MOLECULAR ANALYSIS AND CHARACTERIZATION OF GENES INVOLVED IN GLYCOGEN STORAGE DISEASE TYPE III AND SELECTED INHERITED METABOLIC DISORDERS

ILI SYAZWANA BINTI ABDULLAH

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

MOLECULAR ANALYSIS AND CHARACTERIZATION OF GENES INVOLVED IN GLYCOGEN STORAGE DISEASE TYPE III AND SELECTED INHERITED METABOLIC DISORDERS

ILI SYAZWANA BINTI ABDULLAH

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Ili Syazwana Binti Abdullah

Matric No: SHC120116

Name of Degree: Doctor of Philosophy

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Molecular Analysis and Characterization of Genes Involved in Glycogen Storage Disease Type III and Selected Inherited Metabolic Disorders

Field of Study: Genetics and Molecular Biology

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

MOLECULAR ANALYSIS AND CHARACTERIZATION OF GENES INVOLVED IN GLYCOGEN STORAGE DISEASE TYPE III AND SELECTED INHERITED METABOLIC DISORDERS

ABSTRACT

Inherited metabolic disorders (IMDs) comprise a wide range of diseases, and two groups, namely the glycogen storage diseases (GSDs) and methylmalonic aciduria (MMA) are the focus diseases in this study. Despite the fact that various genetic studies have been carried out and a mutation spectrum have been established in different populations worldwide, information on the genetic causes of these two categories of IMDs in the multi-ethnic Malaysian population is still lacking. In this study, a total of 27 patients were investigated, each suspected of having different types of glycogen storage diseases (namely GSD 1b, GSD III and FBS) or methylmalonic aciduria (MMA mut-type and MMA with homocystinuria cblC-type). Genomic DNA was extracted and the relevant genes were sequenced. Out of 27 patient samples, mutations were identified in 18 patients. In total, two different mutations were identified in the *SLC37A4* gene (for GSD 1b patients), nine different mutations in the *AGL* gene (for GSD III patients), three different mutations in the *SLC2A2* gene (for FBS patients), two different mutations in the *MUT* gene (for one MMA mut-type patient) and one mutation was observed in the *MMACHC* gene (for two MMA cblC-type patients). It is noteworthy that the c.2681+1G>A mutation previously reported to be commonly found for GSD III in Caucasian populations was also recurrently found in the Malay patients (n=7/22 alleles; 31.8%). Mutations that are yet unreported i.e. do not match any mutation in available databases were subjected to further analysis. However, due to time limitation, analysis of unreported mutations only focused on those causing GSD III. A total of six out of nine different mutations identified in GSD III patients were previously unreported (c.1423+1G>T, c.2914_2915delAA, c.3814_3815delAG, c.4333T>G, c.4490G>A,

c.4531_4534delTGTC). Analyses using prediction software Human Splice Finder showed that the mutation c.1423+1G>T disrupts normal splicing motif, and *in vitro* assays later revealed that this mutation causes intron retention. Analyses on mutation c.2914_2915delAA, c.3814_ 3815delAG and c.4490G>A predict formation of truncated proteins and therefore not expected to retain normal protein function. The c.4333T>G substitution results in a missense mutation, and this Tyr to Asp amino acid substitution was predicted to be "damaging" using the Polyphen2 analysis software. The c.4531_4534delTGTC mutation on the other hand leads to a frameshift and codon read-through, resulting in a protein that is longer by 14 amino acids. It is important to note that all unreported mutations were absent in 150 healthy control samples, rejecting the possibility of them being mere polymorphisms. Findings of this study show that the mutation spectrum for GSD III Malaysian patients is heterogeneous but a common mutation for a specific ethnic group may be present. Information gathered in this study will be useful for future diagnosis efforts and subsequent administration of correct treatments. Moreover, the information obtained herein can be used for genetic counselling in families with history of inherited metabolic disorder.

Keywords: Inherited metabolic disorders, glycogen storage diseases, methylmalonic aciduria, mutation

ANALISIS MOLEKUL DAN PENCIRIAN GEN-GEN YANG TERLIBAT DALAM PENYAKIT PENSTORAN GLIKOGEN JENIS III DAN BEBERAPA PENYAKIT METABOLIK TERWARIS YANG TERPILIH

