit (250), Cat. No. 51206 containing

- QIAGEN protease
- Buffer FG1 (Lysis buffer)
- Buffer FG2 (Denaturation buffer)

The kit also contained Buffer FG3 (Hydration buffer). But instead of using Buffer FG3, nuclease free water was used. Buffer FG1, Buffer FG2 and Buffer FG3 were stored at 15-25^{o}C. QIAGEN protease was stored at -20^{o}C.

2.6.2.4 Starting material

Whole blood was used as starting material. Fresh sample or frozen samples could be used for DNA extraction. Frozen sample was thawed at 37^{o}C water bath quickly with mild agitation.

2.6.3 Procedure

500 µL of FG1 buffer was taken in a 1.5 mL micro-centrifuge tube. 200 µL of whole blood was added to the FG1 buffer and mixed by inverting the tube 5-10 times. The mixture was then centrifuged at 10,000 x g for 5 minutes in fixed angle rotor. The supernatant was carefully discarded, and the pellet was dried by inverting the micro-centrifuge tube on a clean tissue paper for 2 minutes (making sure the pellet remained in the tube). In a fresh eppendorf tube, one µl of QIAGEN protease was added to 100 µl of FG2 buffer and mixed by vortexing. Then 100 µL of FG2/QIAGEN protease was added to the pellet and vortexed immediately until the pellet was completely dissolved and the color was changed into olive green. Protease was given in excess amount if the pellet didn't dissolve properly after vortexing. The tube was then spun briefly for 5 seconds. The mixture was then incubated in a water bath at 65°C for 5 minutes. After incubation, 100 µL of isopropanol (100%) was added and mixed by inversion until DNA precipitated as visible threads. Isopropanol amount was increased as much as the amount of extra protease added if DNA didn't precipitate. The tube was centrifuged for 5 minutes at 10000 × g. The supernatant was discarded, and the pellet was dried by keeping the tube inverted state on a clean tissue paper for 1 minute. 100 µL of 70% ethanol was added and vortexed for 5 seconds. The tube was centrifuged for 5 minutes at 10,000 × g. The supernatant was carefully aspirated using a micropipette and keeping the micro-centrifuge tube in the inverted state on tissue paper; the pellet was allowed to air dry for at least 5 minutes. Over-drying was avoided because it would be tough to dissolve over-dried DNA. Depending on the pellet size, 25-50 µL of H_20 was added, and the tube was vortexed briefly for 5 seconds, and the mixture was incubated at 65^{o}C for 1 hour in a water bath for dissolving DNA. FG3 buffer could be used in this step instead of H_20. The mixture was then allowed to stand at room temperature for overnight. Finally, the amount of DNA was quantified and then stored at -700C for downstream use in PCR or other analysis.

2.7 RNA isolation from whole blood mixture using Trizol® LS

2.7.1 Principle

Homogenized sample in Trizol® LS is mixed with chloroform and subsequent centrifugation will cause formation of three phases. The upper layer is the aqueous phase containing the RNA; the interphase contains a small amount of DNA and proteins and the lower organic phase contains the majority of DNA, proteins, lipids and other cell debris. Separation of the aqueous phase and subsequent addition of isopropanol to the separated aqueous phase causes the RNA to precipitate.

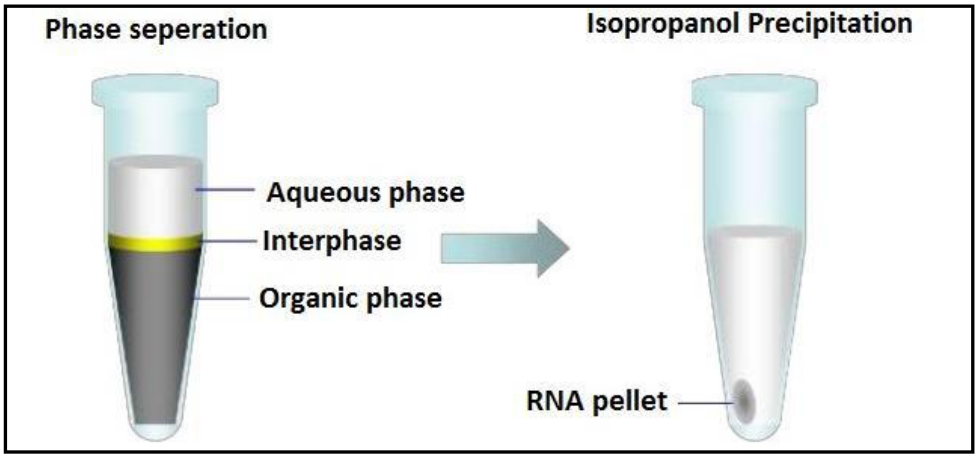

Figure 2.3: Phase separation and isopropanol precipitation of RNA from Trizol® LS

Precipitated RNA is washed with 70% ethanol, and the pellet is air-dried. RNase-free water is added; the pellet is dissolved and incubated in a water bath for 1 hour at 55-60oC. Then the RNA preparation is kept in the -70^{o}C for downstream applications like PCR.

2.7.2 Equipment and reagents for RNA extraction

2.7.2.1 Equipment

- Pipettes
- Sterile, DNase-free pipette tips (with aerosol barrier)
- 1.5 mL and 2.0 mL micro-centrifuge tubes
- Bio-safety hood
- Water bath
- Vortex mixture
- Centrifuge machine

2.7.2.2 Reagents and their Storage temperature

The reagents that were used for RNA extraction are given along with their storage temperature in Table 2.2.

Table 2.2: Reagents used for RNA extraction and their storage temperature

Reagents	Storage temperature
Trizol® LS reagent	4^{o} C
Chloroform	Room temperature
100% Isopropanol	Room temperature
75% Ethanol	Room temperature
RNase-free water	Room temperature

2.7.3 Procedure

2.7.3.1 Sample homogenization

750 µL Trizol® LS reagent (200 mL, Ambion by Life Technology, Cat. No. 10296-028) was taken in a micro-centrifuge tube, and 125 µL of whole blood specimen and 125 µL H_2O was added to it. The tube was inverted several times to homogenize the mixture. The homogenized sample was stored at -70^{o}C.

2.7.3.2 Phase separation

The homogenized sample was incubated at room temperature for 5 minutes. 70 µL of chloroform was added to the homogenized mixture. It was then incubated for 15 minutes at room temperature, and the mixture was centrifuged at 14000 × g for 15 minutes at 4^{o}C. The upper aqueous phase was aspirated and transferred into another micro-centrifuge tube without drawing the interphase or organic phase.

2.7.3.3 RNA isolation

170 µL of isopropanol (100%) was added to the separated aqueous phase containing micro-centrifuge tube. Then the mixture was incubated at room temperature for 10 minutes. Then it was centrifuged at 12000 × g for 10 minutes at 4^{o}C. The supernatant was removed from the tube by pipetting. The pellet was washed with 335 µl of 75% ethanol. The sample was centrifuged at 7500×g for 5 minutes at 4^{o}C and the supernatant was discarded by aspirating with a micropipette. The resultant pellet was air dried for 5-10 minutes. The RNA pellet was then re- suspended in nuclease free water. Then the solution was incubated in water bath at 55^{o}C for 10 minutes, and the RNA was stored at -20^{o}C.

2.8 Real time PCR Protocol

2.8.1 Principle

In molecular biology PCR or polymerase chain reaction is used .Real time PCR is used ….

Real-time allow for the detection of PCR amplification during the early phase of the reaction. Measuring the kinetics of the reaction in the early phases of PCR provides a distinct advantage over traditional PCR detection and this is made possible through the use of fluorescent reporter molecules in the reaction. The fluorescent molecule only fluoresces when bound to double stranded DNA (dsDNA) and then after analyzing the melt curves the specificity of the reaction can be determined. In our study we have used SYBR GREEN technology as a fluorescent dye.

2.8.2. Equipments and reagents for Real-time PCR

2.8.2.1 Equipments

- Real-time PCR machine
- Electrophoresis machine
- Pipettes
- Micro-centrifuge tubes
- PCR cabinet
- Refrigerators(-20^{o}C, +4^{o}C)

2.8.2.2 Reagents

The reagents that were used for SYBR GREEN reaction using so Fast Eva Green master mix along with their storage temperatures are given in table 2.3

Table 2.3: Reagents and storage temperature for real time PCR

Reagents	Storage temperature
Precision melt supermix	-20◦C
Forward primer	-20◦C
Reverse prime	-20◦C
RNase/DNase free water	Room temperature

2.8.2.3 Starting materials

Isolated DNA from blood specimen was used as a starting material.

2.8.3. Procedure

2.8.3.1. Primer designing

Six primer sets were designed by using primer 3 plus tool for mutation analysis, which together covers all 13 exons, some of flanking introns, and also 5´ and 3´untranslated regions. The primer sets were designed based on the human G6PD gene sequence which was retrieved from National Center for Biotechnology (NCBI) database. The primer sets were further checked for its accuracy and specificity by using primer BLAST tool of NCBI to prevent any non-specific binding and amplification. Table 2.4 shows the name and sequence of primers for real-time PCR specific to the G6PD gene along with their mutation, location and product length.

Table 2.4: Primers and Conditions for Screening the G6PD-Deficient Variants by HRM Analysis

Mutation	Primer name	Primer Type	Sequence	Product length
G487A **A493G** **T517C** **C519G**	4A 4B	Sense Antisense	5’-GCAGCTCTGATCCTCACTCC-3’ 5’-GGTTGGACAGCCGGTCA-3’	137
G871A **C1004T** **C1024T**	7A 9B	Sense Antisense	5’-CCCAACTCAACACCCAAGGA-3’ 5’-CTCGAAGGCATCACCTACCA-3’	254
C1311T	10A 10B	Sense Antisense	5’-AGGCAGTGGCATCAGCAAG-3’ 5’-GCAGAAGACGTCCAGGATGAG-3’	88
t93c	12A 12B	Sense Antisense	5’-GCCCTCCCTCCCTGTGTG-3’ 5’CAGCTCAATCTGGTGCAGCAGT-3’	111

2.8.3.2 Master Mix preparation

For SYBR Green reactions using Precision melt supermix, the protocols that accompany the master mix kit is followed. The amount of master mix is fixed (the precision melt super mix is 2x, so for a 20 µL reaction volume, we will have to use 10 µL and for 10 µL reaction volume we will have to use 5 µL), the amount of primers, nuclease-free water, and DNA

template will vary. The following table provides the amount of reagents for 10 µL volume reaction.

Table 2.5: Volume of reagents for Real-time PCR

Reagents	Volume per reaction
Precision melt super mix	5.0 µL
Forward primer	0.2 µL
Reverse primer	0.2 µL
DNA	1 µL
Nuclease free water	Up to 10 µL
Total Volume	**10 µL**

Before starting all the components should be thawed and mixed well at room temperature. The gene specific forward and reverse primer pair is used in each reaction. For more accurate replication, first, correct volume of supermix, primers and water was pipette into a microfuge tube to make 9 µL volumes and mixed. Then appropriate volume of the mixture was transferred to the bottom of each well of the PCR plate, and then the DNA template is added to each well.

The PCR plate was then centrifuged at 4000rpm for 1-3 min to remove bubbles and assure that all of the components are mixed at the bottom of the well, and the plate was placed in the machine.

The following thermal cycling profile was used.

1. 95.0°C for 3 minutes, 1 cycle
2. 95.0°C for 10 seconds, 1 cycle
3. 60.0 °C for 10 seconds, 1 cycle
4. 72.0◦C for 15 seconds, 1 cycle
   +Plate Read
5. GO TO 2, 39 more times
6. 95.0°C for 3 minutes, 1cycle
7. 60.0°C for 1 minute, 1 cycle
8. Melt Curve 65.0°C to 95.0°C: Increment 0.2°C 0:10

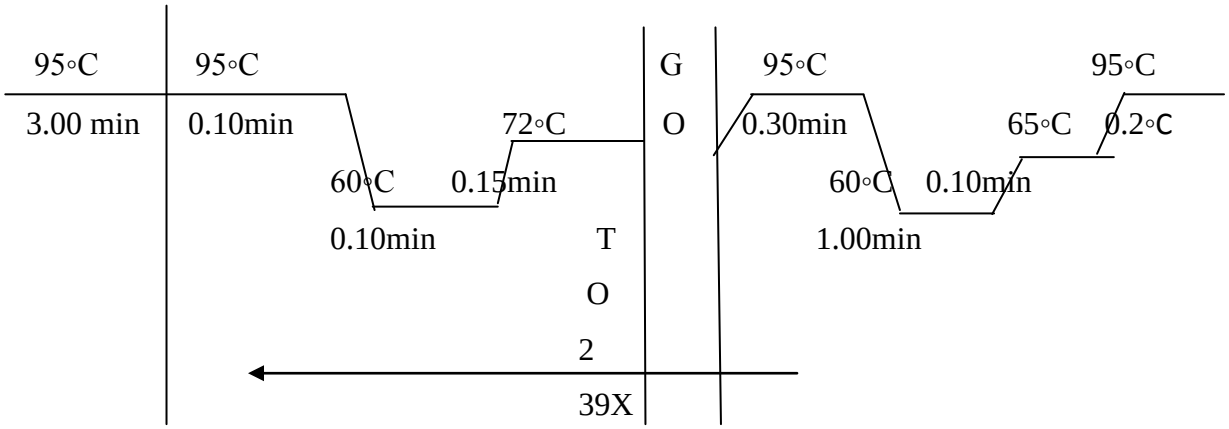

Figure 2.4: Thermal cycling Protocol for Real time PCR-HRM

2.9 Polymerase Chain Reaction (PCR)

2.9.1 Principle

In molecular biology, polymerase chain reaction (PCR) is a technique used for polymerization of nucleotides in a nucleic acid sequence and for amplification of a single copy or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence of interest. The nucleic acid sequence of interest is amplified using specific primers, which can later be used for analysis and for drawing a conclusion. By using PCR, it is possible to produce 236 copies of desired DNA fragments from a single template DNA after 36 cycles. Thus, it is possible to amplify DNA copy numbers from a very small amount of specimen (Figure 2.5).

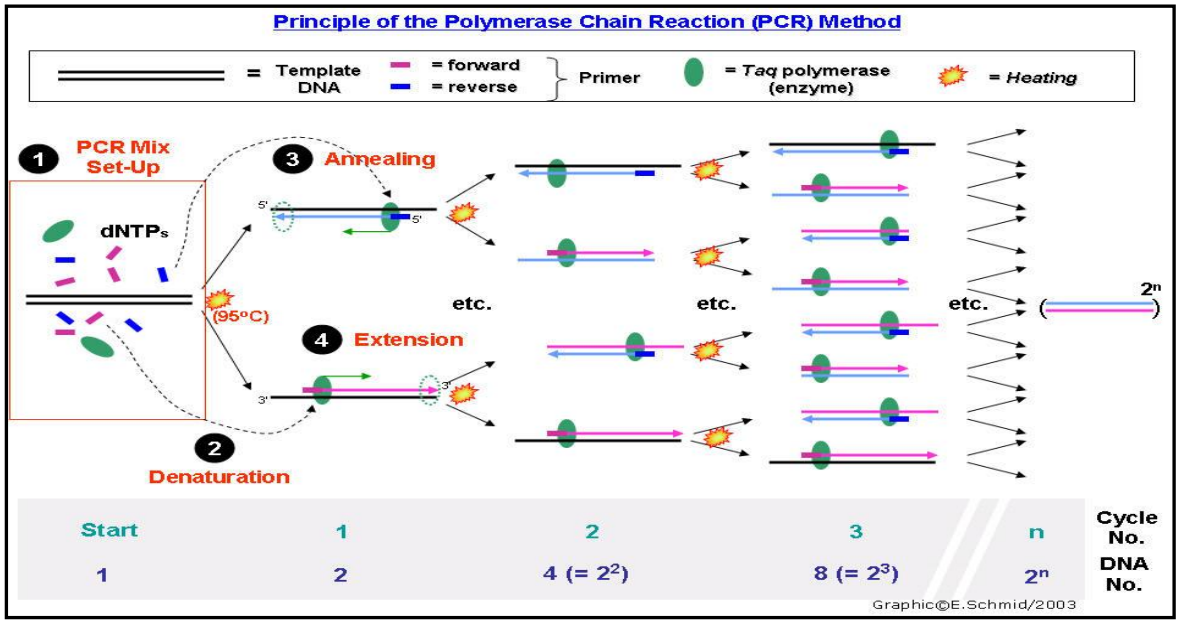

Figure 2.5: Principle of Polymerase chain reaction (source:Graphic©E.Schmid.2003)

2.9.2 Equipment and Supplies for polymerase chain reaction.

2.9.2.1 Equipment

- PCR machine
- Electrophoresis machine
- Pipettes
- Micro-centrifuge tubes
- PCR cabinet
- Refrigerators (-20oC, +4oC)

2.9.2.2 Reagents

The reagents that were used for polymerase chain reaction are given along with their storage temperature in Table 2.6.

Table 2.6:Reagents used for PCR and their storage temperature

Reagents	Temperature
10X Buffer (MgCl2)	-20◦C
25 mM MgCl2	-20◦C
2.5 mM dNTPs	-20◦C
Primer (Forward and reverse)	-20◦C
Taq polymerase (Takara Taq and Hot start Taq)	-20◦C
Q-solution (provided with Hot tart Taq polymerase)	-20◦C
Template DNA	-20◦C
Nuclease-free water	Room temperature

2.8.2.4 Starting materials

Isolated DNA from whole blood was used as starting material.

2.9.3 Procedure

2.9.3.1 Primer designing

The DNA sequence of human G6PD gene was retrieved from the nucleotide database of National Center for Biotechnology Information (NCBI). The retrieved FASTA sequence was then used to design six sets of G6PD gene specific primer that together could cover all 13 exons, some of flanking introns, and also 5′ and 3′untranslated regions (UTR) using the Primer3 tool (Untergasser et al., 2012; Koressaar et al., 2007). The specificity of the designed primers was confirmed by further analysis using the primer BLAST tool of NCBI (Stephen et al., 1990) to prevent any non-specific binding and amplification. Table 2.7 shows the name and sequence of primers specific to the G6PD gene along with their melting temperature (Tm), guanine-cytosine content (GC%), product length from DNA, and product length from RNA (SN means serial number).

Table 2.7: List of primers specific to the G6PD gene along with sequence, Tm, product length from DNA and product length from RNA.

SN	Primer Name	Primer sequence	GC%	Tm	Product length from RNA (bp)	Product length from DNA (bp)
1st set	G6PD_EX1F2	AAGCCGGCGAGAAGTGTGAGG	61.9	64.97	740	2476
	G6PD_Ex6R1	GCACCATGAGGTTCTGCACCAT	54.55	63.43		
2nd set	G6PD_Ex5F	CTACGAGGCCGTCACCAAGAAC	59.09	63.15	739	2538
	G6PD_Ex10R	GATCACCAGCTCGTTGCGCTTG	59.09	64.79		
3rd set	G6PD_Ex9F	CACTTTTGCAGCCGTCGTCCTC	59.09	64.43	642	1087
	G6PD_Ex13R1	GTGCAGCTGAGGTCAATGGTCC	59.09	63.70		

2.9.3.2 PCR optimization

In case of Ex9F-Ex13R1 and Ex9F-Ex13R1primer sets, DNA was used for amplification. However, in case of the primer set Ex5F-Ex10R, the DNA sequence from exon-1 to exon-10 being too huge to amplify, DNA was first converted to mRNA sequence which was then converted to cDNA sequence before amplification.