ABSTRAK

Penyakit metabolik terwaris (*IMDs*) merangkumi pelbagai jenis penyakit dan dua kumpulan yang merupakan fokus di dalam pengajian ini ialah penyakit-penyakit yang melibatkan penstoran glikogen dikenali sebagai GSDs (bagi *glycogen storage diseases*) dan penyakit *methylmalonic aciduria.* Walaupun pelbagai kajian telah dijalankan dan spektrum mutasi gen-gen terlibat telah dikenalpasti di dalam pelbagai populasi berbeza di dunia, informasi tentang faktor genetik untuk dua kumpulan IMD bagi populasi Malaysia yang pelbagai etnik masih berkurangan. Di dalam kajian ini, sejumlah 27 pesakit yang disyaki menghidap pelbagai jenis GSD yang berbeza (GSD 1b, GSD III dan FBS) atau penyakit methylmalonic aciduria (MMA jenis mut dan MMA jenis cblC) telah dikaji. DNA genomik telah diekstrak dan gen-gen berkaitan telah dijujuk. Di kalangan 27 sampel pesakit, mutasi dijumpai di dalam 18 pesakit. Secara keseluruhan, dua mutasi yang berbeza telah dijumpai di dalam gen *SLC37A4* (untuk pesakit GSD 1b), sembilan mutasi berbeza di dalam gen *AGL* (untuk pesakit GSD III), tiga mutasi berbeza di dalam gen *SLC2A2* (untuk pesakit FBS), dua mutasi berbeza di dalam gen *MUT* (untuk seorang pesakit MMA jenis mut) dan satu mutasi di dalam gen *MMACHC* (untuk dua pesakit MMA jenis cblC). Penting untuk dinyatakan bahawa mutasi c.2681+1G>A yang dilaporkan kerap dijumpai di kalangan pesakit GSD III Kaukasia turut dijumpai dengan kekerapan tertinggi di kalangan pesakit Melayu (n=7/22 alel; 31.8%). Mutasi-mutasi yang masih belum dilaporkan (iaitu mutasi yang tiada padanan dengan maklumat dari pangkalan-pangkalan data sedia ada) telah dianalisa dengan lebih lanjut. Oleh kerana kekangan masa, analisis ke atas mutasi yang belum dilaporkan hanya ditumpukan kepada mutasi yang menyebabkan penyakit GSD III sahaja. Secara

keseluruhan, sejumlah enam dari sembilan mutasi berbeza yang telah dijumpai di kalangan pesakit GSD III masih belum dilaporkan (c.1423+1G>T, c.2914_2915delAA, c.3814_3815delAG, c.4333T>G, c.4490G>A dan c.4531_4534delTGTC). Analisa menggunakan perisian ramalan Human Splice Finder menunjukkan bahawa mutasi c.1423+1G>T akan menggangu motif penyambatan sedia ada, dan cerakinan/assai *in vitro* menunjukkan mutasi ini menyebabkan retensi intron. Analisa ke atas mutasi c.2914_2915delAA, c.3814_ 3815delAG dan c.4490G>A meramalkan protein terpangkas dan dijangka tidak mampu manjalankan fungsi seperti biasa. Mutasi c.4333T>G menyebabkan mutasi salah erti dan gantian asid amino Tir kepada Asp ini diramal oleh perisian PolyPhen-2 sebagai memberi kesan yang buruk ("*damaging"*). Manakala mutasi c.4531_4534delTGTC pula menyebabkan anjakan rangka translasi dan seterusnya mengakibatkan langkauan kodon penamat yang akan menghasilkan protein yang mempunyai lebihan 14 asid amino. Penting dinyatakan bahawa semua mutasi tidak dijumpai di dalam 150 sampel kawalan, menolak kemungkinan bahawa mutasi-mutasi ini tidak berlaku secara rawak dan adalah sekadar polimorfisme. Dapatan kajian ini menunjukkan spektrum mutasi untuk pesakit GSD III di Malaysia adalah heterogen tetapi satu atau lebih mutasi yang kerap dijumpai mungkin wujud bagi sesuatu kumpulan etnik. Informasi yang dikumpul dari kajian ini akan digunakan untuk usaha-usaha diagnosis dan seterusnya pemberian rawatan yang betul. Malahan, informasi yang didapati disini boleh digunakan untuk kaunseling genetik bagi keluarga yang mempunyai sejarah penyakit metabolik terwaris.

Kata kunci: Penyakit-penyakit metabolik terwaris, *glycogen storage diseases*, *methylmalonic aciduria*, mutasi

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my warm gratitude to my supervisor, Professor Dr. Zulqarnain Mohamed for his intellectual guidance and moral support throughout my study journey. Dr Zul, thank you for always being so positive and patient, for noticing my potentials and working with my flaws, for all of that I am forever grateful.

I would also like to extend my eternal gratitude to my family, especially to both my parents. Their dedication to provide many opportunities, unconditional love and unwavering faith in me are the true source of my determination. Ayah and Mak, thank you for being the most supportive loving parents and just thank you for everything. Words will not suffice to convey my gratitude to both of you. To my sisters Titi, Yeen and Tina, thank you for all the jokes and humour, fights and tantrums, and most importantly for the countless emotional support sessions. To Azam, Justin, Kayla, Micah, Kai and Noah, thank you for all the fun and joy each of you brings to the family.