PCR with G6PD Ex9F and Ex13R1 primers and Hot start Taq DNA polymerase from DNA

PCR was performed with G6PD Ex9F and Ex13R1 primers using a T100TM thermal cycler (Bio-Rad). The final reaction volume was 10 µL containing 1.0 µL PCR buffer (1 x PCR buffer with MgCl2 salt), 0.25 µL MgCl2 (25mM), 2.0 µL Q-solution, 1.6 µL dNTPs mixture (2.5 mM), 0.2 µL forward (Ex9F) and 0.2 µL reverse (EX13R1) primers (10mM), 0.05 µL Hot Start Taq DNA Polymerase (Cat. No. 203203) and 90 ng of genomic DNA, total volume was made 10 µL with nuclease-free water (Table 2.8).

Table 2.8: Reagents used for PCR with G6PD Ex9F and Ex13R1 primers and Hot Start Taq polymerase from DNA

Reagents	Final volume 10 µL
10X Buffer (MgCl2)	1.0
25 mM MgCl2	0.25
2.5 mM dNTPs	1.6
Q-solution	2.0
Forward Primer(Ex9F)	0.2
Reverse Primer(Ex13R1)	0.2
H2O	3.95
Template	0.75
Hot Start Taq DNA polymerase	0.05

The optimized conditions used for PCR were: pre-denaturation at 95°C for 15 minutes, 35 cycles of denaturation at 94°C for 45 seconds, annealing at 59°C for 30 seconds and extension at 72°C for 1 minute 20 seconds; and a final extension at 72°C for 10 minutes (Table 2.9).

Table 2.9: PCR conditions and cycle for G6PD primers

Annotation	Steps	Temperature	Time	Cycle
1	Initial denaturation	95◦C	15 mins	
2	Denaturation	94◦C	45sec	35
3	Annealing	59◦C	30sec	
4	Elongation	72◦C	1min20sec	
5	Final Elongation	72◦C	10 mins	

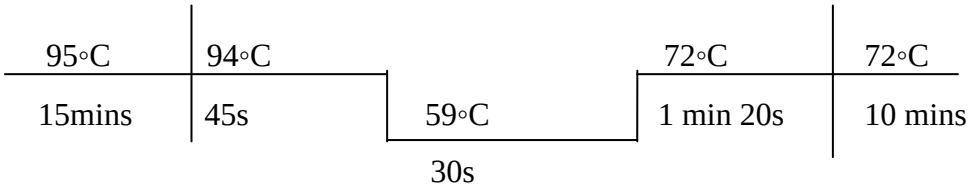

Figure 2.6: Thermal cycling profile for G6PD primers

2.10 Reverse transcription and polymerase chain reaction

2.10.1 Principle

Reverse transcription is a process of producing cDNA from mRNA using random hexamer or oligo-dT primer. After production of cDNA, mRNA from cDNA-mRNA duplex is degraded by RNase H. Then using gene specific primer, PCR is accomplished to amplify the cDNA.

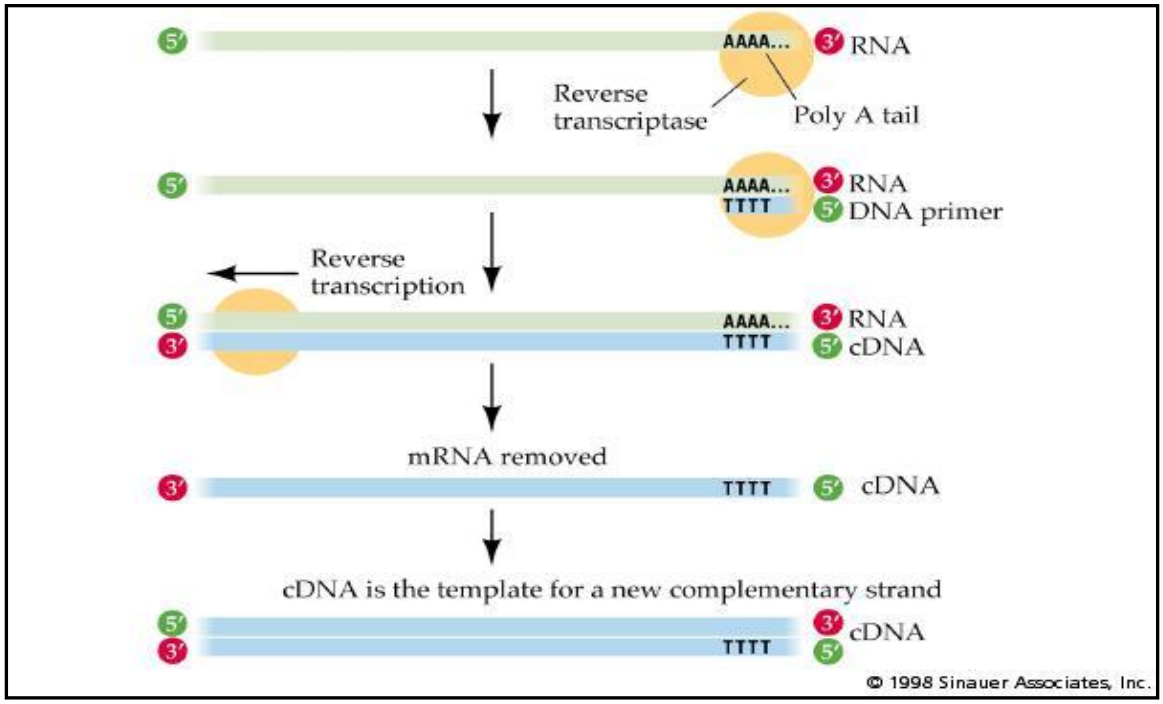

Figure 2.7: Reverse transcription from PCR mRNA to DNA.

2.10.2 Equipment and reagents for reverse transcription and polymerase chain reaction

2.10.2.1 Equipment

- PCR machine
- PCR hood
- Micro-centrifuge tubes
- Pipettes
- Tips
- Electrophoresis machine
- Gel documentation system
- Refrigerator (-20oC)

2.10.2.2 Reagents and storage temperature

The reagents for reverse transcription from mRNA are given along with their storage temperature in Table 2.10.

Table 2.10: Reagents used for Reverse transcription from mRNA

Reagents	Storage temperature
10X Buffer (MgCl2)	-20^{o}C
25 mM MgCl2	-20^{o}C
2.5 mM dNTPs	-20^{o}C
Primer (forward and reverse)	-20^{o}C
Taq polymerase (Takara Taq or Hot Start Taq	-20^{o}C
Reverse transcriptase	-20^{o}C
Nuclease-free water	Room temperature

2.10.2.3 Starting materials

Isolated mRNA from blood specimen was used as starting materials.

2.10.4 Procedure

2.10.4.1 First-Strand cDNA Synthesis

Each component was mixed and briefly centrifuged before use. First strand cDNA synthesis was performed mixing following components in a 0.5 mL tube: 1 µL Oligo-dT (50µM,

Invitrogen, Cat. No. 12574), 1 µL dNTPs mixture (10 mM) and >500ng of total RNA. The final volume was made 10 µL with nuclease-free water (Table 2.11).

Table 2.11: Reagents used for first strand cDNA synthesis

Components	Amounts
Total RNA	>500 ng
Oligo dT (50µM) or Random hexamer (50 ng/µl)	1 µL
10mM dNTP mixture	1 µL
Nuclease-free water	Upto 10 µL

The mixture was incubated at 65°C for 5 minutes in a T100TM thermal cycler (Bio-Rad, USA) and then placed on ice for 1 minute. cDNA synthesis mix was prepared by adding each component in the indicated order: 2.0 µL RT buffer (10X, Invitrogen, Cat. No. 12574-018), 4 µL MgCl2 (25 mM), 2 µL dithiothreitol (0.1 M), 1 µL RNaseOUT (40 U/µL, Invitrogen, Cat. No. 12574-018) and 1 µL SuperScript® III RT (200 U/µL, Invitrogen, Cat. No. 12574-018) (Table 2.12).

Table 2.12: Reagents used for preparation of cDNA synthesis mix

Component	Amounts
10xRT buffer	2
25mM MgCl2	4
0.1 M DTT	2
RNaseOUT (40 U/µL)	1
SuperScript® III RT (200 U/µL)	1

10 µL of cDNA synthesis mix was added to each RNA primer mixture, mixed gently, spun briefly to collect all the solution at the bottom of the tube. Then it was incubated for 50 minutes at 50°C (when Oligo dT was primed) or 10 minutes at 25oC followed by 50 minutes at 50oC (when Random hexamer was primed). The reaction was terminated at 85°C for 5 minutes and chilled on ice. The reactions were collected following brief centrifugation and 1 µL of RNase H was added to each tube which was then incubated for 20 minutes at 37°C. The synthesized cDNA was then used immediately or stored at -20°C.

PCR with G6PD Ex1F2 and EX6R primers with Hot start Taq polymerase from cDNA

PCR was performed with G6PD Ex1F2 and Ex6R primers using a T100TM thermal cycler (Bio-Rad, USA). The final reaction volume was 10 µL containing 1.0 µL PCR buffer (10X PCR buffer with MgCl2 salt, Clontech, Code No. R001A), 0.25 µL MgCl2 (25mM), 1.6 µL dNTPs mixture (2.5 mM), 2 µL Q-solution, 0.2 µL forward (Ex1F2) and 0.2 µL reverse (Ex6R) primers (10mM), 0.05 µL of Hot Start Taq™ DNA Polymerase (5 U/µL, Clontech, Code No. R001A) and 90 ng of genomic DNA, total volume was made 10 µL with nuclease-free water (Table 2.13).

Table 2.13: Reagents used for PCR with G6PD primers and Hot Start Taq polymerase from cDNA

Reagent	Final volume 10µl
10X Buffer (MgCl2)	1.0
25 mM MgCl2	1.0
dNTPs	1.6
Q-solution	2.0
Forward Primer (Ex1F2)	0.2
Reverse Primer (EX6R)	0.2
Taq	0.05
cDNA Template	2.0 (90ng)
H2O	up to 10 uL

The optimized conditions used for PCR were: pre-denaturation at 95°C for 5 min; 35 cycles of denaturation at 95°C for 35 seconds, annealing at 63°C for 40 seconds and extension at 72°C for 1 minute 10 seconds; and a final extension at 72°C for 10 minutes (Table 2.14).

Table 2.14: PCR conditions and cycle for G6PD Ex1F2 and Ex6R primers

Annotation	Steps	Temperature	Time	Cycle
1	Initial denaturation	95^{o}C	15 minutes	
2	Denaturation	95^{o}C	35 seconds	35
3	Annealing	63^{o}C	40 seconds	
4	Elongation	72^{o}C	1:10 minutes	
5	Final elongation	72^{o}C	10 minutes	

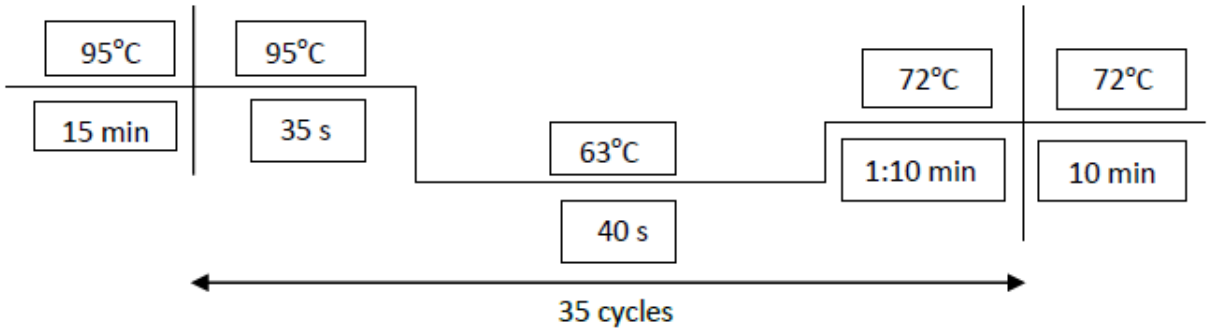

Figure 2.8: Thermal cycling profile for G6PD primer for Ex1F2 and Ex6R primers PCR with G6PD Ex5F and Ex10R primers and Hot Start Taq polymerase from cDNA

PCR was performed with G6PD Ex5F and Ex10R primers using a T100TM thermal cycler (Bio-Rad, USA). 1.0 µL PCR buffer (10X PCR buffer with MgCl2 salt, Clontech, Code No. R001A), 0.25 µL MgCl2 (25mM), 1.6 µL dNTPs mixture (2.5 mM), 2 µL Q-solution, 0.2 µL forward (Ex1F2) and 0.2 µL reverse (Ex6R) primers (10mM), 0.05 µL of Hot Start Taq™ DNA Polymerase (5 U/µL, Clontech, Code No. R001A) and 90 ng of genomic DNA was mixed. The total reaction volume was made 10 µL with nuclease-free water (Table 2.15).

Table 2.15: Reagents used for PCR with G6PD primers and Hot Start Taq polymerase from cDNA

Reagent	Final volume 10µl
10X Buffer (MgCl2)	1.0
25 mM MgCl2	1.0
dNTPs	1.6
Q-solution	2.0
Forward Primer (Ex5F)	0.2
Reverse Primer (Ex10R)	0.2
Taq	0.05
cDNA Template	2.0 (90ng)
H2O	up to 10 uL

The optimized conditions used for PCR were: pre-denaturation at 95°C for 5 min; 35 cycles of denaturation at 95°C for 35 seconds, annealing at 63°C for 40 seconds and extension at 72°C for 1 minute 10 seconds; and a final extension at 72°C for 10 minutes (Table 2.16).

Table 2.16: PCR conditions and cycle for G6PD Ex5F and Ex10R primers

Annotation	Steps	Temperature	Time	Cycle
1	Initial denaturation	95°C	5 minutes	
2	Denaturation	95°C	35 seconds	35
3	Annealing	63°C	40 seconds	
4	Elongation	72°C	1:10 minutes	
5	Final elongation	72°C	10 minutes	

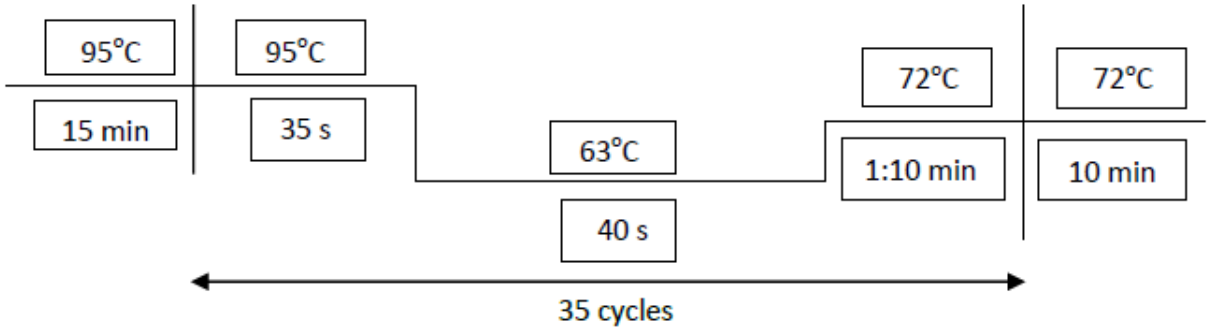

Figure 2.9: Thermal cycling profile for G6PD primer for Ex5F and Ex10R primers

2.11 PCR Purification

2.11.1 Principle

MinElute® PCR purification kit can be used to purify 5 µg PCR products (70 bp – 4 kb) using the column-based method. When sample is loaded on a spin column, dsDNA binds to the column. The adsorption of DNA to the membrane is efficient only in acidic condition at a pH < 7.5 that is maintained by the addition of binding buffer. An optional pH indicator added to the binding buffer allows easy determination of the optimal pH for DNA binding to the spin column. The binding buffer contains chaotropic salts, which disrupts protein structure by destabilizing hydrophobic interactions and removes all the DNA-binding proteins (such as -polymerase) and other contaminants. The high salt also decreases the negative charge on the DNA, allowing stronger interactions with the column membrane. Ethanol is added to the washing buffer addition of which efficiently removes dNTPs, primer dimers (which are usually around 70bp), salt, buffer or anything except dsDNA from the column. Centrifugation is done for complete removal of the residual ethanol from buffer PE. Finally, the neutral pH and higher affinity of nuclease-free water for dsDNA is utilized to elute dsDNA or PCR

products from the column. All centrifugation steps are carried out at 17,900 x g (13000 rpm) in a conventional micro-centrifuge at room temperature (15-25oC, usually 21oC) (Source**:** Instruction manual for MinElute® PCR Purification Kit). Figure 2.10 shows the steps for purification of PCR products (Figure reprinted from www.bioneer.com.au).

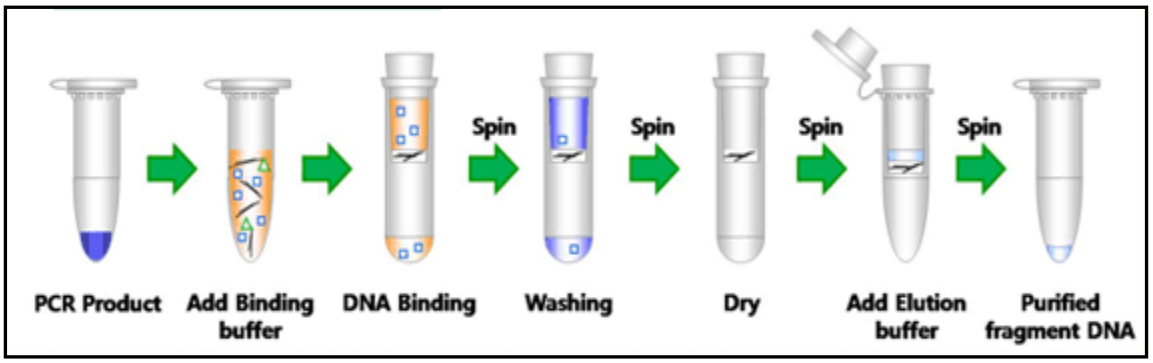

Figure 2.10: Purification of PCR products

2.11.2 Equipment and reagents for purification of PCR products

Equipment

- Biosafety cabinet (Level II)
- Micro-centrifuge
- Water bath
- Spectrophotometer or EON ELISA reader (BioTek, USA)
- Refrigerator (-70oC)

Materials

- MinElute Column
- 1 mL Pipettes
- 20-200 µL Pipettes
- 1-20 µL Pipettes
- Take3 plate (BioTek, USA)

Reagents

- MinElute® PCR Purification Kit (Cat. No. 28004 and 28006, QIAGEN, Germany)

- Buffer PB

- Buffer PE

- 3 M Sodium acetate (pH=5.0)

- Nuclease-free water

All reagents were stored according to the instructions on the label, given in Table 2.17.

Table 2.17: Storage temperature of Reagents and Materials used for PCR Purification

Name of Reagent and Material	Storage temperature
MinElute PCR Purification Kit	Room Temperature (15-25oC)
Spin Columns	2-8oC

2.11.3 Procedure

All the PCR products were spun and then transferred to an Eppendorf tube. 1:250 volume pH indicator-I was added to the Buffer PB. pH indicator I was added to the entire buffer contents, not to buffer aliquots. Ethanol (96-100%) was added to Buffer PE concentrate before use according to the bottle label for volume. Five volumes of Buffer PB were added to 1 volume of the PCR products and mixed properly. It was ensured so that the color of the mixture was yellow (similar to Buffer PB). If the color of the mixture was orange/violet, 10 µL 3 M sodium acetate (pH 5.0) was added and mixed well by pipetting up and then inverting. The mixture was transferred to a MinElute Column that was placed in a 2 mL collection tube. The preparation was allowed to stand at room temperature for 5 mins so that it could pass through the column properly. Then it was centrifuged at 17,900 × g for 1 min. The flow-through was discarded, and the MinElute Column was placed back into the same collection tube. 750 µL Buffer PE was used to wash the column. The tube was allowed to stand at room temperature for 5 minutes and then centrifuged for 1 min. The flow-through was discarded and the column was placed back into the same collection tube followed by another centrifugation for another 1 min. Each column was then placed in a new 1.5 mL micro-centrifuge tube. Depending n the amount of PCR products, nuclease-free water was applied to the center of the column membrane for complete elution of the bound DNA and was allowed to stand at room temperature for 5 mins and then centrifuged at 17,900 x g for 1 min. Then the column was discarded as the purified PCR products were in the flow-through. The amount and purity of the dsDNA were measured using Take3 plate and spectrophotometer. The purified template were stored in the refrigerator (-70°C) before downstream analysis.

2.12 Sequencing

2.12.1 Principle

Fluorescence-based automated cycle sequencing entails a DNA template, a sequencing primer (forward or reverse), a thermal stable DNA polymerase, deoxynucleotide triphosphate (dNTPs), sequencing buffer and fluorescent dye-labeled 2′,3′-dideoxynucleotide triphosphates (ddNTPs). All the components are mixed and subjected to cycles of denaturation, annealing, and extension in a thermal cycler. DNA polymerase incorporates either a dNTP (A, C, G, or T) or the corresponding ddNTPs (nucleotide base analogs that lack the 3′-hydroxyl group essential for phosphodiester bond formation) at each step of chain extension depending on the relative concentration of both molecules. When a dNTP is added to the 3′ end, chain extension is continued. However, when a ddNTP terminator (either ddA, ddC, ddG, or ddT each tagged with a different fluorescent dye) is added to the 3´ end, chain elongation is terminated, forming labeled extension products of various lengths. Figure 2.11 shows the general overview of dye terminator cycle sequencing.

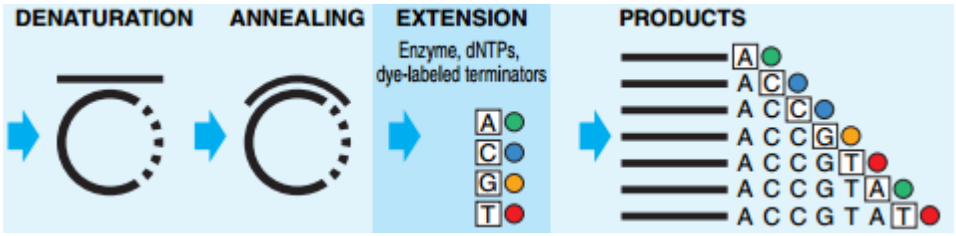

Figure 2.11: Steps of sequencing of PCR product

For automated separation and analysis of fluorescently labeled DNA fragments, the products are subjected to capillary electrophoresis that involves a denaturing polymer. It has largely replaced the polyacrylamide gel electrophoresis technique due to significant gains in workflow, throughput, and ease of use as there is no need to pour gels between two glass plates with capillary electrophoresis and more samples can be processed at once. A high voltage charge is applied to the buffered sequencing reaction that forces the negatively charged fragments into the capillaries (Applied Biosystems, 2009). Shortly before reaching the positive electrode, the fluorescently labeled DNA fragments, separated by molecular weight or size, move across the path of a laser beam. When excited by the laser beam, each fluorescent dye on the fragments emits light at a unique wavelength, and all four colors and therefore, all four bases, can be detected in one capillary injection using an optical detection device. Finally, the fluorescence signal is converted to digital data using Data Collection Software.

Big Dye Terminators version 3.1 is formulated with dITP in place of dGTP to reduce peak compressions. They utilize a single dye molecule for efficient transfer of energy. A tremendously efficient energy transfer linker combines an energy donor dye (fluorescein) and an energy acceptor dye (dichlororhodamine).

2.12.2 Reagents and apparatus for sequencing

Apparatus

- Applied Biosystems 310 Genetic Analyzer

- No. of capillaries: 1

- Capillary Array Length: 47 or 61 cm

- Sample capacity: 48 or 96 sample tubes

- Cycle sequencing machine Mastercycler® gradient)

- Centrifuge machine

- 8-tube PCR strip

Reagents

The reagents needed for sequencing are given along with their storage temperature in Table 2.18.

Table 2.18: Reagents needed for sequencing and their storage temperature

Reagents	Storage temperature
Sequencing buffer	4°C
BigDye® Chain Terminator v3.1 Ready Reaction (RR) mix	-20°C
Forward primer or Reverse primer	-20°C
SAM solution (Applied Biosystems, USA)	4°C
X-terminator solution (Applied Biosystems, USA)	4°C

2.12.3 Cycle Sequencing Master-mix Preparation

Purified PCR products for all deficient samples were sent to Institute of Epidemiology, Disease Control, and Research (IEDCR), Dhaka, Bangladesh. Direct sequencing of specific regions of PCR products was done using BigDye Chain Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, USA) and ABI PRISM 310 automated sequencer (Applied Biosystems, USA) according to manufacturers' instruction. For each sample, a

master-mix was prepared in an Eppendorf tube and mixed by vortexing. For each sample, 2.0 µL of 5X sequencing buffer and 0.5 µL of BigDye® chain terminator version 3.1 Ready Reaction mix and 0.3 µL of individual primer (forward or reverse) was added to an 8-tube PCR strip (Table 2.19).

Table 2.19: Reagents and their amounts used for Cycle Sequencing Master-mix Preparation

Reagent	Concentration	Amount of single reaction (µL)
Sequencing buffer	5X	2
BigDye® Chain Terminator v3.1 RR mix	2.5X	0.5
Forward primer or Reverse primer	0.2 µM	0.3

The tubes containing the template were spun, and 10-20 ng/µL (depending on the band size and density) of each of the purified PCR products were added to the 8-tube PCR strip. Then nuclease free water was added to the mixture to make the total volume 10 µL. The PCR tubes were vortexed and centrifuged at 4000 rpm for 3 minutes. Then the PCR strip was placed in the cycle sequencing machine Mastercycler® gradient (Cat. No. 4095-0015, USA Scientific) Thermal Cycler and subjected to following thermal cycling profile: pre-denaturation at 94^{o}C for 1 minute; 25 cycles of denaturation at 94°C for 10 seconds, annealing at 58°C for 5 seconds and extension at 60°C for 4 minutes; and a final extension at 60°C for 10 minutes (Table 2.20).

Table 2.20: Cycle sequencing thermal condition cycle

Stage	Step	Temperature (°C)	Duration	Cycle
I	1 (Initial denaturation)	94	1 min	
II	1 (Denaturation)	94	10 s	25
	2 (Annealing)	58	5 s	
	3 (Extension)	60	4 min	
	Final extension	60	10 min	
III	1 (Hold)	4	Hold	

After the completion of cycle sequencing, products were centrifuged for 2 mins at 4100g. Purification of cycle sequencing products were performed by mixing with 45 µL of SAM solution (Applied Biosystems, USA) and 10 µL of X-terminator solution (Applied Biosystems, USA) in an automated shaker for 30 minutes. Before addition, the X-terminator solution was vortexed properly at maximum speed for at least 30 seconds, until it became

homogenous. As it was difficult to pipette the highly dense X-terminator solution out from the bottom of its container, the narrow outer end of the micropipette tip was cut a little to make the opening a little wider. If any particulate was present in SAM solution, it was heated at 37^{o}C, mixed and cooled to room temperature before using it. The PCR strip was vortexed for 30 mins continuously (1500-2500 rpm). Following centrifugation at 4000 X g for 2minutes, 10 μL supernatant from each tube was transferred to a new 0.2 mL PCR tube strip. The solutions were subjected to capillary electrophoresis through (denaturing) POP-6TM polymer in an automated sequencing machine ABI PRISM® 310 Genetic Analyzer (Sequencing by Sanger Method).

2.12.4 Sequencing analysis

Sequencing results were analyzed by Chromas Lite 2.4 software to identify the sequences that are reliable and conclusive for mutational analysis and to eliminate sequences with noises that give inconclusive and unreliable data. Reliable and conclusive sequences obtained after analysis were used to locate the mutation using nucleotide BLAST tool of the National Center for Biotechnology Information (NCBI). If any mutation was found, then the mutant base was again checked in the Chromas Lite tool to see whether any significant noise was there or not. If no significant noise was present and the peak for nucleotide was conclusive, then the changed base was considered as a mutant base. FASTA format of the sequence was used to perform alignment of the query sequence with wild-type sequence of mRNA or gene stored in NCBI database using BLAST (Basic Local Alignment Search Tool). After performing the alignment of query and wild-type sequence (subject), the BLAST result was scrutinized to see whether there was any mismatch. If any mismatch was found between query and wild-type sequence, and the corresponding base was checked in the .ab1 file on Chromas Lite tool again for quality score and noise below the peak. When no or negligible noise was found and the quality value was above 15, the mismatched base was confirmed as the mutant or substituted base. The substituted base was then taken into the spotlight to see whether it causes any change in the codon. From the BLAST result, the mismatched base was checked for the corresponding position of the mRNA. Then the position in CDS (coding sequence) was recorded and divided by 3 to get the position of amino acid in peptide. This was also in attention whether substitution of base resulted in synonymous codon (different codon but coding same amino acid) or a codon that code for different amino acid.

Chapter 3: Results

3.1 Mutation detection using High Resolution Melt (HRM) curve analysis

3.1.1 Common G6PD gene mutations in Dhaka, Bangladesh

The current study discovered 12 G6PD deficient samples through fluorescent spot test and spectrophotometric enzyme assay. From these samples, three different mutations have been detected and confirmed. These mutations include G6PD Mahidol (c.G487A), Kalyan-Kerala (c.G949A) and Orissa (C131G). Sequencing and BLAST results for all three mutations are shown in figure-3.1 to figure-3.6. As sequencing is a time consuming process, finding an easy but reliable method for mutation detection is of practical importance. This work assessed the potential of real time PCR-HRM analysis to detect the mutations using existing primers.

3.1.1.1 G6PD Orissa sequencing result

Figure-3.1 and figure-3.2 illustrate sequencing and BLAST results respectively for deficient sample harboring Orissa mutation. The base substitution that is responsible for G6PD Orissa mutation correspond to 248th base of mRNA transcript variant-2 (NM_001042351.2) and 131th base in coding sequence.

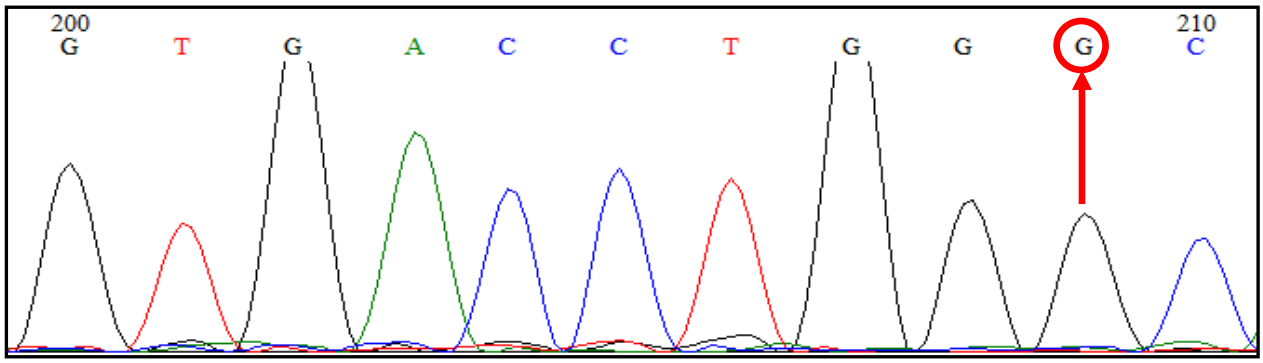

Figure-3.1: Illustration of a chromatogram segment containing mutated base responsible for G6PD Orissa mutation in exon-3 of the deficient sample. Sequencing data and BLAST result reveals a single base substitution (c.C131G) which corresponds to 209th base of chromatogram. The mutated base and corresponding peak are highlighted by a circle and an arrow. In chromatogram the bases are represented in the colors green (adenine), black (guanine), red (thymine), and blue (cytosine).

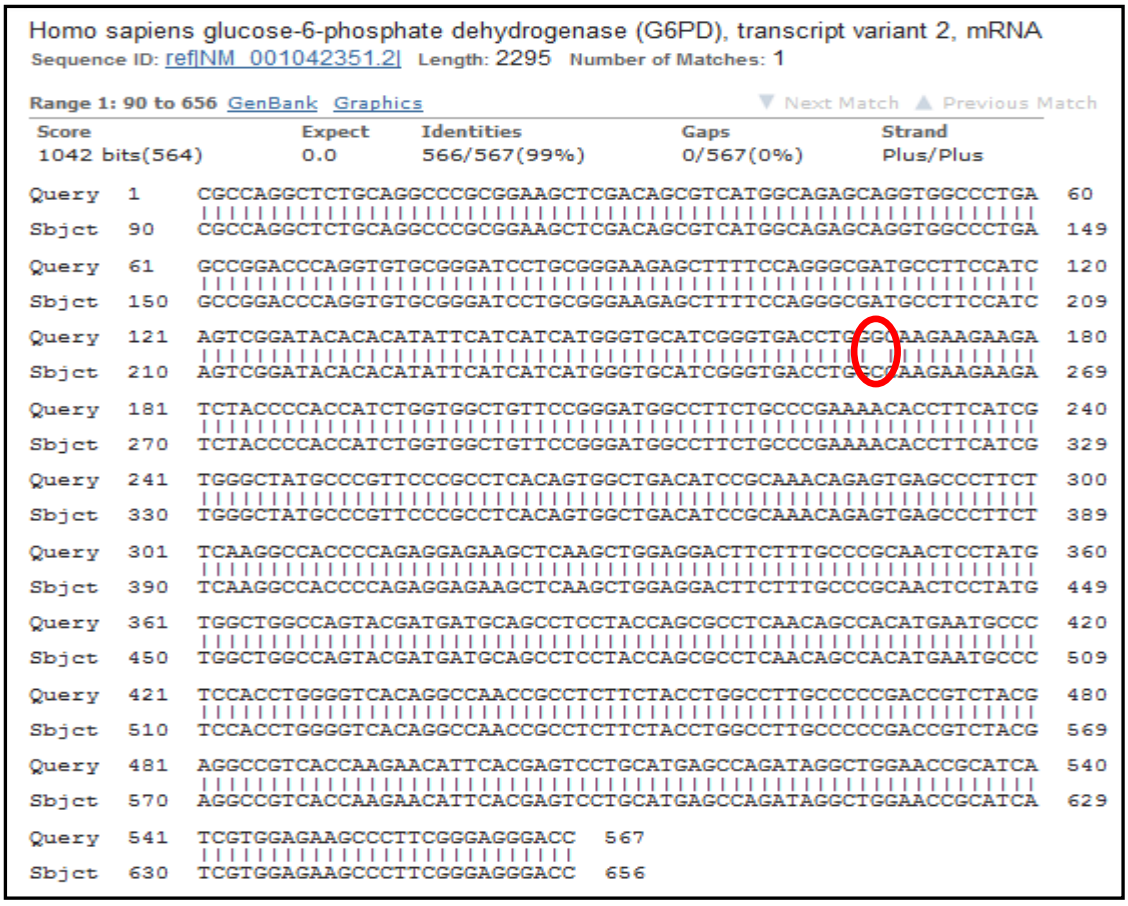

Figure-3.2: Illustration of Blast result for one deficient sample harboring mutation in exon-3. Blast result shows that deficient individual harbors a mutation in exon-3 at 248th base of mRNA (NM_001042351.2) and this mutated base position accounts for coding position 131 (c.131) implying a c.C131G substitution.

The base substitution at 131st position of coding sequence result codon change from: GCC (Alanine) to GGC (Glycine). The resultant amino acid change is p.Ala44Gly in G6PD peptide. The wild type alanine residue is a part of the NADP binding site that spans between residues 38 and 44 (GASGDLA), and so is crucial for the human enzyme, both for activity and for the long-term stability of the enzyme in physiological conditions. So, in case of the G6PD-Orissavariant, the Ala44Gly substitution alters the tertiary structure of the protein and eventually,results in impaired NADP-binding. In addition, a mutation in any of NADP binding site residues dramatically increases the Michaelis constant of NADP (Km NADP). In the G6PD-Orissa variant, the Km of NADP is 5-fold higher than that of the wild-type enzyme, and there is also an increase in the thermo-stability of the enzyme.

3.1.1.2 G6PD Mahidol sequencing result

Figure-3.3 and figure-3.4 illustrate sequencing and BLAST results respectively for deficient sample harboring Mahidol mutation. The base substitution that is responsible for G6PD Mahidol mutation corresponds to 614th base of mRNA transcript variant-2 (NM_001042351.2) and 487th base in coding sequence.

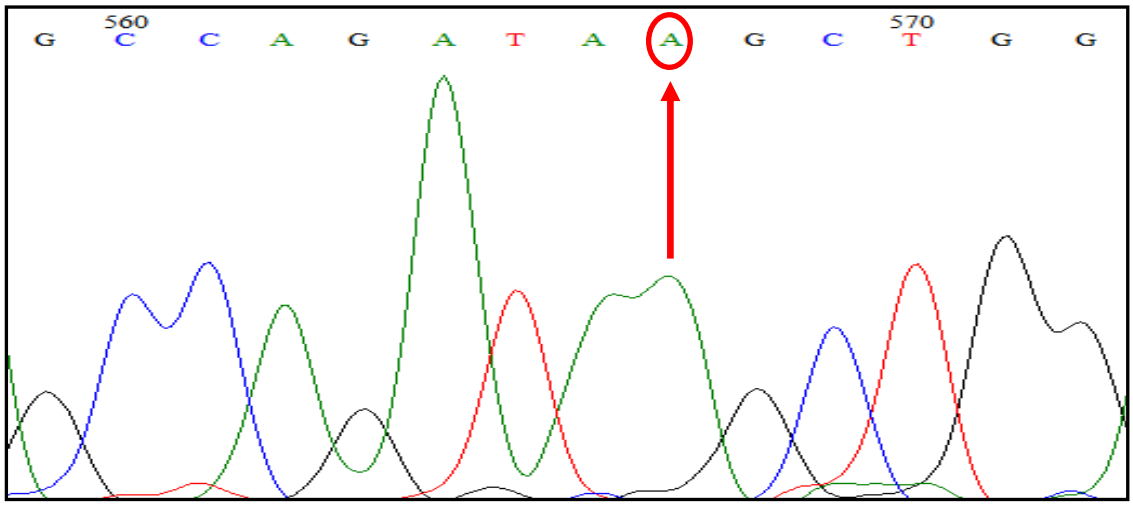

Fig 3.3: Illustration of a chromatogram segment containing mutated base responsible for G6PD Mahidol variant in exon-6 of the deficient sample. Sequencing data and BLAST result reveals a single base substitution (c.G487A) which corresponds to 567th base of chromatogram. The mutated base and corresponding peak are highlighted by a circle and an arrow. In chromatogram the bases are represented in the colors green (adenine), black (guanine), red (thymine) and blue (cytosine).