Special thanks also go to Dr Sally Teh Ser Huy who has been an invaluable mentor in the laboratory and a great friend. Thank you also to the rest of MP2.0 laboratory members for their technical expertise and help, but also for their friendship developed over the years. To Pah, Fiq, Taufiq, Mira and Von, thank you for sharing many memorable times and great laughs.

I must extend my appreciation to my collaborators, Professor Dr. Sufin Yap, Dr. Keng Wee Teik and Dr. Ngu Lock Hock for providing the samples as well as for their input on the medical aspects of this study. I also like to thank Associate Professor Dr. Ng Ching Ching and Dr Azlina Ahmad Annuar for providing the healthy control samples. Lastly, I would like to thank University Malaya and Government of Malaysia for the grants and scholarship used for the completion of this study. Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

AML	acute myeloid leukemia
Ado-Cbl	adenosine cobalamin
α	alpha
A	ampere
AGL	amyloglycosidase-1, 4-alphatransferase
&	and
β	beta
BSA	bovine serum albumin
cAMP	cyclic adenosine monophosphate
cblC	cobalamin C
cDNA	complementary DNA
cfp	cyan fluorescence marker
ºC	degree celcius
ddNTPs	dideoxynucleotides
DGGE	denaturing gradient gel electrophoresis
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleoside triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ESI/MS/MS	electroscopy tandem mass spectrometry
FBS	Fanconi-Bickel syndrome
FPR	false positive rate
G6P	glucose-6-phosphate
G6Pase	glucose-6-phosphatase

G6PD	glycogen-6-phosphate dehydrogenase deficiency
GC-SF	granulocyte colony-stimulating factor
GDE	glycogen debranching enzyme
GLUT	glucose transporter protein
GlpT	*Escherichia coli* glycerol-3-phosphate transporter
GSD	glycogen storage disease
≥	greater than or equal to
HBS	hepes buffer solution
HepG2	human hepatocellular carcinoma cells
HGMD	Human Gene Mutation Database
HMG-CoA	5-hydroxy-3-methylglutaryl-coenzyme A
HSF	Human Splice Finder
IMDs	inherited metabolic disorders
LB	Luria Bertoni
MaxEnt	maximum entropy
MCM	methylmalonyl CoA mutase enzyme
MMA	methylmalonic aciduria
mRNA	messenger RNA
N-terminal	amino terminal
OHCbl	hydroxocobalamin
%	percentage
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffer solution
PCR	polymerase chain reaction
PMSF	phenylmethanesulphonylfluoride
R	side chain

RE	restriction enzyme
rcf	relative centrifugal force
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNase A	ribonuclease A
rpm	revolutions per minute
RT-PCR	reverse transcription polymerase chain reaction
sdH_20	sterile distilled water
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium-dodecyl sulphate polyacrylamide gel electrophoresis
SNPs	single nucleotide polymorphism
SSCP	single strand conformational polymorphism
Taq	*Thermus aquaticus*
TCA	trichloroacetic acid
TBE	Tris Borate EDTA
TBS	Tris-buffered saline
TBST	Tris-buffered saline with Tween 20
TE	Tris-EDTA
TPR	true positive rate
UDP-glucose	uridine diphosphate glucose
UhpT	*Escherichia coli* hexose phosphate transporter
UTR	untranslated region
Vmax	maximum velocity
WT	wild type
-COOH	carboxyl group
$-NH_2$	amine group

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

This chapter provides a synopsis of the research work undertaken in this study. The research focused on identifying the mutations that causes glycogen storage disease type III and other related inherited metabolic disorders, with the hope of aiding accurate diagnosis as well as establishing the mutation spectrum for these disorders. The overview of the work is explained in several sub-sections as follows.

Section 1.1 explains the background of the study as an introductory account to the thesis. Section 1.2 justifies the undertaking of the research work, Section 1.3 provides the statement of research problems and Section 1.4 outlines the statement of research objectives. Section 1.5 discusses the significance of the research work and finally in Section 1.6, the layout of the thesis was provided.

1.1 Background Information

Inherited metabolic disorders (IMDs), which are a group of diseases that affects various cellular metabolism processes was first documented in the 19th century by British physician Sir Archibald Edward Garrod (Dronamraju, 1992). During their early discovery, IMDs were found to be closely related to the breakdown of lipids and amino acids. As time progresses, more types were discovered and amounted to significant causes of morbidity and mortality despite individual types being considered rare (Mak et al., 2013). In the recent years, IMDS have been classified into several categories, and examples include disorders of carbohydrate metabolism, amino acid metabolism disorder, urea cycle disorder and organic acid metabolism disorder. The categories were made according to the metabolic pathway in which they affected.