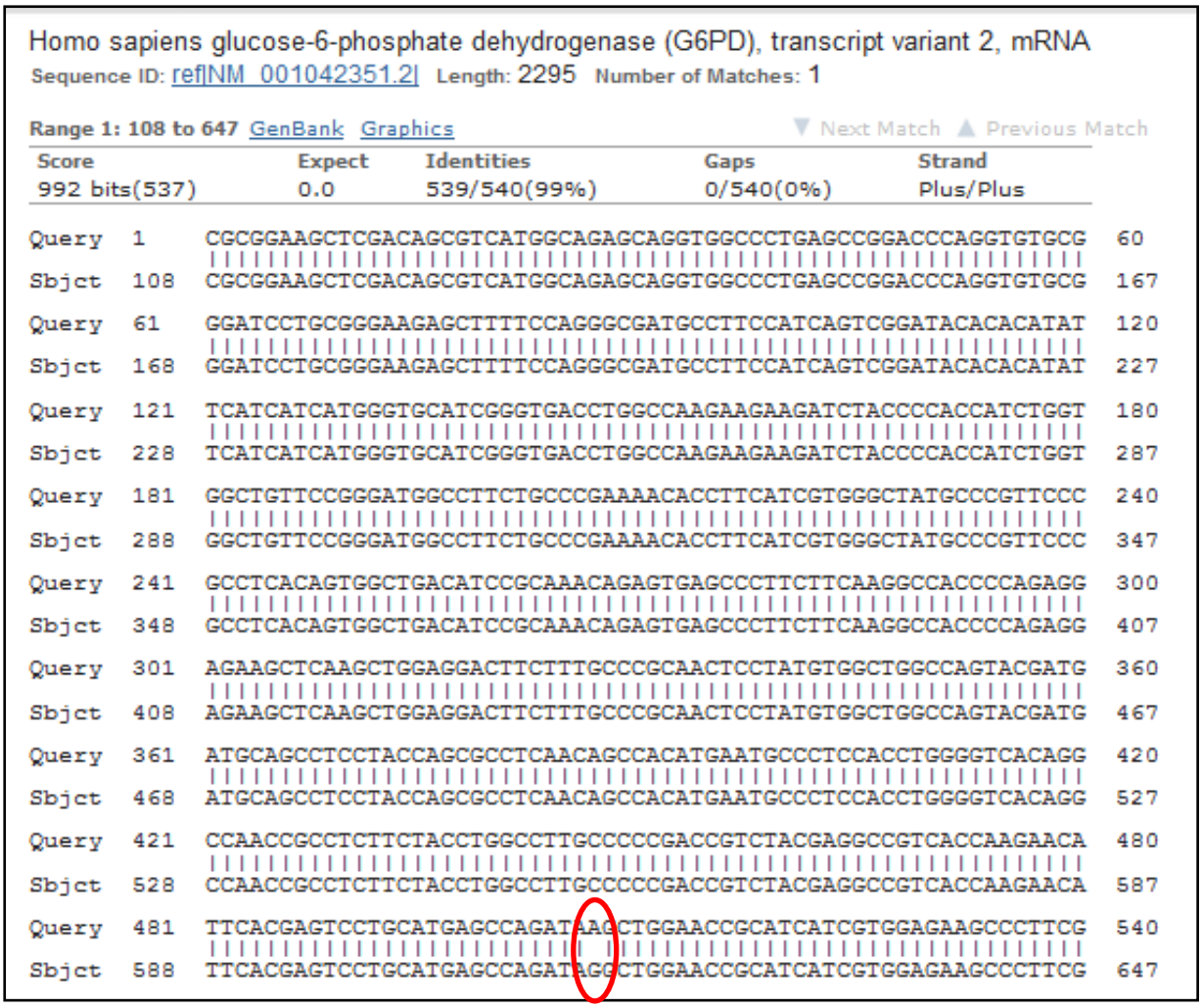

Fig 3.4: Illustration of Blast result for one deficient sample harboring mutation in exon-6. Blast result shows that the deficient individual harbors a mutation in exon-6 at 614th base of mRNA (NM_001042351.2) and this mutated base position accounts for coding position 487 (c.487) implying a c.G487A substitution.

The c.G487A substitution causes change of 163th codon of the G6PD peptide. The wild type codon at position 163 of the peptide is GGC (glycine) which alters to AGC (serine) due to the c.G487A mutation. The p.gly163ser mutation is known as Mahidol mutation and associated with protein stability. As the wild type non-polar amino acid glycine is replaced by polar serine residue, protein structure undergoes disruption and Mahidol G6PD protein becomes less thermostable than wild type G6PD enzyme. Reduction in protein stability accounts for the reduced G6PD enzyme activity in Mahidol deficient.

3.1.1.3 G6PD Kalian-Kerala sequencing result

Figure-3.5 and figure-3.6 illustrate sequencing and BLAST results respectively for deficient sample harboring Kalian-Kerala mutation. The base substitution that is responsible for G6PD Kalyan-Kerala mutation correspond to 1076th base of mRNA transcript variant-2 (NM_001042351.2) and 949th base in coding sequence.

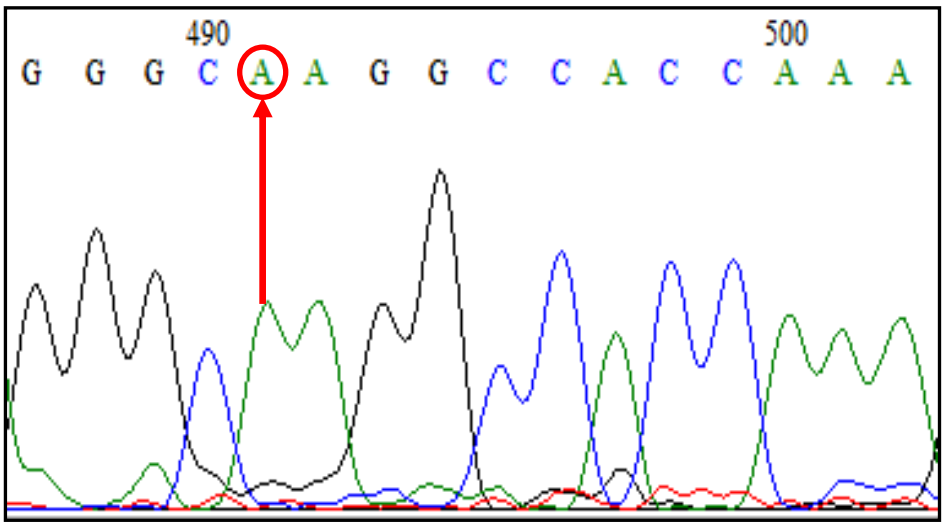

Fig 3.5: Illustration of a chromatogram segment containing mutated base responsible for G6PD Kalian-Kerala variant in exon-9 of the deficient sample. Sequencing data and BLAST result reveals a single base substitution (c.G949A) which corresponds to 491th base of chromatogram. The mutated base and corresponding peak are highlighed by a circle and an arrow. In Chromatogram the bases are represented in the colors green (adenine), black (guanine), red (thymine) and blue (cytosine).

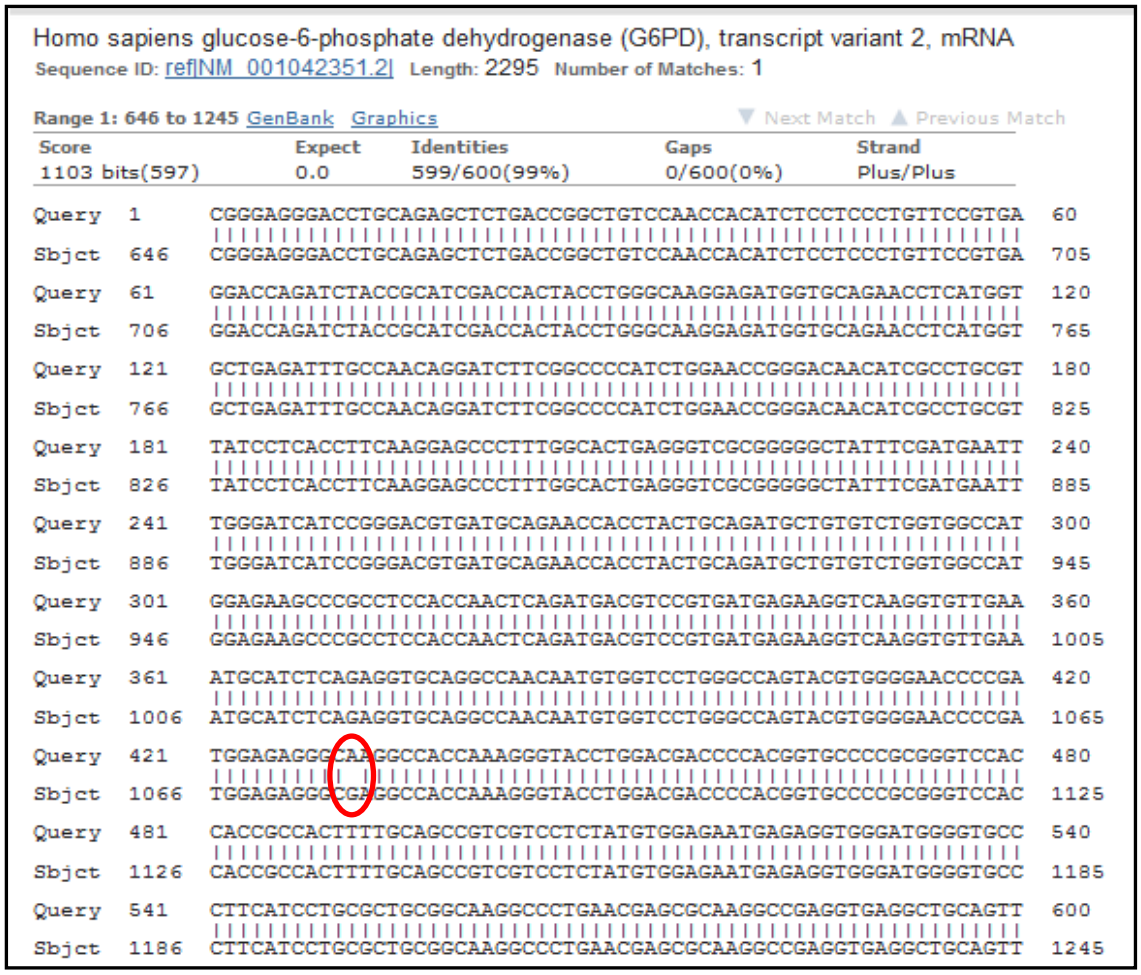

Fig 3.6: Illustration of BLAST results of the deficient sample found which harbors a mutation in exon-9 after sequencing using the EX5F primer. Blast result shows that deficient individual harbors a mutation in exon-9 at 1078th base of mRNA (NM_001042351.2) and this mutated base position accounts for coding position 949 (c.949) implying a c.G949A substitution.

The base substitution (G>A) was found to be at position 949 of the coding sequence (CDS). The c.G949A substitution causes change of 317th amino acid codon of G6PD peptide. The wild type codon at position 317 of peptide is GAG (Glutamic acid) which alters to AAG (Lysine) due to the c.G949A mutation. The p.glu317lys mutation is known as Kalyan-Kerala mutation and associated with protein stability. As the wild type acidic amino acid glutamate is replaced by basic lysine residue, protein structure undergoes disruption and Kalyan-Kerala G6PD protein becomes less stable than wild type G6PD enzyme. Reduction in protein stability accounts for the reduced G6PD enzyme activity in Kalyan-Kerala mutant.

3.1.2 Detection of G6PD Mahidol (c.G487A) mutation in five deficient samples using HRM

Gene sequencing already confirmed G6PD Mahidol (c.G487A) mutation in exon-6 region. Therefore primers (5'-GCAGCTCTGATCCTCACTCC-3' and 5'-GGTTGGACAGCCGGTCA-3') covering G487A mutation were used to analyze exon-6 in deficient samples. Samples that were defined as G6PD normal by sequencing were run in parallel with the known deficient samples to find distinction in melting property between normal and deficient samples. All five c.G487A deficient samples differed by melting curves from normal samples (Figure 3.7-Normalized melt curves and Figure3.8- difference curves).

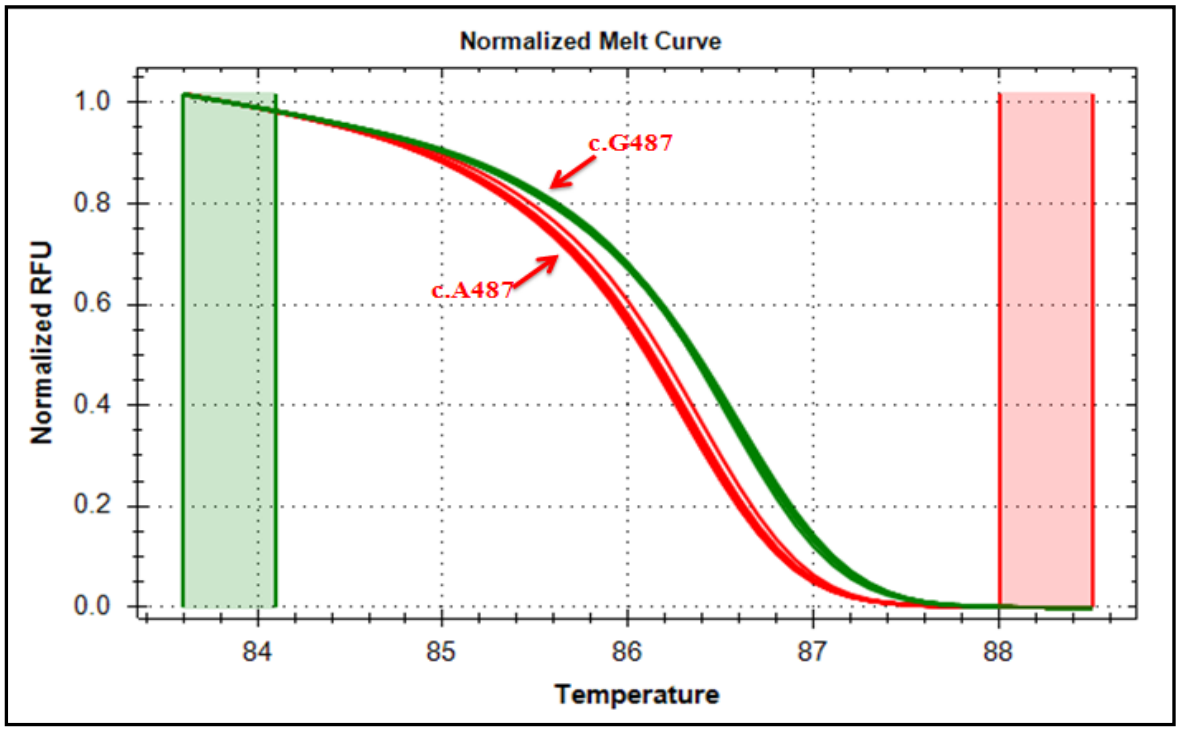

Fig 3.7: Normalized melt curves of G6PD deficient and G6PD normal samples for portion of exon-6. Melt curve effectively distinguished difference between the wild type and normal alleles. Five of the c.G487A deficient samples differed in this region by wild type allele.

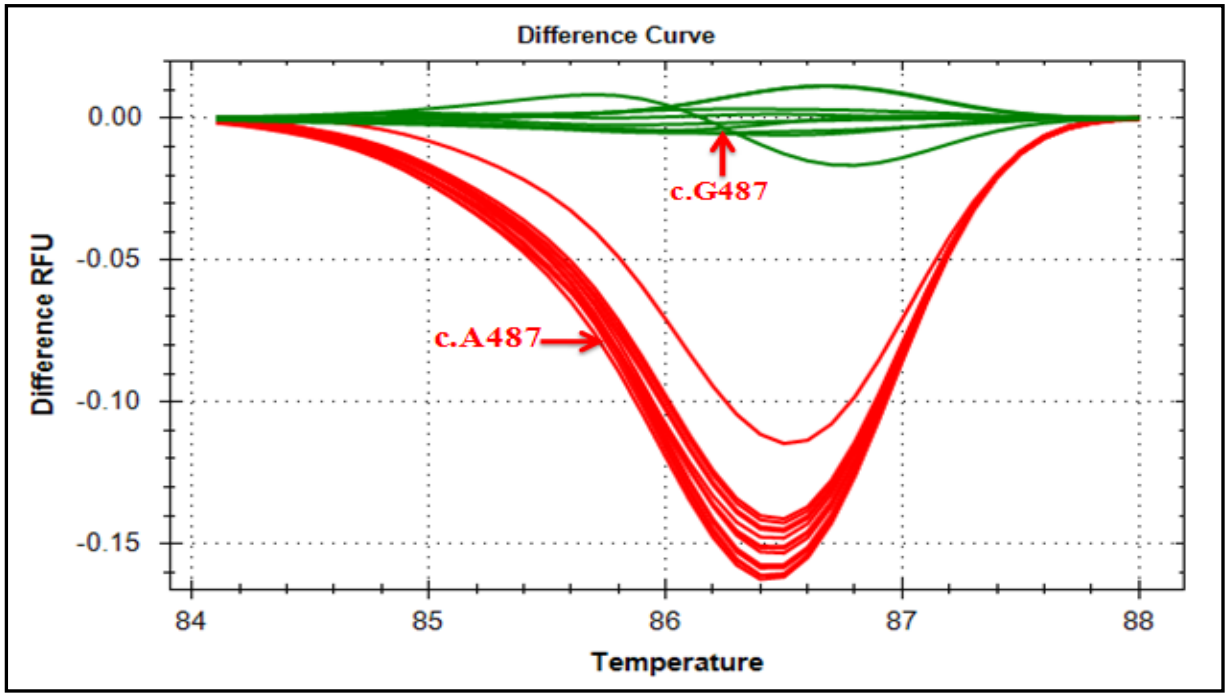

Figure-3.8: Difference curves of deficient and normal samples for portion of exon-6. Discernable changes in relative fluorescence units were observable between the five deficient samples and normal samples.

From HRM analysis (fig 3.7) using deficient samples and non-deficient controls, two types of normalized melting curves were observed, indicating presence of two types of alleles in the target region among samples. All non-deficient controls fell among the green normalized melting curves (Figure-3.7) whereas five deficient samples fell under the red normalized melting curves. HRM result sufficing adenine (A) / thymine (T) to guanine (G) / cytosine (C) transversion (G ↔ T or C ↔ A) or transition (G ↔ A or C↔T).

3.1.3 Detection of G6PD Kalyan-Kerala (c.G949A) mutation in one deficient sample using HRM.

Two primers (5'-CCCAACTCAACACCCAAGGA-3' and 5'-CTCATTCTCCACATAGAGGACGAC-3') covering G871A mutations were used to analyze exon-9 of the c.G949A deficient sample along with samples that were defined as G6PD normal by sequencing. The c.G949A deficient sample differed in melting property from normal samples (Figure 3.9-Normalized melt curve and Figure3.10- difference curve).

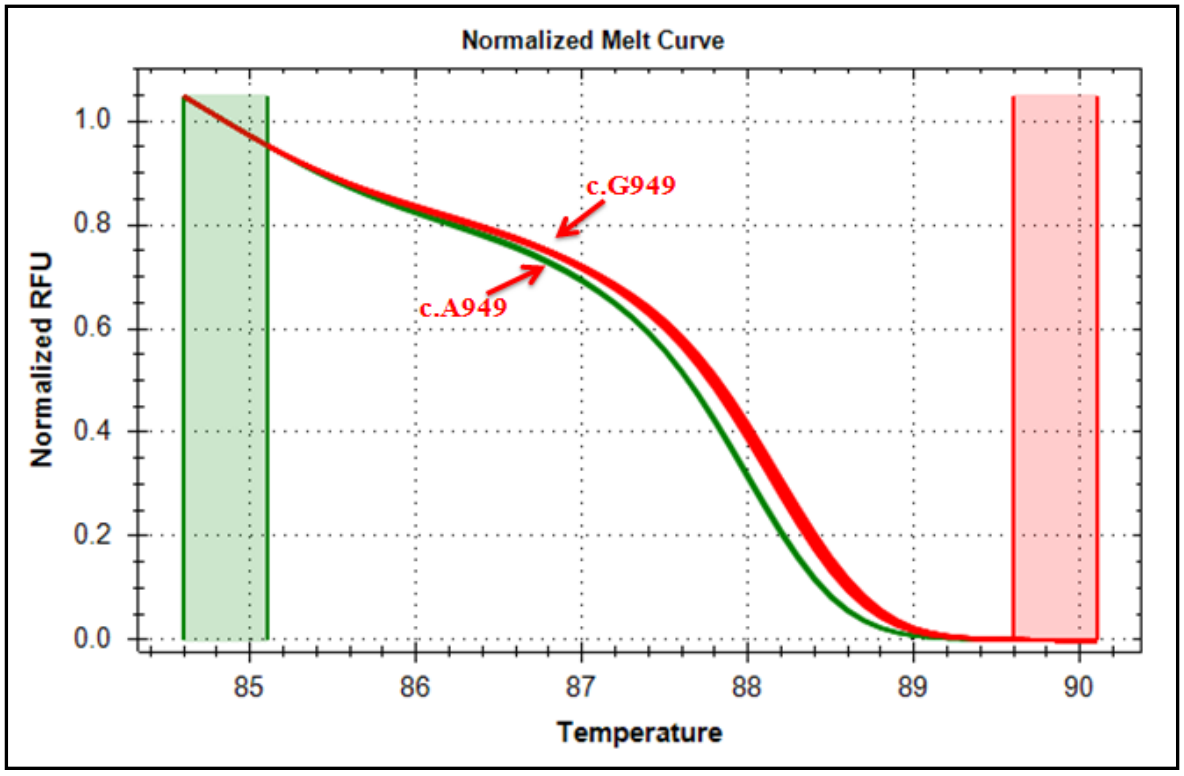

Fig 3.9: Normalized melt curve of G6PD deficient and G6PD normal samples for portion of exon-9 containing c.G949A. Melt curve efficiently distinguish difference between wild type and normal alleles. One of the deficient samples differed in this region by comparison with the wild type allele.

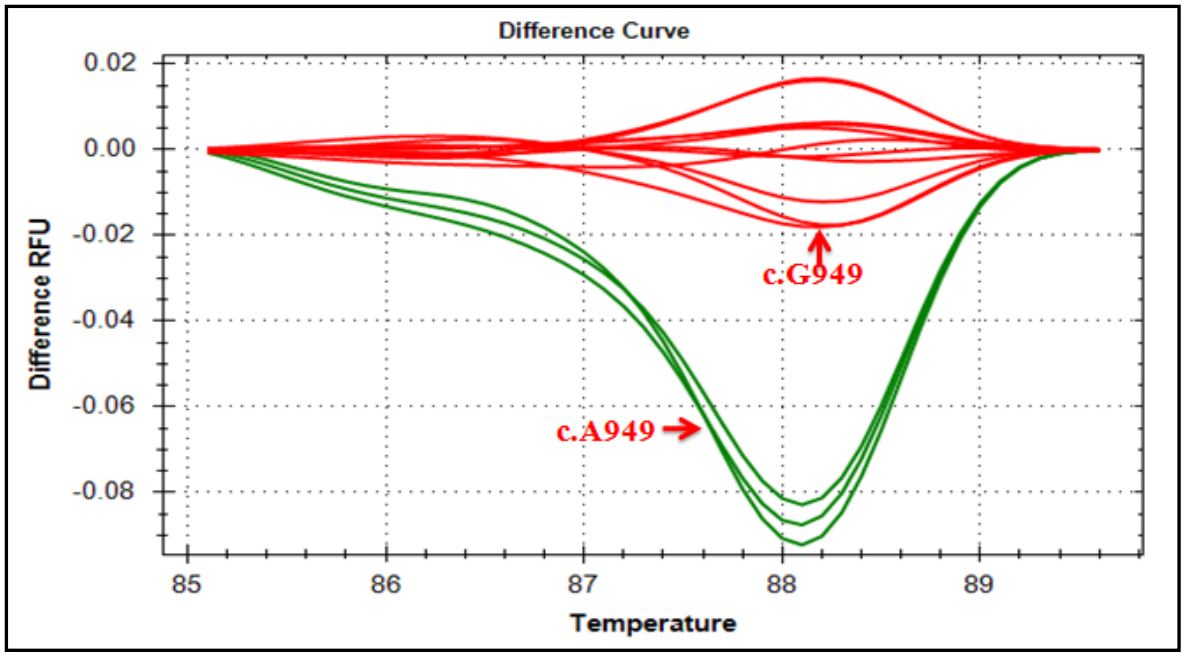

Figure-3.10: Difference curve of deficient and normal samples for portion of exon-9. Discernable change in relative fluorescence unit is observable in one deficient sample and normal samples.

From HRM analysis (fig 3.9) using deficient samples and non-deficient controls, two types of normalized melting curves were observed, indicating presence of two types of alleles in the target region among samples. All non-deficient controls fell among the red normalized melting curves (figure-3.9) whereas the c.G949A deficient samples fell under the green normalized melting curves.

3.1.4 Detection of G6PD Polymorphism in exon-11 or c.1311T/C using HRM

In case of c.1311T/C SNP detection we initially did HRM analysis and then confirmed the alleles by sequencing. Two primers (5'-AGGCAGTGGCATCAGCAAG-3' and 5'-GCAGAAGACGTCCAGGATGAG-3') were used for the detection of G6PD polymorphism in exon-11 of the G6PD gene of study participants which revealed that HRM can distinguish between heterozygous and homozygous states. This suggests that HRM can be employed as a useful method for heterozygous mutation detection which is not possible to diagnosis by conventional screening methods like FST or spectrophotometric G6PD enzyme assay.

Initially hemizygous male samples were run for HRM analysis to see whether HRM can differentiate between c.1311T and c.1311C alleles. Figure-3.11 and figure-3.12 demonstrates two types of products from the target region that was amplified from extracted DNA samples.

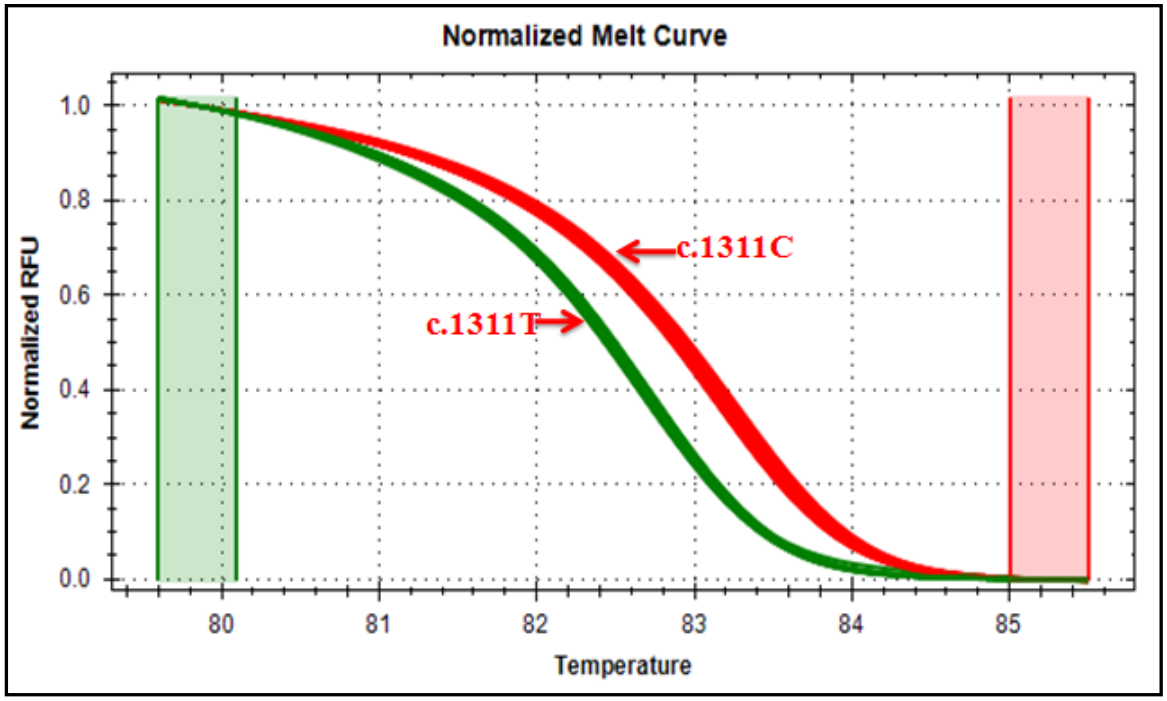

Fig 3.11: Normalized melt curve of hemizygous male participants for analysis of c.1311C/T in exon-11. HRM analysis of hemizygous males produces two types of normalized melt curve.

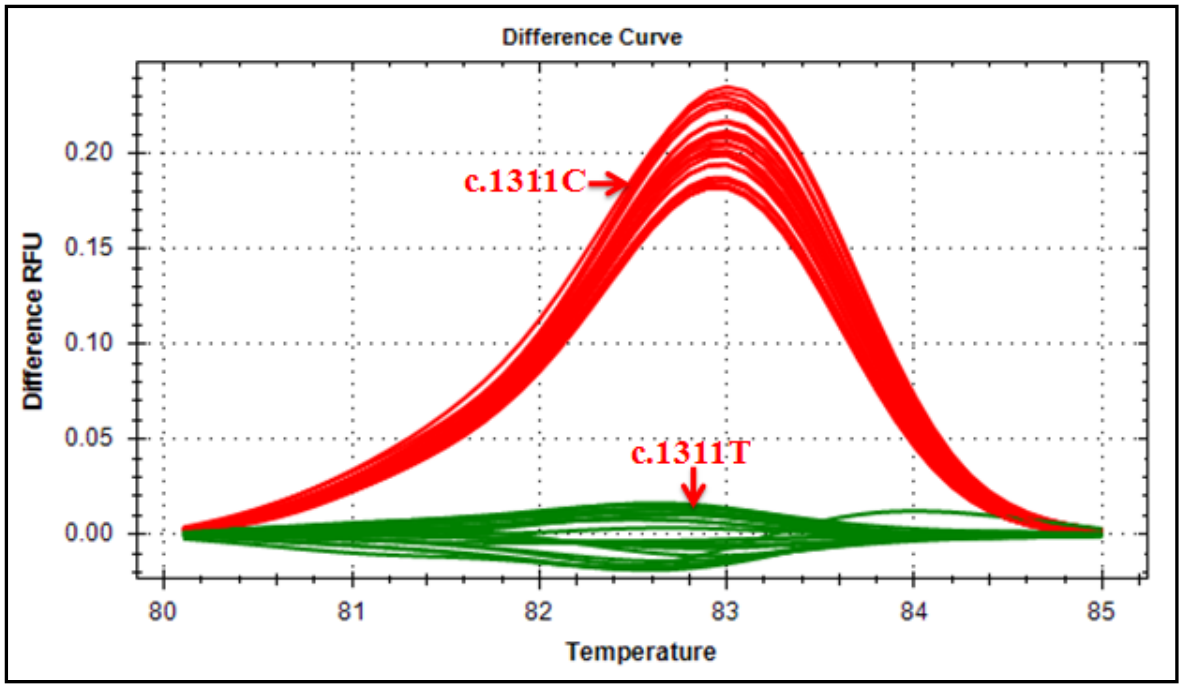

Figure-3.12: Difference curve of hemizygous male samples for analysis of c.1311C/T polymorphism of exon-11. Discernable change in relative fluorescence unit is observable among samples from hemizygous males.

HRM results (figure-3.11 and figure-3.12) clearly indicate difference in amplified fragments. To find out the difference in the alleles, sequencing was performed using samples from each cluster. Sequencing of samples with red normalized melt curve revealed a single base mismatch with the reference G6PD mRNA sequence (NM_001042351.2). Query sequence contained c.1311C allele whereas c.1311T allele was present in the reference sequence.

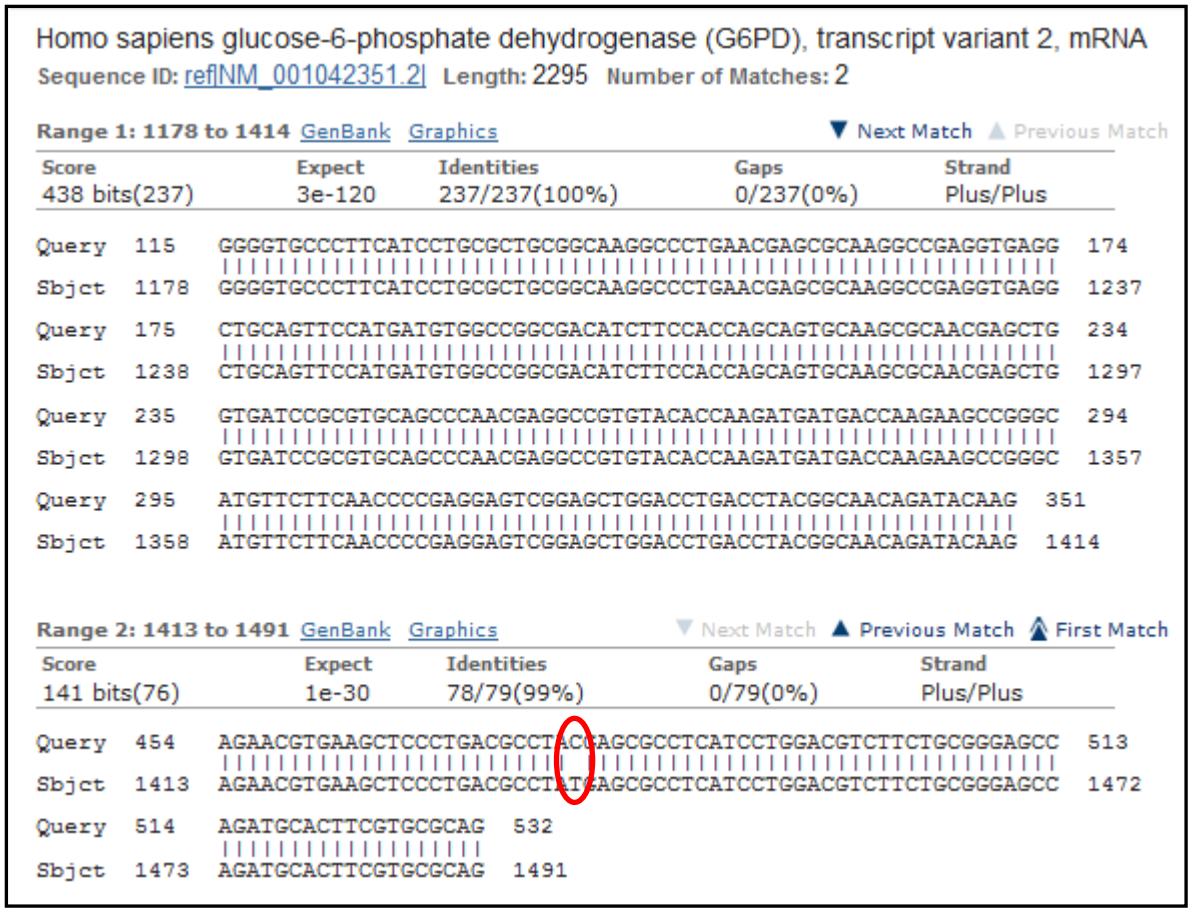

Fig 3.13: BLAST result of a hemizygous male participant with red normalized melt curve. Sequencing data and BLAST result represent c.1311C allele (1438th position of mRNA) in the participant (shown by red circle) falls under red normalized melt curve.

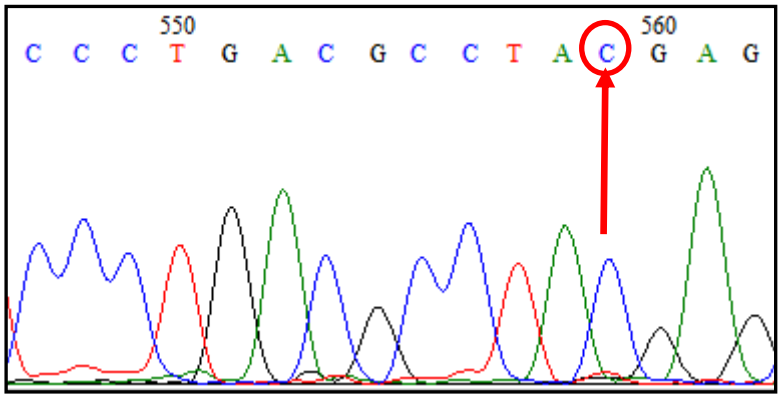

Fig 3.14: Partial chromatogram of sequencing data for exon-11. Mismatched base of BLAST result (figure-24) corresponds to 559th cytosine of chromatogram (479 base of query sequence).

On the other hand, sequencing of samples with green normalized melt curve revealed a c.1311T allele. Figure-3.15 represents the BLAST result, showing no mismatch between query sequence and reference mRNA sequence (NM_001042351.2).

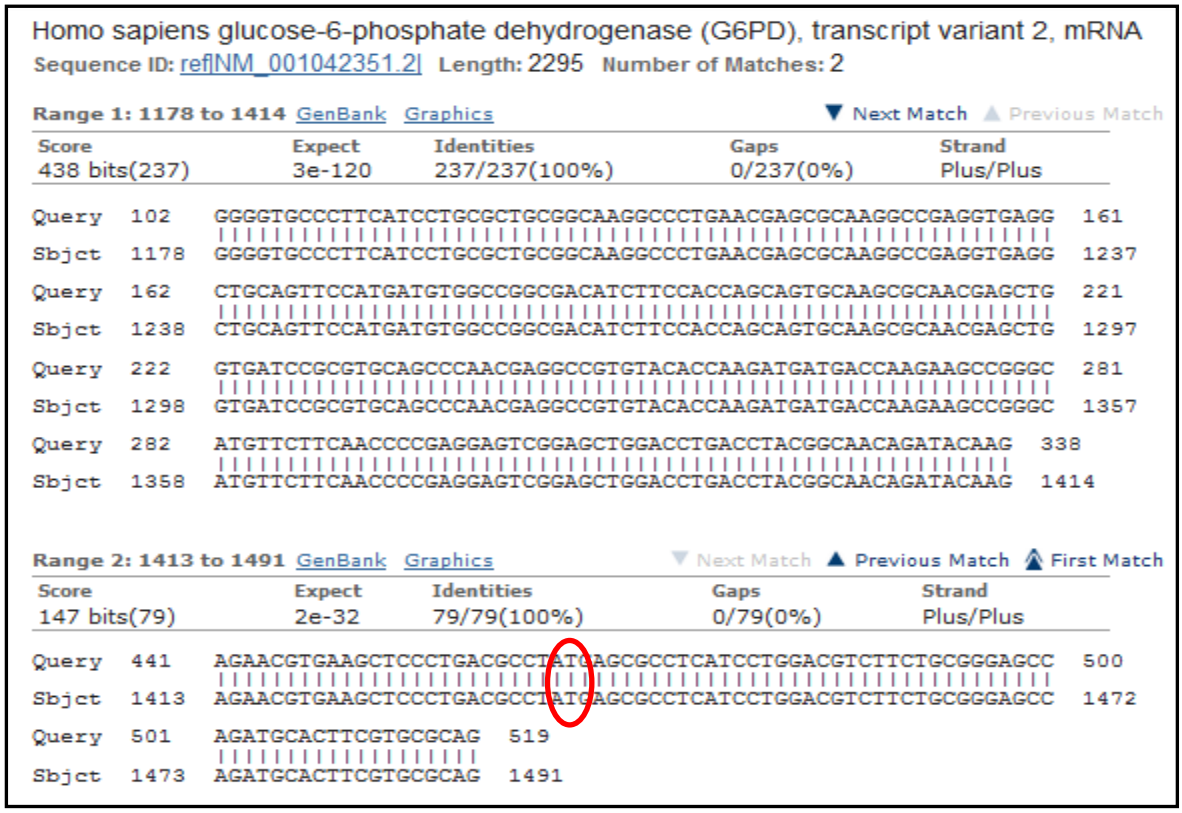

Fig 3.15: Illustration of BLAST results of a hemizygous male participant fall under red normalized melt curve. The hemizygous participant harbors a T base at c.1311 (1438th base of mRNA) position of query sequence which is same as the base in reference sequence.