The present study focuses on two types of IMDs which are the disorders of carbohydrate metabolism (glycogen storage diseases) and disorders of organic acid

metabolism (methylmalonic aciduria). Glycogen storage diseases (GSDs) are disorders that affect the break down and/or storage of glycogen. The glycogen metabolism pathway involves various enzymes, and mutations in any of these enzymes may cause different types of GSDs. On the other hand, methylmalonic aciduria (MMA) is a metabolic disorder that affects the breakdown of amino acids and certain fatty acids. Similar to GSDs, there are different types of MMA that are caused by mutations in different genes. In this study GSD types that are being investigated include GSD type 1b, GSD type III and FBS while for MMA, only MMA mut-type and MMA with homocystinuria cblC-type are included.

Including the disorders mentioned above, there are at present more than 6,000 known single gene disorders, involving more than 8000 genes with an excess of 203,000 genetic lesions, and all these information is made available to researchers through the establishment of the Human Gene Mutation Database (HGMD, http://www.hgmd.cf.ac.uk/ac/all.php). With relevance to this study, the database had collated more than 101 reported mutations for GSD 1b, 155 reported mutations for GSD III, 66 reported mutations for FBS, 243 mutations for MMA mut-type and 73 mutations for MMA cblC-type. These data are clearly indicative of high levels of heterogeneity in each disease. However, data collated in the HGMD largely represent the European and East Asian populations, and there is not much data available for the South East Asian populations in general or Malaysian population specifically, and this fact was one of the motivation for the current study to be carried out.

1.2 Motivation of Research

Comprehensive information on IMDs cases in Malaysia is not available, and pockets of information if available are mostly incomplete. The lack of data covers a wide range of information category, from the number of incidences for each disease,

rate of mortality and morbidity, or data on common mutations for a particular disease is unknown. Thong and Yunus (2008) for example reported that the actual number of IMD patients in Malaysia is unknown and with only several short reports on IMDs available. IMDs were previously perceived as rare in many populations but the fact that it was not recognized and documented, especially in Asian countries may be the underlying factor of the low number of cases reported (Tan et al., 2006; Yunus et al., 2016). This underlines the need for more study to be carried out to contribute to the knowledge of IMDs in Malaysia.

With limited knowledge of the clinical spectrum of IMDs locally, diagnosis of closely related disease within each category of IMDs still presents a major challenge for clinicians. GSD patients for example, are mostly presented with hepatomegaly (enlargement of the liver) but diverse phenotypes of different GSD subtypes complicate the diagnosis process and subsequently delay the administration of appropriate treatment. Low awareness of atypical and variable presentations in IMD (Leong et al., 2014), along with high index of suspicion required to make diagnosis increase the challenge among clinicians to come up with a diagnosis. While various research effort have been channelled towards understanding this group of disorders, the molecular basis as well as the reason for variability in the phenotypes are not clearly understood (Lucchiari et al., 2007). Studies to understand such variation will provide definitive diagnosis as well as better diagnostic and preventive measures to be undertaken.

The distribution and spectrum of mutations causing various IMDs in Malaysia where the population comprises various ethnic groups are also still poorly known (Leong et al., 2014; Yunus et al., 2016). Various studies have suggested that the types of disorders, the gene mutations found and the mutation spectrum differs according to different ethnic groups (Okubo et al., 2000; Rampal et al., 2012; Shaiu et al., 2000). In

populations where high rate of consanguinity occurs, repeated occurrence of certain mutations have been reported (Aoyama et al., 2009; Santer et al., 2001). Results obtained in this study should provide important insights on the molecular genetic aspect of the disorders in question, and highlight any possible correlation with any ethnicity, if any. Identification of ethnic-specific mutation will be useful for future targeted treatment efforts. Proper and early diagnosis is crucial as appropriate treatment can then be provided, preventing serious clinical complications.

1.3 Statement of Problem

IMDs are presented with both heterogeneous features as well as features that are common across different types (Mak et al., 2013). Heterogeneous clinical features seen in a particular disease which can be due to varied residual enzyme activity or involvement of various tissues slows down diagnostic efforts (Özen, 2007). Common clinical features across different types of metabolic disorder or similar biochemical profiles across the same group of IMDs, on the other hand, prevent fast diagnosis to be made. Confirmatory enzyme assays, while offering definitive results, are very invasive and traumatic for the patients, and parents especially when infants or young children are involved. Genetic analysis is therefore the preferred method for confirmatory diagnosis (Kishnani et al., 2010; Mahdieh & Rabbani, 2013). Efficient genetic analyses however require pre-existing information on the possible mutations that is prevalent for a particular disease. The fact that some mutations have been found to be prevalent in certain populations, concise information on the common mutation in a particular population is also crucial. As mentioned in the preceding section, this information is at present lacking for the multi-ethnic Malaysian population.

In addition, the apparent lack of information also hampers development of specific treatments to target specific physiological reaction (Burton, 1998; Hörster &

Hoffmann, 2004). The level of responsiveness towards a particular treatment varies and this may be dependent on the mutations present (Crushell et al., 2010). Determining the correct treatment is crucial as genetic disorders cannot be cured but managing the symptoms can avoid life-threatening disease manifestation. Narrowing the potential mutation that affecting particular patient provides better chance of administering the correct treatment.