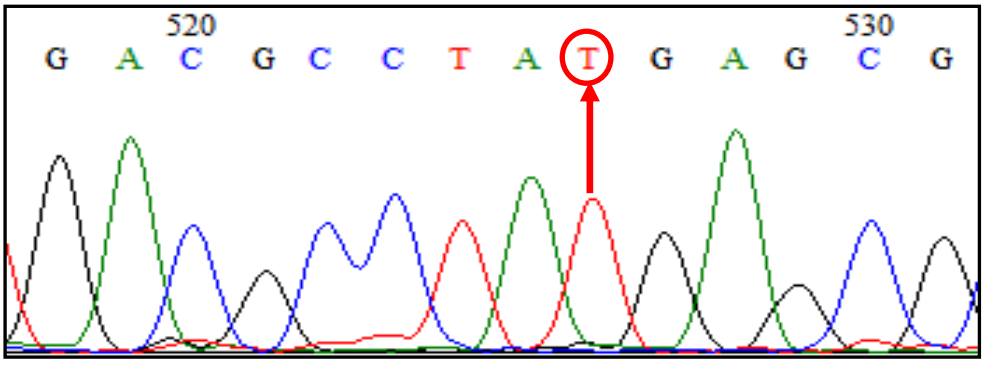

Fig 3.16: Chromatogram of portion of exon-11 for a hemizygous male that fall under green normalized melt curves. The 526th base in the chromatogram highlighted corresponds to 1438th base (c.1311) of mRNA which is same like the base of reference mRNA (NM_001042351.2).

After confirming the fact that HRM efficiently differentiates between c.1311C and c.1311T alleles using hemizygous males, extracted DNA samples from female participants were subjected to HRM analysis to see whether it can differentiate between heterozygous and homozygous states. Figure-3.17 illustrates three types of normalized melt curves. Sequencing of samples from each clusters revealed the three states: blue, red and green normalized melt curves correspond to homozygous c.1311C, homozygous c.1311T and heterozygous c.1311 (T+C) respectively shown in the figure-3.17. Difference curves in figure-3.18 makes the three states even more distinguishable.

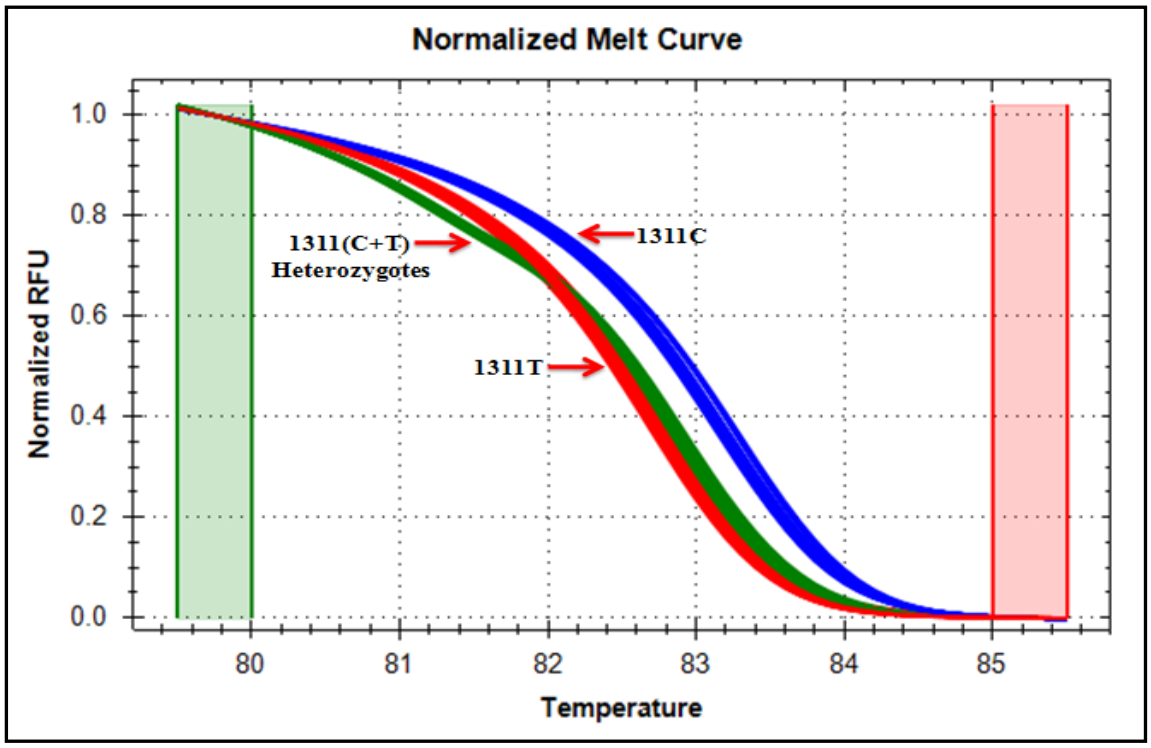

Fig 3.17: Normalized melt curves for female participants for exon-11. HRM analysis can efficiently differentiate between homozygous c.1311C, homozygous c.1311T and heterozygous c.1311 (T+C).

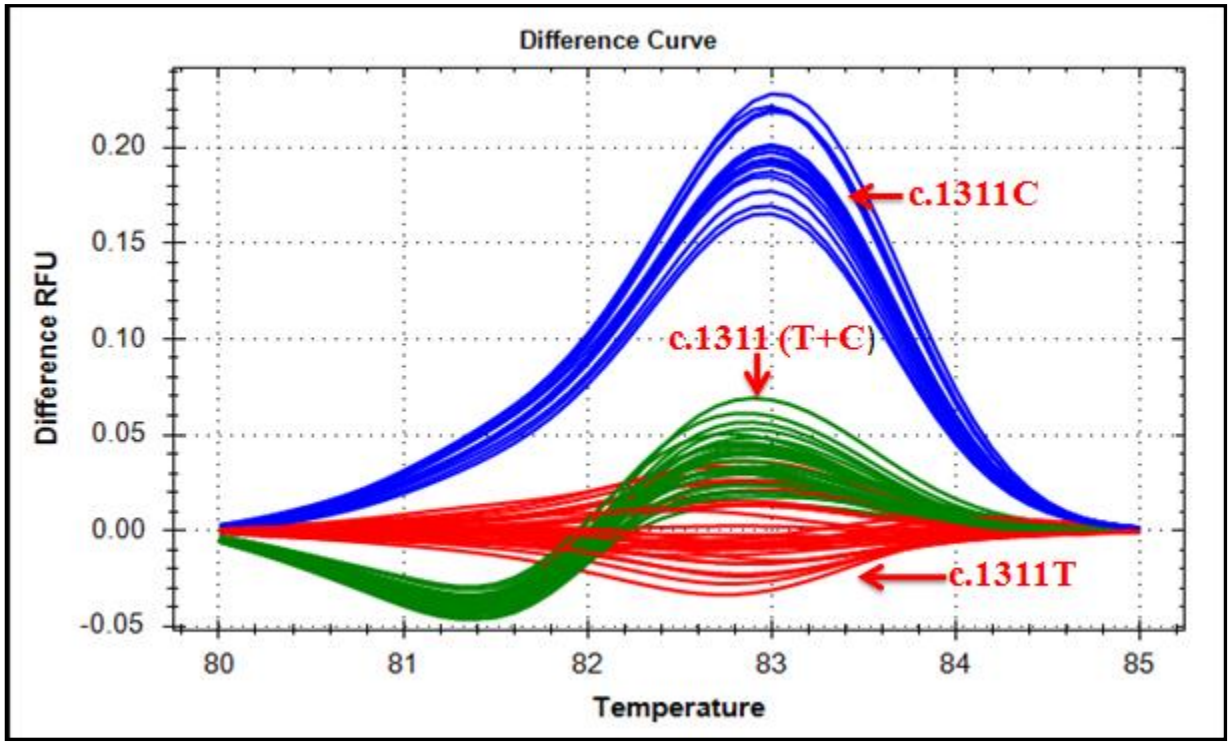

Fig 3.18: Difference curves for exon-11 in female participants. Discernable change in relative fluorescence unit is observable among homozygous c.1311C, homozygous c.1311T and heterozygous c.1311 (T+C) alleles.

Blast result of sequencing data for heterozygous samples is shown in the figure-3.19. As sequencing data and BLAST results for c.1311C and c.1311T alleles are shown in figure-3.13 and figure-3.15 respectively, those are not repeated again for homozygous c.1311C, homozygous c.1311T states. Blast result of participant which falls under green normalized melt curve shows a mismatch at 1438th base (c.1311) of mRNA. The mismatched base of query sequence is Y which actually represents combination of C plus T bases. In chromatogram the mismatched base corresponds to 526th base (figure-3.20).

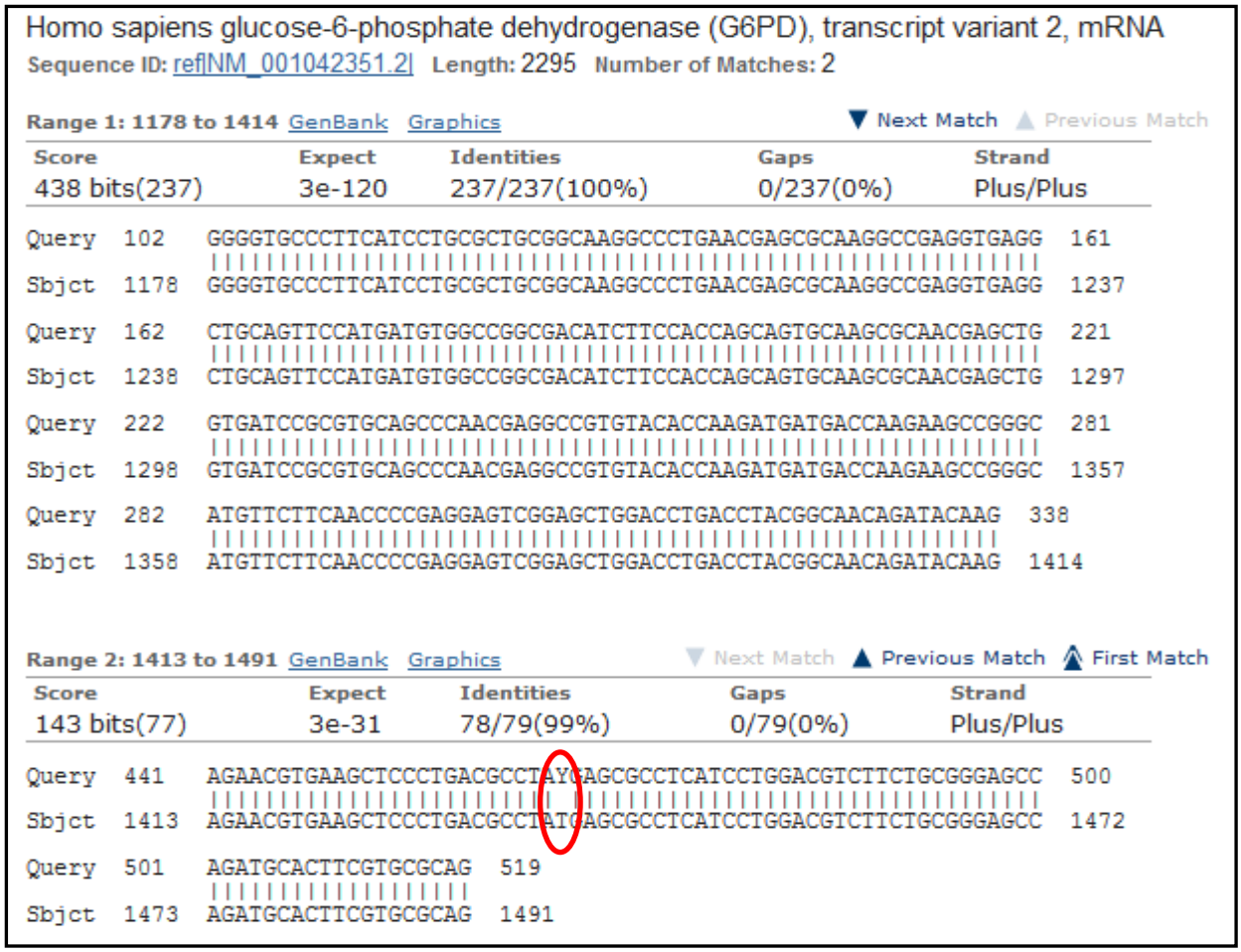

Fig 3.19: BLAST result for a sample with green normalized melt curve for exon-11. The mismatched base at 1438th position of mRNA is highlighted by a red circle.

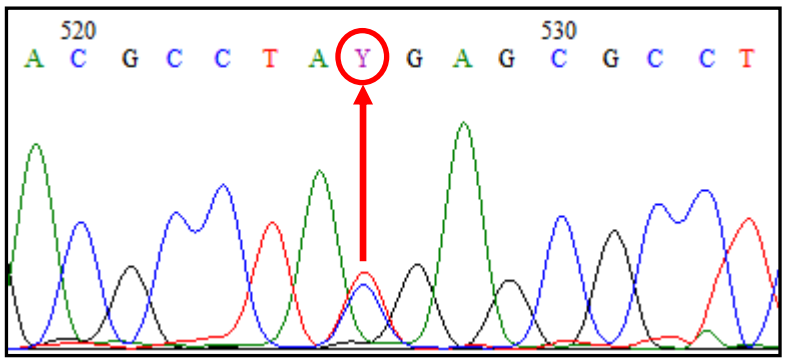

Fig 3.20: Chromatogram segment of exon-11 of a participant with green normalized melt curve. The mismatched base of BLAST result (figure-3.19) corresponds to 526th base in the chromatogram.

Polymorphism (c.1311T/C) in the 1438th base of mRNA or 1311th base of coding sequence does not change the 437th amino acid in the protein. Both codons, TAC and TAT, code for serine at 437th position of G6PD peptide. It is not clear whether T to C transition or C to T transition resulted in this polymorphism, as in some study c.1311T was more frequent but opposite result observed in other studies.

3.1.5 Detection of G6PD Polymorphism in intron-11 or IVS-11 t93c

To further confirm the reliability of HRM in heterozygous mutation detection, IVS-11 t93c polymorphism was targeted. Forward primer 5'-GCCCTCCCTCCCTGTGTG-3' and reverse primer 5'CAGCTCAATCTGGTGCAGCAGT-3' was used to serve the purpose. Analysis of samples using HRM produced following results shown in figure-3.21 and figure-3.22. Three types of normalized melt curves were generated from samples that were subjected to real time PCR and subsequent high resolution melting (HRM) analysis. Samples from all three normalized melt curves were sequenced to find out corresponding allele differences.

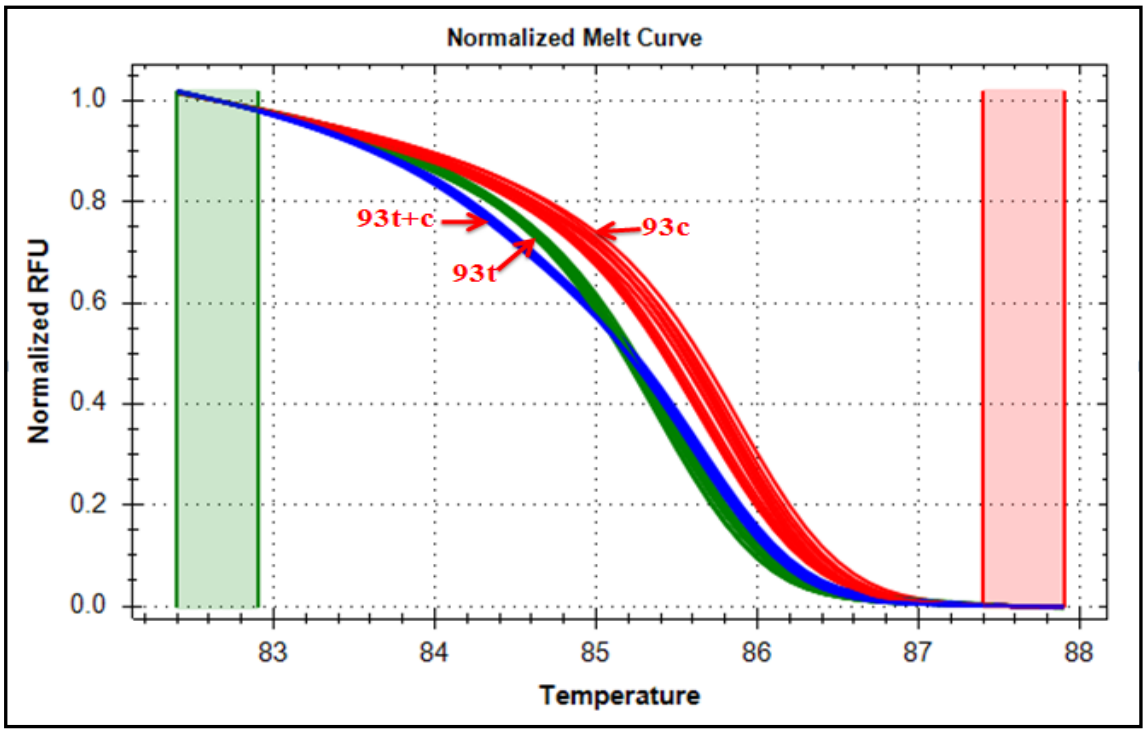

Fig 3.21: Normalized melt curve for intron-11 for study participants. Melt curve analysis produced three types of normalized melting curves; sequencing of samples from all three clusters reveals red, green and blue normalized curves corresponding to IVS-11 93c, 93t and 93t+c respectively (figure-3.23, figure-3.25 and figure-3.27).

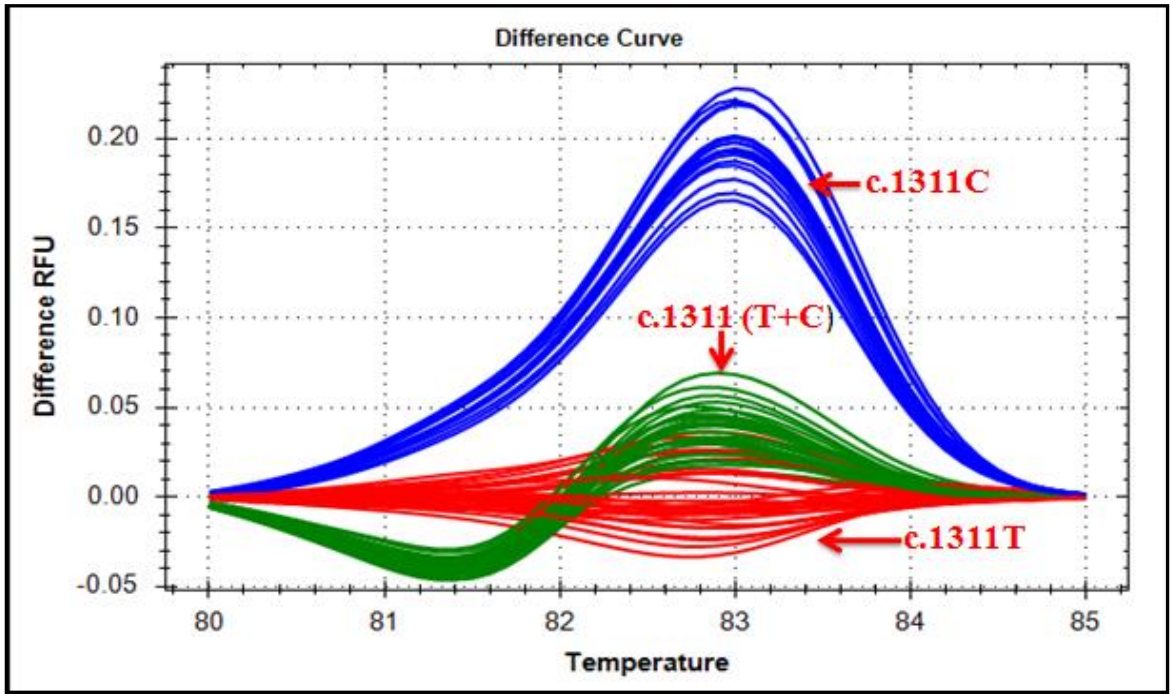

Fig 3.22: Difference curves for intron-11 for study participants. Three distinct alleles (IVS-11 93t, 93c and 93t+c) become easily distinguishable by HRM generated difference curves.

Homozygous allelic variants, IVS-11 93t and 93c, could be characterized by a temperature shift on the resulting melt curve produced by HRM analysis. In comparison, IVS-11 93t+c heterozygotes are characterized by changes in melt curve shape. This is due to base-pair mismatching generated as a result of destabilized heteroduplex annealing between wild-type and variant strands. These differences can be easily seen on the resulting melt curve and the melt profile differences between the different genotypes can be amplified visually via generating a difference curve (figure-3.22).

BLAST results of samples sequenced from each cluster are shown in figure-3.23, figure-3.25 and figure-3.27. Figure-3.23 demonstrates presence of IVS-11 93c in samples under red normalized curves (figure-3.21 and figure-3.22) which is 153672423th base of reference sequence of X-chromosome (NC_018934.2). The IVS-11 93c base of a sample under red curve is in 214th position in chromatogram (figure-3.24). Sequencing of samples under green normalized melt curve revealed IVS-11 93t allele (figure-3.25) which corresponds to 200th base of chromatogram for representative sample (figure-3.26). Samples under blue normalized melt curve contain both IVS-11 93t and IVS-11 93c (read as Y by data collection

software integrated with the sequencer according to settings) alleles (figure-3.27) and the base coincides in 385th position of chromatogram in representative sample (figure-3.28).

Homo sapiens chromosome X, alternate assembly CHM1_1.1

Sequence ID: ref|NC_018934.2| Length: 155181468 Number of Matches: 1

Range 1: 153672217 to 153672481 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
479 bits(259)	8e-133	263/265(99%)	0/265(0%)	Plus/Minus

Features: glucose-6-phosphate 1-dehydrogenase isoform a
glucose-6-phosphate 1-dehydrogenase isoform b

```
Query  1          GCAGTGGCATCAGCAAGACACTCTCTCCCTCACAGAACGTGAAGCTCCCTGACGCCTACG  60
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Sbjct  153672481  GCAGTGGCATCAGCAAGACACTCTCTCCCTCACAGAACGTGAAGCTCCCTGACGCCTATG  153672422

Query  61         AGCGCCTCATCCTGGACGTCTTCTGCGGGAGCCAGATGCACTTCGTGCGCAGGTGAGGCC  120
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  153672421  AGCGCCTCATCCTGGACGTCTTCTGCGGGAGCCAGATGCACTTCGTGCGCAGGTGAGGCC  153672362

Query  121        CAGCTGCCGGCCCCTGCATACCTGTGGGCTATGGGGTGGCCTTTGCCCTCCCTCCCTGTG  180
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  153672361  CAGCTGCCGGCCCCTGCATACCTGTGGGCTATGGGGTGGCCTTTGCCCTCCCTCCCTGTG  153672302

Query  181        TGCCACCGGCCTCCCAAGCCATACTATGTCCCCTCAGCGACGAGCTCCGTGAGGCCTGGC  240
                  |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  153672301  TGCCACCGGCCTCCCAAGCCATACCATGTCCCCTCAGCGACGAGCTCCGTGAGGCCTGGC  153672242

Query  241        GTATTTTCACCCCACTGCTGCACCA  265
                  |||||||||||||||||||||||||
Sbjct  153672241  GTATTTTCACCCCACTGCTGCACCA  153672217
```

Fig 3.23: BLAST result for a sample with red normalized melt curve for intron-11. The IVS-11 93c base of sample is highlighted using a red circle along with the IVS-11 93t base of reference sequence of X-chromosome (NC_018934.2).

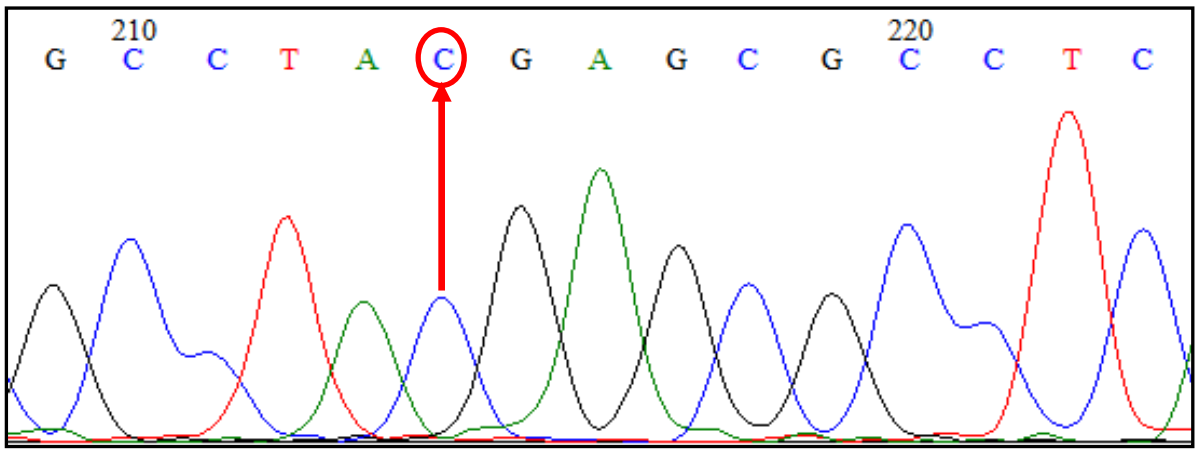

Fig 3.24: Chromatogram segment of intron-11 of a participant with green normalized melt curve. The IVS-11 93c of sample under red normalized curve corresponds to 214th base of chromatogram.

Homo sapiens chromosome X, alternate assembly CHM1_1.1

Sequence ID: ref|NC_018934.2| Length: 155181468 Number of Matches: 1

Range 1: 153672023 to 153672482 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
850 bits(460)	0.0	460/460(100%)	0/460(0%)	Plus/Minus

Features: glucose-6-phosphate 1-dehydrogenase isoform a
glucose-6-phosphate 1-dehydrogenase isoform b

```
Query  1          GGCAGTGGCATCAGCAAGACACTCTCTCCCTCACAGAACGTGAAGCTCCCTGACGCCTAT  60
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  153672482  GGCAGTGGCATCAGCAAGACACTCTCTCCCTCACAGAACGTGAAGCTCCCTGACGCCTAT  153672423

Query  61         GAGCGCCTCATCCTGGACGTCTTCTGCGGGAGCCAGATGCACTTCGTGCGCAGGTGAGGC  120
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  153672422  GAGCGCCTCATCCTGGACGTCTTCTGCGGGAGCCAGATGCACTTCGTGCGCAGGTGAGGC  153672363

Query  121        CCAGCTGCCGGCCCCTGCATACCTGTGGGCTATGGGGTGGCCTTTGCCCTCCCTCCCTGT  180
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  153672362  CCAGCTGCCGGCCCCTGCATACCTGTGGGCTATGGGGTGGCCTTTGCCCTCCCTCCCTGT  153672303

Query  181        GTGCCACCGGCCTCCCAAGCCATACCATGTCCCCTCAGCGACGAGCTCCGTGAGGCCTGG  240
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  153672302  GTGCCACCGGCCTCCCAAGCCATACCATGTCCCCTCAGCGACGAGCTCCGTGAGGCCTGG  153672243

Query  241        CGTATTTTCACCCCACTGCTGCACCAGATTGAGCTGGAGAAGCCCAAGCCCATCCCCTAT  300
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  153672242  CGTATTTTCACCCCACTGCTGCACCAGATTGAGCTGGAGAAGCCCAAGCCCATCCCCTAT  153672183

Query  301        ATTTATGGCAGGTGAGGAAAGGGTGGGGGCTGGGGACAGAGCCCAGCGGGCAGGGGCGGG  360
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  153672182  ATTTATGGCAGGTGAGGAAAGGGTGGGGGCTGGGGACAGAGCCCAGCGGGCAGGGGCGGG  153672123

Query  361        GTGAGGGTGGAGCTACCTCATGCCTCTCCTCCACCCGTCACTCTCCAGCCGAGGCCCCAC  420
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  153672122  GTGAGGGTGGAGCTACCTCATGCCTCTCCTCCACCCGTCACTCTCCAGCCGAGGCCCCAC  153672063

Query  421        GGAGGCAGACGAGCTGATGAAGAGAGTGGGTTTCCAGTAT  460
                  |||||||||||||||||||||||||||||||||||||||
Sbjct  153672062  GGAGGCAGACGAGCTGATGAAGAGAGTGGGTTTCCAGTAT  153672023
```

Fig 3.25: BLAST result for a sample with green normalized melt curve for intron-11. The IVS-11 93t base of sample is shown using a red circle along with the IVS-11 93t base of reference sequence of X-chromosome (NC_018934.2).

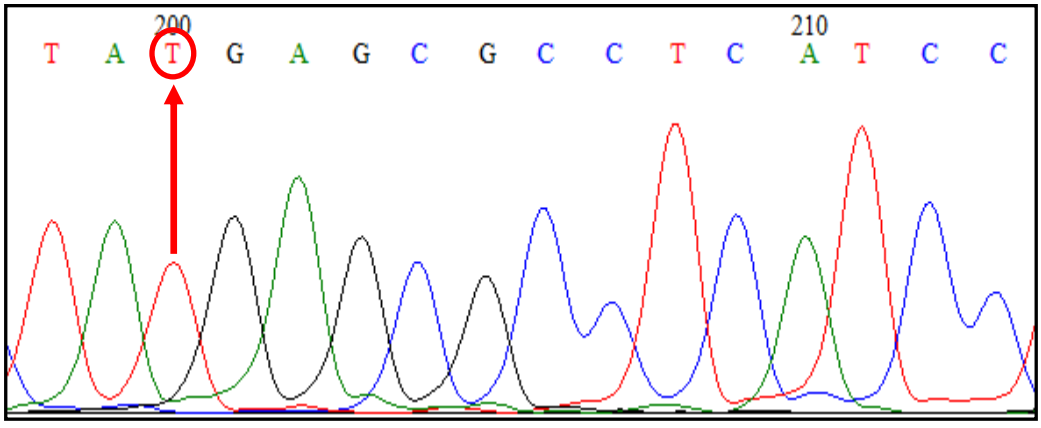

Fig 3.26: Chromatogram segment of intron-11 of a participant with green normalized melt curve. The IVS-11 93t of sample under green normalized melting curve corresponds to 200th base of chromatogram.

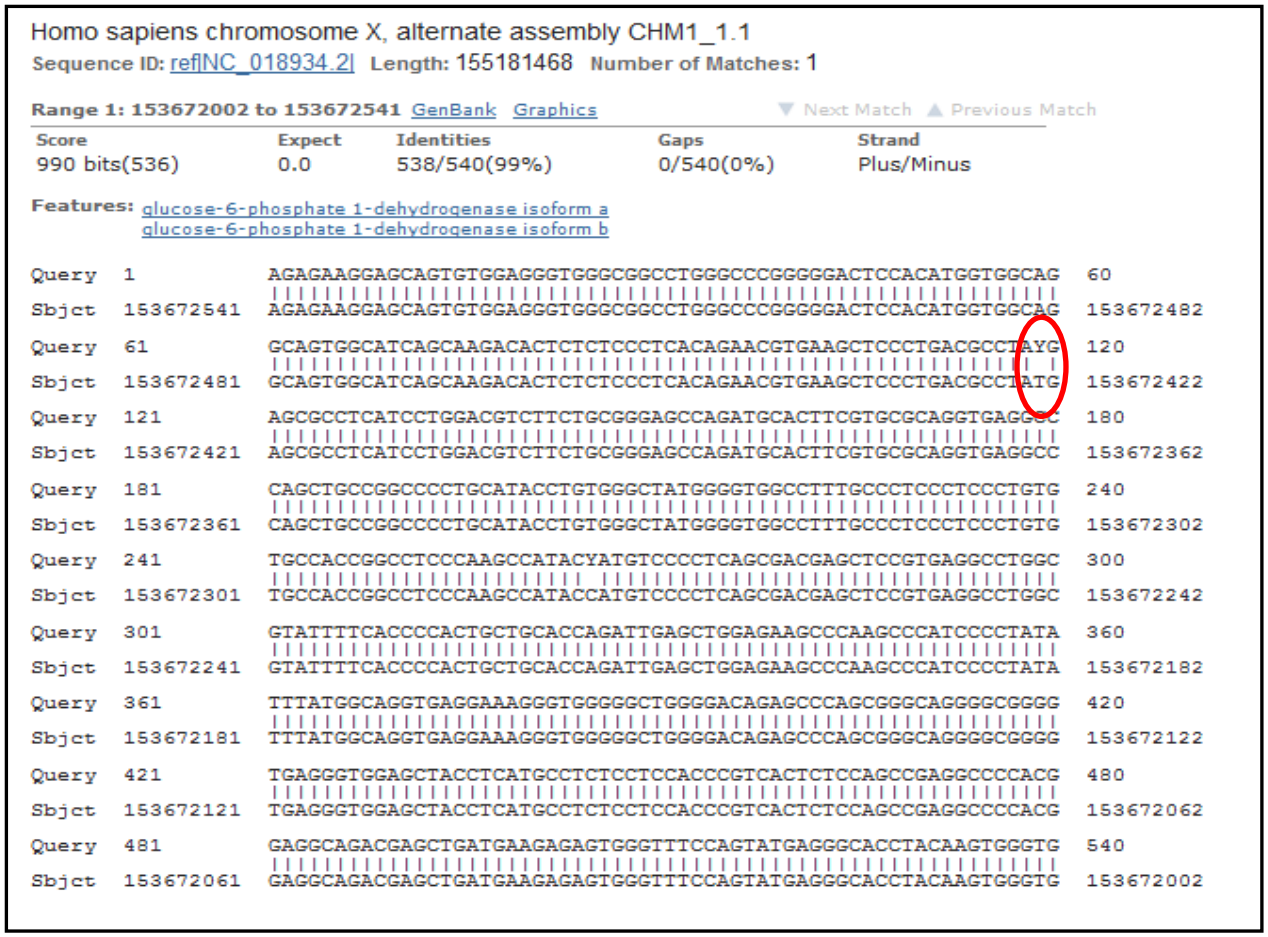

Fig 3.27: BLAST result for a sample with blue normalized melt curve for intron-11. The presence of IVS-11 93t+c alleles in sample is highlighted along with the IVS-11 93t base of reference sequence of X-chromosome (NC_018934.2).

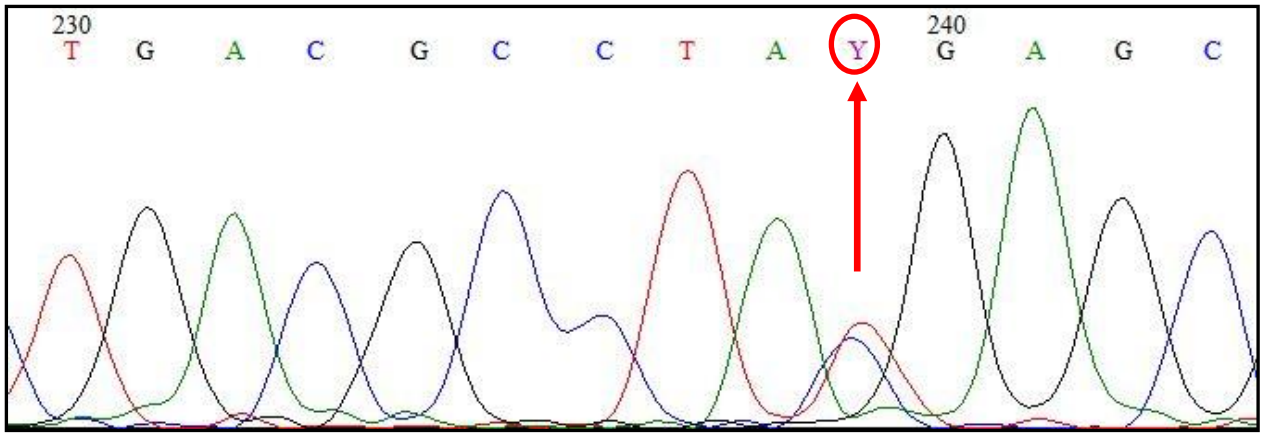

Fig 3.28: Chromatogram segment of intron-11 of a participant with blue normalized melt curve. The presence of heterozygous alleles IVS-11 93t+c in sample under blue normalized melting curve corresponds to 239th base of chromatogram.

The influence of polymorphism IVS11_93 in G6PD enzyme activity is controversial. There have been reports that combination of c. C1311T/IVS-XI t93c is associated with mild form of G6PD deficiency. Other studies suggested that C1311T/IVS-XI t93c is common in normal population who had normal G6PD enzyme activity. Our study is also in coherence with the reports that IVS-XI t93c is common in people with normal G6PD study as it was found in participants with normal G6PD enzyme activity.

Chapter 4: Discussion

Glucose 6 phosphate dehydrogenase (G6PD) deficiency, an X-linked inherited disease, is one of the most common enzymopathies, affecting over 400 million people worldwide. To date, more than 150 (Yan, Xu et al. 2010) different G6PD mutations have been identified among different ethnic populations and each ethnic population has a characteristic mutation profile. G6PD gene mutations cause deficiency of the enzyme and a large spectrum of diseases.

There are widely used methods for the screening and diagnosis of G6PD deficiency. Such methods include qualitative fluorescent spot test and the spectrophotometric quantitative enzyme activity assay. In these methods, end products such as NADPH are used to measure enzyme activity. End product targeting is less cumbersome, more cost effective and rapidly performable, but end product targeting may not be useful sometimes and misdiagnoses may occur. This is because these methods cannot detect deficiencies in some cases, such as heterozygous female specimens. This may expose any misdiagnosed persons to accidental clinical conditions. Qualitative fluorescent spot test and spectrophotometric quantitative enzyme activity assays use NADPH for G6PD enzyme activity assessment. These methods are more reliable for detection of hemizogous and homozygous mutations, but they are not suitable for heterozygous mutations. Sometimes heterozygous mutations show mild deficiency or normal enzyme activity which is misleading. The fluorescent spot test is inexpensive and easy to perform but it is only reliable for distinguishing hemizygous G6PD deficient men from non-deficient men. Current qualitative tests can only accurately diagnose G6PD deficiency in people with G6PD enzyme activity below 30–40 % normal activity. It is unable to detect heterozygote deficient females and may misclassify them with having intermediate G6PD activity. This is because; G6PD normal and G6PD deficient red blood cell (RBC) populations can co-exist within a heterozygous female. As a consequence of lyonization (inactivation of one X chromosome) heterozygous women have two red blood cell populations, each resulting from the expression of one of two G6PD alleles: one population may have normal G6PD level, whereas the other population expresses the G6PD deficient allele. In these women, G6PD normal RBCs may mask G6PD deficient cell populations and can result in G6PD normal test results.(Ley, Luter et al. 2015).