1.4 Statement of Objectives

This study aims at facilitating the diagnostic process and treatment administration of IMD specifically the carbohydrate metabolic disorders and the organic acids metabolic disorders. This can be achieved with the following set of objectives.

1. To determine the molecular genetic cause(s) of GSD III in the Malaysian patients.
2. To determine the molecular genetic causes of less common GSD types (GSD 1b and FBS) in the Malaysian patients.
3. To determine the molecular genetic causes of selected IMDs (MMA mut-type and MMA with homocystinuria cblC-type) in the Malaysian patients.
4. To verify the pathogenicity of unreported mutation.
5. To examine the effects of unreported mutation on protein function.

1.5 Significance of Research

The causative mutations affecting glycogen storage diseases (type Ib, type III and FBS) as well as methylmalonic aciduria (mut-type and cblC-type) among Malaysian patients were identified. In Malaysia, information on IMDs is scarce and scientific research on the molecular aspect of GSD III specifically has not been carried out. This reflects a gap in the genetic knowledge of this disease that is crucial for future treatment plans and prevention. With the current direction towards personalized

medicine, albeit less active in the field of inherited metabolic disorders, genetic information is invaluable.

Mutation identification also allows for common mutations in a population to be deduced. This narrows the range of test to be carried out and minimize cost for diagnostic purposes, as well as facilitating genetic counselling or testing for family members of the affected individual. Early recognition and treatment correlate with better prognosis. Other than that, knowledge on the mutations supports advances in treatment as the mechanism of which a mutation causes the disease can be studied.

1.6 Thesis Layout

This thesis is comprised of six chapters.

Chapter 1: Introduction. It briefly introduces the research topic and potential research problem to be addressed. This chapter also includes the statement of research objectives and its significance.

Chapter 2: Literature Review. It provides a comprehensive review of the research domains; glycogen storage diseases and methylmalonic aciduria. This was carried out to identify a potential knowledge gaps in IMDs, which led to further investigation among local patients. The literature reveals that IMDs reports was largely available, however local reports are scarce.

Chapter 3: Materials and Methods. This chapter explains the experimental procedures and standard molecular techniques used in this study. The methods were organized as to describe the mutation screening work and further analysis on unreported mutations.

Chapter 4: Results. It reports the gathered data for mutation screening of the following genes; *AGL, SLC37A4, SLC2A2, MUT* and *MMACHC*. Predicted effects of unreported mutations on protein function were also included.

Chapter 5: Discussion. This chapter discusses the findings in line with the research objective achievements. Findings of this study were compared with known data and argued.

Chapter 6: Conclusion. It concludes the research achievements by revisiting the objectives of this study. A short account of research contributions and suggestions for future work were also included.

CHAPTER 2: LITERATURE REVIEW

This chapter presents an overview of the two categories of IMDs and its associated diseases addressed in this study, with the objective to give brief descriptions of each disease and to highlight the importance of mutation screening.

Section 2.1 provides description about carbohydrate, the nutrient that is affected in carbohydrate metabolic disease. Section 2.2 describes the glycogen storage disease in general and the different types included in this study. Section 2.3 recount lipids and protein involvement in organic acid metabolism disorder. Section 2.4 describes the methylmalonic aciduria and the different types included in this study. Section 2.5 briefly discusses about mutation screening of inherited diseases, validation of polymorphism and functional study on mutated gene.

2.1 Carbohydrates as Source of Energy

The human body utilizes food that we eat to produce energy, to regulate metabolism as well as to build and repair tissues. Energy production is however the main priority as it is needed to fuel cells and carries out all other body functions. There are various sources of foods that can be used to generate energy. Carbohydrate is the primary source of energy (Jequier, 1994) while lipids and proteins are the alternative sources.

Carbohydrates or saccharides are biological molecules consisting of carbon (C), hydrogen (H) and oxygen (O) atoms with $C_m(H_2O)_n$ as the compound formula. It is divided into four chemical groups: monosaccharide (sugar), disaccharides, oligosaccharides and polysaccharides. Monosaccharides (such as glucose, dextrose and galactose) are the most basic units of carbohydrates and form the building blocks of disaccharides (e.g. sucrose and lactose), oligosaccharides (e.g. fructooligosaccharides

and galactooligosaccharides) and polysaccharides (e.g. cellulose and starch) (Ghazarian et al., 2011). The human body utilizes carbohydrate in the form of glucose and stored it in the form of glycogen as blood glucose concentration has to be maintained (Jensen et al., 2011).

Glycogen is a multi-branched polysaccharides of glucose that forms a spherical shape structure (Adeva-Andany et al., 2016). The centre of a glycogen molecule is glycogenin, a self-glucosylating protein (Lomako et al., 2004) that transfer glucose residues from activated form of glucose, uridine diphosphate glucose (UDP-glucose) to itself. Glucose residues linked to each other forming the glycogen molecule, where Whelan's model for glycogen described two different kinds of glucose chains: the linear non-branched A-chains and the branched B-chains (Figure 2.1a). These residues are linked by 1, 4-glycosidic linkage to forms the linear chains while the branches are formed by 1, 6-glycosidic linkage (Roach et al., 2012) (Figure 2.1b). Each B-chain has two branches on it, creating further A or B-chains. The two branches are separated with four glucose residues and there are tails of four glucose unit after the second branch in the B-chains (Melendez-Hevia et al., 1993). With this model, glucose can be released from the non-reducing end of all the non-branched outer A-chains. Only a fraction of these outer A-chains glucose residues are accessible by degradative enzyme as the reaction stalls four residues from a branch before action of the debranching enzyme took place (Roach et al., 2012). The whole molecule of glycogen is composed of over 55000 glucose residues with average diameter of 44 nm. Electron microscopy detection has also shown that the glycogen molecules are larger in the liver than in the muscle (Adeva-Andany et al., 2016).

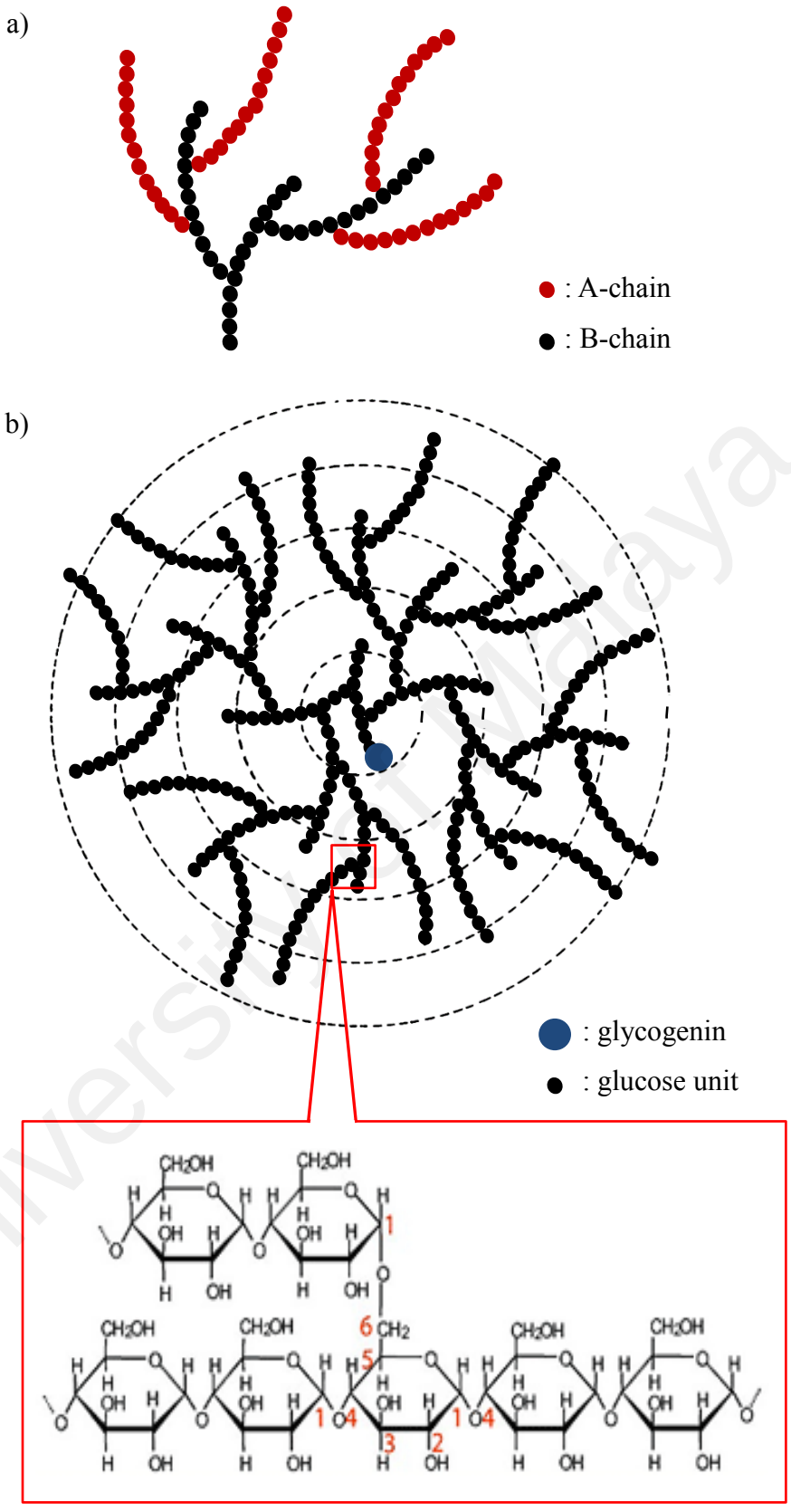


Figure 2.1: Schematic diagram of glycogen. (a) The unbranched A-chains and branched B-chains of glycogen (b) The glycogen molecule is spherical and organized into concentric tier. Glycogenin, primer in glycogen synthesis is at the core. Inset showing linear chains formed via α-1,4 linkages and the branch point formed via α-1,6 linkages.