Similarly, like fluorescent spot test, spectrophotometric G6PD enzyme assay fails to produce the exact status of G6PD heterozygotes due to the co-existence of G6PD normal and G6PD deficient cell population. Co-existence of two types of cell populations is not the only barrier for G6PD enzyme assay using spectrophotometric methods. Other contributing factors that leads to false G6PD activity include acute hemolysis, blood transfusion etc. Enzyme activity test should be performed when patients are in remission as results may be falsely negative during acute hemolysis since the erythrocytes undergo destruction. On the other hand, enzyme activity can be high in patients just after recovering from acute hemolysis as the number of reticulocytes increases, contributing to increased G6PD enzyme activity levels. After blood transfusion, donor RBCs circulate in the recipient and G6PD testing gives a false positive result. G6PD deficiency in newborns is also misleading as reticulocyte count is high and their turnover is also rapid, thus showing falsely elevated G6PD enzyme activity levels. Thus it necessitates genetic testing for exact diagnosis of females, neonates and other cases as mentioned above when FST and enzyme assay may not be reliable. Otherwise, without knowing their G6PD status, they are at risk of hemolysis if accidentally or unknowingly exposed to certain foods such as fava beans, legumes, soya foods as well as drugs like primaquine, antipyrine and sulfa drugs that lead to oxidative damage. Thereafter it is urgent to detect the mutation in such a way which can produce the authentic result, and can overcome the limitations associated with other existing methods.

Various molecular methods have been developed to detect G6PDmutations; these include PCR-single strand conformational polymorphism analysis, DNA sequencing, amplification refractory mutation system, gradient gel electrophoresis (DGGE), probe melting curve, microarray, matrix–assisted laser desorption/ ionization –time of flight mass spectrometry, reverse dot blot assay, restriction fragment length polymorphism (RFLP) and denaturing high-performance liquid chromatography. However, these approaches are either expensive or technically challenging.

Hence we need to apply a technique to detect G6PD mutations which will be more cost-effective, reliable, rapid and easy. High-resolution DNA melting (HRM) assay has been recently introduced as a rapid, inexpensive and effective method for mutation detection in which PCR and mutation scanning are performed simultaneously in a single procedure within two and half hours. Sensitivity and specificity for mutation detection are extremely high, and this technique also has the advantages of being both low cost and high throughput. Recently, a growing number of potential causative mutations for different diseases have been measured

by HRM analysis, which can identify hundreds of mutations in many different genes. The detection of human genetic diseases, specifically those that are autosomal dominant, recessive, or X chromosome-linked, has been one of the largest applications of this technique. To avoid such problems associated with fluorescent spot test and enzyme activity assay, and also to make mutation detection easier, we wanted to employ real time PCR-HRM analysis. Differentiation between mutant and wild type alleles was made based on melting curve shapes. In our current study we have identified three different G6PD mutations; G6PD Mahidol (c.G487A), Kalyan-Kerala (c.G949A) and Orissa (C131G) in the Bangladeshi population through gene sequencing. For PCR-HRM analysis, primers spanning c.G487A and c.G949A mutations were used for analyzing five deficient samples with c.G487A mutation and one deficient sample with c.G949A mutation. G6PD non-deficient samples were used as controls. Melt curves for all five c.G487A deficient specimens differed significantly from that of samples with normal G6PD activity. Similarly, PCR-HRM analysis could identify c.G949A mutation. HRM findings for each mutation were consistent with the sequencing result, supporting the applicability of the method for easier and more reliable detection of G6PD mutations. Although PCR-HRM analysis of c.C131G mutation was not done in the present study, it can be expected that it will also enable distinction of deficient samples from normal samples.

Even though we did not find any heterozygous mutation in deficient samples, we had demonstrated the usefulness of PCR-HRM for heterozygous mutation detection utilizing c.1311 T/C and IVS-11_93t/c polymorphisms. Analysis of c.1311 single nucleotide polymorphism (SNP) in 87 samples determined that c.1311T/C heterozygotes could produce a different pattern of melting curve for c.1311C or c.1311T alleles. The reason behind this phenomenon is that there are steps in melting curve analysis that involves initial denaturation and subsequent annealing of the PCR product. As the annealing process is random and irrespective of single base pair differences, hetero-duplex formation occurs between two types of alleles during this annealing step. This hetero-duplex formation affects the melting property of PCR product of the heterozygote, and produces a melt curve that is distinct for homozygous c.1311C and c.1311T alleles. At the c.1311 position of exon-11, 32 samples had a homozygous Thymine base (36.78% frequency), 35 had a homozygous Cytosine base (40.23%), and 20 specimens showed heterozygosity in terms of T/C (22.98%). Similar to detection of c.1311T/C alleles, 21 samples were subjected to PCR-HRM analysis and heterozygous and homozygous states for IVS-11_93t/c SNP were successfully differentiated.

Frequencies for IVS-11 93(t+c), 93t and 93c alleles were present at 18%, 32% and 50% respectively.

Although HRM has many beneficial applications, it has some limitations as well. To find an unknown mutation it is necessary to scan all the exons one by one, as the amplicon size should be no more than ~200 bp to get better resolution and reproducible results. Even if all exons are scanned, a mutation may not be found if present in the intron which is not considered for melt curve analysis. Another limitation of HRM analysis is the possibility that another unexpected SNP may interfere in the mutations of interest. So, the designed amplicons should be in the range of 75bp to 200bp for proper analysis of mutation/polymorphism by PCR-HRM. In addition, the designed primers should flank the exon or intron as closely as possible so that no portion of the gene is missed if searching for an unknown mutation. Additionally, effective amplicon designing is an important consideration for achieving robust and reproducible results in HRM analysis. Moreover, different heterozygotes may produce melting curves that are very similar to each other, although the HRM analyses of heterozygotes are excellent. Yet there is a misclassification risk which is at about 10%. (Tindall, Petersen et al. 2009). So, it is necessary to perform direct DNA sequencing to confirm the mutation site if different heterozygotes produce melting curve very similar to each other. The biggest problem of HRM is that it cannot detect the site of unknown mutation.

Even though it has some limitations, real time PCR-HRM has shown its dynamics in successful detection of G6PD mutations, polymorphisms and zygosity-especially the heterozygosity. The greatest benefit of melt curve analysis is that it can be a very useful tool for detecting unknown mutations, as it does not directly involve the mutant base as other methods like restriction digestion, amplification refractory mutation system, etc. HRM will give indication of unknown base substitutions, deletions or insertions and sequencing of the fragment showing distinct melting properties revealing the exact position of aberration. Thus it reduces the burden of sequencing of the whole gene and saves cost. Another usefulness of HRM is that it can be used for the screening of known G6PG mutations. Even though there are 150 known G6PD mutations, only a small fraction are commonly observed within any particular cast, ethnic group or geographical location. So, designing HRM primers for specific mutations for particular groups can help screening of G6PD deficiency at the molecular level, especially for heterozygous mutation detection.

In epidemiological studies it is seen that the prevalence of Glucose-6-phosphate dehydrogenase deficiency is significantly frequent in malaria endemic areas. It is found that malaria and G6PD share the same geographical area. So it can be easily said that there is a strong correlation between malaria and G6PD deficiency. This correlation reveals two important results- (1) G6PD deficiency provides protection against malaria and (2) G6PD deficiency causes harm in fighting against malaria. It is very important to detect G6PD status in malarial endemic areas because the drugs used to treat malaria may cause oxidative stress or even acute hemolytic anemia in G6PD deficient patients. In this case real-time PCR-HRM can be a great support in helping to detect G6PD conditions in emergencies. So PCR-HRM approach can be really helpful in malaria endemic areas to screen for G6PD mutations where people are more frequently taking antimalarial drugs. This will help to ensure that G6PD deficient people should take alternative antimalarial drugs that do not cause oxidative damage.

Researchers have worked on cancer where PCR-HRM technique has shown to be suitable for the detection of mutations in breast cancer (Krypuy, Ahmed et al. 2007). Accordingly, the detection of human genetic disorders, especially autosomal recessive, dominant, and X-linked disorders, has been one of the largest applications of HRM analysis. Fabry disease (FD) is an X-linked disorder, associated with GLA gene mutations. To date more than 500 mutations in the GLA gene have been associated with FD (Hwu, Chien et al. 2009). HRM analysis of the GLA gene has been demonstrated as a reliable pre- sequencing screening tool. Spinal muscular atrophy (SMA) is the most common neuromuscular autosomal recessive disorder. Recently, the HRM platform has been applied in the detection of survival motor neuron 1 (SMN1) gene homozygous deletion for the diagnosis of SMA .To detect the mutation of SMA the variants with homozygous deletion of SMN1 exon 7 produced a distinctive melt profile that identified SMA patients. In near future HRM analysis will be a feasible method for use as a rapid and large-scale newborn screening technique for SMA. Hemoglobinopathies like α and β-thalassemia are the most common inherited disorders in humans due to globin gene mutation. Around 7% of the world populations is a carrier of globin gene mutation (Steinberg, Forget et al. 2009). HRM has also been used for the detection of HBB gene in β-thalassemia and as a reliable method for β-globin gene cluster haplotyping. It is also used in finding α-hemoglobin variants and HRM can easily identify the Hb variants using melting curves and the variants can be easily distinguished to be located at either HBA1 or HBA2 gene. Another genetic disease is cystic fibrosis, in which HRM has

been applied successfully. Other recent achievements of HRM are in the detection of mutations of Chronic Granulomatous Disease, Prader-Willi syndrome, Angelman syndrome, Cystic Fibrosis, Duchenne/Becker muscular dystrophy (BMD) and many others(Er and Chang 2012).

Favism or glucose-6-phosphate dehydrogenase deficiency is one of the common enzymopathises in Bangladesh. It is a disease which is involved in the inactivation of the gene responsible for coding the enzyme glucose-6-phosphate dehydrogenase which helps red blood cell work properly. Too little G6PD may lead to the destruction of red blood cells (hemolysis). Acute hemolysis can even cause death and that is why detection of G6PD status is important.

Our study was successful for detection of common G6PD mutations by this molecular method, HRM. Moreover, we can implement this molecular method for heterozygous G6PD mutations detection in Bangladesh as no other molecular method is in existence for G6PD mutations detection. As in case of neonatal screening the existing qualitative methods may have high rate of misinterpretation of G6PD status, PCR-HRM analysis will be very helpful method to get the authentic results during neonatal testing for G6PD status in Bangladesh. Also real-time PCR-HRM can be put into use to perform mutation detection on a large scale in malaria-endemic regions like Bandarban. In our study we have employed HRM analysis only for G6PD mutation screening but this technique can be extended to detect other mutations associated with common genetic disorders like Thalassemia, Hopothyroidism, Sickle cell anemia etc. in Bangladeshi population. We expect that the PCR-HRM analysis will be a very useful tool for use in medical or research sectors in Bangladesh.

This is the first study on G6PD deficiency conducted in Bangladesh which provides a genetic and molecular basis of the disorder. We hope the results and outcomes of this research will be helpful to medical institutions, researchers, and students who wish to study G6PD or any hemolysis related diseases.

References:

A Geostatistical Model-Based Map. PLoS Med 9(11): e1001339. doi:10.1371/journal.pmed.1001339

Acero, L. H. (2013). "Growth Response of Brassica Rapa on the Different Wavelength of Light." International Journal of Chemical Engineering and Applications 4(6): 415.

Allahverdiyev, A. M., E. S. Abamor, et al. (2012). Glucose-6-phosphate dehydrogenase deficiency and malaria: a method to detect primaquine-induced hemolysis in vitro, INTECH Open Access Publisher.

Ayene, I. S., T. D. Stamato, et al. (2002). "Mutation in the glucose-6-phosphate dehydrogenase gene leads to inactivation of Ku DNA end binding during oxidative stress." Journal of Biological Chemistry **277**(12): 9929-9935.

Acero, L. H. (2013). "Growth Response of Brassica Rapa on the Different Wavelength of Light."International Journal of Chemical Engineering and Applications **4**(6): 415.

Beutler, E., S. Duparc, et al. (2007). "Glucose-6-phosphate dehydrogenase deficiency and antimalarial drug development." The American journal of tropical medicine and hygiene **77**(4): 779-789.

Bonilla, J. F., M. C. Sánchez, et al. (2007). "Glucose-6-phosphate dehydrogenase (G6PD): Response of the human erythrocyte and another cells to the decrease in their activity." Colombia Médica 38(1): 68-75.

Chen, E. Y., A. Cheng, et al. (1991). "Sequence of human glucose-6-phosphate dehydrogenase cloned in plasmids and a yeast artificial chromosome." Genomics 10(3): 792-800.

Er, T.-K. and J.-G. Chang (2012). "High-resolution melting: applications in genetic disorders." Clinica Chimica Acta 414: 197-201.

Farhud, D. and L. Yazdanpanah (2008). "Glucose-6-phosphate dehydrogenase (G6PD) Deficiency." Iranian journal of public health 37(4): 1-18.

Finley, D. S. (2008). "Glucose-6-phosphate dehydrogenase deficiency associated stuttering priapism: Report of a case." The journal of sexual medicine 5(12): 2963-2966.

Fusco, F., M. Paciolla, et al. (2012). "Genomic architecture at the Incontinentia Pigmenti locus favours de novo pathological alleles through different mechanisms." Human molecular genetics 21(6): 1260-1271.

Goss, S. and H. Harris (1977). "Gene transfer by means of cell fusion I. Statistical mapping of the human X-chromosome by analysis of radiation-induced gene segregation." Journal of cell science 25(1): 17-37.

Hale, J. P., C. P. Winlove, et al. (2011). "Effect of hydroperoxides on red blood cell membrane mechanical properties." Biophysical journal 101(8): 1921-1929.

Hwu, W. L., Y. H. Chien, et al. (2009). "Newborn screening for Fabry disease in Taiwan reveals a high incidence of the later-onset GLA mutation c. 936+ 919G> A (IVS4+ 919G> A)." Human mutation 30(10): 1397-1405.

Kotaka, M., S. Gover, et al. (2005). "Structural studies of glucose-6-phosphate and NADP+ binding to human glucose-6-phosphate dehydrogenase." Acta Crystallographica Section D: Biological Crystallography 61(5): 495-504.

Krenková, P., P. Norambuena, et al. (2009). "Evaluation of high-resolution melting (HRM) for mutation scanning of selected exons of the CFTR gene." Folia biologica 55(6): 238.

Krypuy, M., A. A. Ahmed, et al. (2007). "High resolution melting for mutation scanning of TP53 exons 5–8." BMC cancer **7**(1): 1.

LaRue, N., M. Kahn, et al. (2014). "Comparison of quantitative and qualitative tests for glucose-6-phosphate dehydrogenase deficiency." The American journal of tropical medicine and hygiene 91(4): 854-861.

Lee, S. W. H., N. M. Lai, et al. (2016). "Adverse effects of herbal or dietary supplements in G6PD deficiency: a systematic review." British journal of clinical pharmacology.

Ley, B., N. Luter, et al. (2015). "The challenges of introducing routine G6PD testing into radical cure: a workshop report." Malaria journal 14(1): 1.

Luzzatto, L. (2006). "Glucose 6-phosphate dehydrogenase deficiency: from genotype to phenotype." Haematologica 91(10): 1303-1306.

Mamlok, R. J., V. Mamiok, et al. (1987). "Glucose-6-phosphate dehydrogenase deficiency, neutrophil dysfunction and Chromobacterium violaceum sepsis." The Journal of pediatrics 111(6): 852-854.

Martini, G., D. Toniolo, et al. (1986). "Structural analysis of the X-linked gene encoding human glucose 6-phosphate dehydrogenase." The EMBO journal 5(8): 1849.

Minucci, A., B. Giardina, et al. (2009). "Glucose-6-phosphate dehydrogenase laboratory assay: How, when, and why?" IUBMB life 61(1): 27-34.

Nkhoma, E. T., C. Poole, et al. (2009). "The global prevalence of glucose-6-phosphate dehydrogenase deficiency: a systematic review and meta-analysis." Blood Cells, Molecules, and Diseases 42(3): 267-278.

Pai, G., J. A. Sprenkle, et al. (1980). "Localization of loci for hypoxanthine phosphoribosyltransferase and glucose-6-phosphate dehydrogenase and biochemical evidence of nonrandom X chromosome expression from studies of a human X-autosome translocation." Proceedings of the National Academy of Sciences **77**(5): 2810-2813.

Rosse, W. and H. Bunn (1998). "Hemolytic anemias and acute blood loss." HARRISONS PRINCIPLES OF INTERNAL MEDICINE: 659-671.

Stanton, R. C. (2012). "Glucose-6-phosphate dehydrogenase, NADPH, and cell survival." IUBMB life 64(5): 362-369.

Steinberg, M. H., B. G. Forget, et al. (2009). Disorders of hemoglobin: genetics, pathophysiology, and clinical management, Cambridge University Press.

Szabo P, Purrello M, Rocchi M, Archidiacono N, Alhadeff B, Filippi G, Toniolo D, Martini G, Luzzatto L, Siniscalco M. Cytological mapping of the human glucose-6-phosphate dehydrogenase gene distal to the fragile-X site suggests a high rate of meiotic recombination across this site. Proc Natl Acad Sci USA. 1984; 81(24):7855-7859.

Takizawa, T., I.-Y. Huang, et al. (1986). "Human glucose-6-phosphate dehydrogenase: primary structure and cDNA cloning." Proceedings of the National Academy of Sciences 83(12): 4157-4161.

Tindall, E. A., D. C. Petersen, et al. (2009). "Assessing high-resolution melt curve analysis for accurate detection of gene variants in complex DNA fragments." Human mutation 30(6): 876-883.

Vandaveer, W. R., D. J. Woodward, et al. (2003). "Redox cycling measurements of a model compound and dopamine in ultrasmall volumes with a self-contained microcavity device." Electrochimica acta 48(20): 3341-3348.

Yan, J.-b., H.-p. Xu, et al. (2010). "Rapid and reliable detection of glucose-6-phosphate dehydrogenase (G6PD) gene mutations in Han Chinese using high-resolution melting analysis." The Journal of Molecular Diagnostics 12(3): 305-311.

Vivek Chowdhry, Samarjit Bisoyi, Banabihari Mishra. "Perioperative challenges in a patient of severe G6PD deficiency undergoing open heart surgery". The journal of Annals of Cardiac Anaesthesia, Vol. 15, No. 1, January-March, 2012, pp. 50-53.

Von Seidlein L, Auburn S, Espino F, Shanks D, Cheng Q, McCarthy J, Baird K, Moyes C, Howes R, Ménard D, Bancone G. Review of key knowledge gaps in glucose-6-phosphate dehydrogenase deficiency detection with regard to the safe clinical deployment of 8-aminoquinoline treatment regimens: a workshop report. Malaria journal. 2013; 12(1):1.

WHO Working Group. Glucose-6-phosphate dehydrogenase deficiency. Bull World Health Organ 1989;67:601–11.

Yan J-b, Xu H-p, Xiong C, Ren Z-r, Tian G-l, Zeng F, et al. Rapid and reliable detection of glucose-6-phosphate dehydrogenase (G6PD) gene mutations in Han Chinese using high-resolution melting analysis. The Journal of Molecular Diagnostics. 2010;12(3):305-11.
(Krenková, Norambuena et al. 2009).

Youngster I, Arcavi L, Schechmaster R, et al. Medications and glucose- 6-phosphate dehydrogenase deficiency: an evidence-based review. Drug Saf 2010;33(9):713–26.

Zinkham, W. H. and Oski, F. A. (1996) Henna: a potential cause of oxidative hemolysis and neonatal hyperbilirubinemia. Pediatrics 97, 707–709.

Zollo M, Mazzarella R, Bione S, Toniolo D, Schlessinger D, D'urso M and Chen EY. Sequence and gene content in 52 kb including and centromeric to the G6PD gene in Xq28. Mitochondrial DNA. 1995; 6(1): 1-11.