Liver and muscles are two major organs that stores glycogen but with varying concentrations. The liver has higher concentrations of glycogen compared to the skeletal muscle. However, in term of overall mass, more glycogen is stored in the skeletal muscle due to greater mass (Jensen et al., 2011; Taylor et al., 1996). Minor glycogen stores are present in other tissues such as the heart and brain but only the liver glycogen directly contributes to glucose release into the blood (Jensen et al., 2011). Lack of glucose-6-phosphatase prevent skeletal muscles to release glucose, instead muscle glycogen can be broken down into lactate that serve as energy in 'fight or flight' situations. Heart and brain glycogen is also used for survival as it can be used to generate anaerobic energy during short-term oxygen deficiency (Falkowska et al., 2015).

Glycogen metabolism follows distinct regulatory pathways with various enzymes involved (Figure 2.2). During glycogen synthesis, glucose is initially converted to glucose-6-phosphate by hexokinase and then converted to glucose-1-phosphate by phosphoglucomutase. Glycogen synthase then catalyse the formation of α-1, 4 linkages. Branching enzyme forms the α-1, 6 linkages that generates the branches of glycogen. Meanwhile, degradation of glycogen involves hydrolysis of glycogen that is catalysed by glycogen phosphorylase and debranching enzyme. Glycogen phosphorylase release the terminal glucose residue from outer A-chain until only four glucose units left on the outer chain. Finally, the debranching enzyme will release the remaining glucose units ready to be utilised by the body (Bollen et al., 1998). Mutations of the genes encoding the various enzymes involved in glycogen metabolism cause glycogen storage diseases.

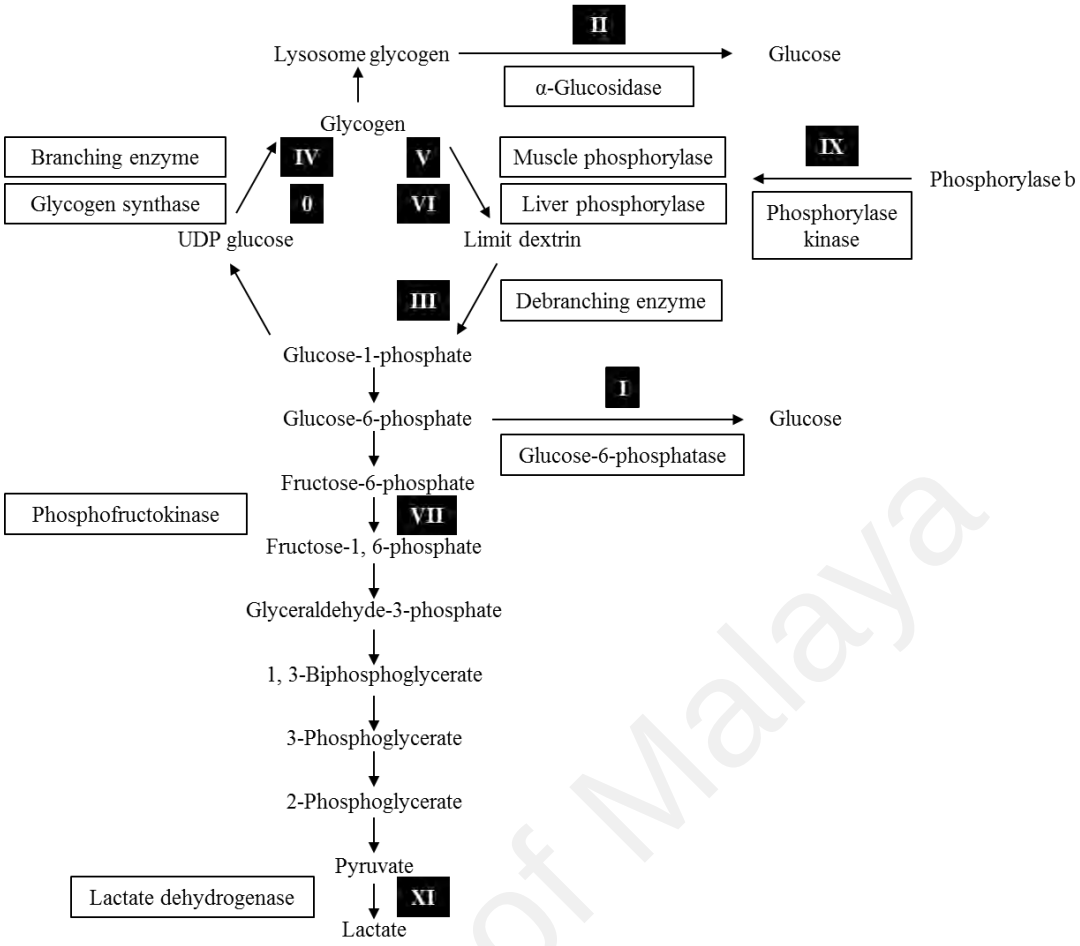


Figure 2.2: Glycolytic pathways. The texts in boxes indicate the affected enzymes associated with different types of glycogen storage diseases and the types are designated with Roman numerals.

2.2 Glycogen Storage Diseases

Glycogen storage diseases (GSDs) are inherited metabolic disorders that disrupt the glucose homeostasis. There are various types of GSDs, with each type are considered rare. In general there are two broad GSDs categories, those affecting synthesis of glycogen and those affecting its breakdown. Absence or defect in a particular enzyme determines the type of GSD (Chen, 2011) (Table 2.1) (Figure 2.2).

GSD type I for example is reported to occur at 1 in 100,000 to 300,000 (Froissart et al., 2011) while type III occurs at 1 for every 100,000 birth in the United States (Hendriksz & Gissen, 2014; Parvari et al., 1997). Overall incidence of GSDs had been reported as approximately 1 case per 20,000-43,000 live births (Özen, 2007) and

this disease is considered as a major category of human inborn errors of metabolism that affect mostly infants.

Patients suffering from GSD are presented with different clinical, biochemical and molecular features depending on the type of GSD (Chen, 2011; Hendriksz & Gissen, 2014). The organs involved in glycogen storage are mainly the liver and muscle. In cases where liver is involved, hepatomegaly and hypoglycaemia are the common clinical presentations of all types of GSDs (Özen, 2007). Infants and children patients with muscle involvement usually presented with muscle cramps and growth retardation. In severe cases, muscle weaknesses and heart problems develop that leads to death. Meanwhile, suspected juveniles or adults usually presented with fatigue, myositis or myopathy (Das et al., 2010).

Diagnosis of GSD is made based on the symptoms presented as well as biochemical assay findings. Hepatomegaly and ketotic hypoglycaemia with fasting as well as elevated transaminases and CK is the telltale sign of GSD III. Biopsy of the involved organ and biochemical assay performed to confirm the diagnosis. Biochemical assay is the method of choice for confirmatory diagnosis but the assay should be performed immediately after tissues biopsy to prevent tissue deterioration. Furthermore, the biopsy procedure itself is very invasive and risky to the patient (Shin 1990). Molecular detection, on the other hand allows diagnosis to be carried out without the need for extensive diagnostic profiles as well as prevention of invasive biopsy procedures to be undertaken.

Table 2.1: Glycogen storage diseases (GSDs) types, their affected enzymes and mutated genes.

Type	Affected enzyme	Mutated gene (s)
Type 0	Glycogen synthase	*GYS1, GYS2*
Type I	Glucose-6-phosphatase	*SLC37A4*
Type II	Acid maltase	*GAA*
Type III	Glycogen debranching enzyme	*AGL*
Type IV	Glycogen branching enzyme	*GBE1*
Type V	Muscle phosphorylase	*PYGM*
Type VI	Liver phosphorylase	*PYGL*
Type VII	Muscle phosphofructokinase	*PFKM*
Type IX	Phosphorylase kinase	*PHKA1, PHKA2*
Type XI	Lactate dehydrogenase	*LDHA*
FBS	Glucose transporter	*SLC2A2*

2.2.1 Glycogen Storage Disease Type III (GSD III)

Glycogen storage disease type III (GSD III; Cori-Forbes Disease; MIM #232400) is an autosomal recessive metabolic disorder. It is caused by deficiency of the glycogen debranching enzyme (GDE) which causes accumulation of abnormally structured glycogen with truncated outer chains (Okubo et al., 1998). The disease prevalence is about 1/100,000 but occur at higher rate in North African Jewish communities with prevalence of about 1/5000 (Hendriksz & Gissen, 2014).

2.2.1.1 Clinical Symptoms of GSD III

Patients of GSD III are presented with heterogeneous clinical manifestation that includes hepatomegaly (enlarged liver), hypoglycaemia (low blood sugar), hyperlipidaemia (high blood lipid), hypotonia (decreased muscle tone), short stature, progressive myopathy and cardiomyopathy (Coleman et al., 1992; Ko et al., 2013). Hepatomegaly occurs due to accumulation of abnormally structured glycogen in the liver. Glycogen is stored in the liver, during normal condition glycogen is broken down

and utilised. However, defective enzyme in GSD III patients prevents the breaking down hence the glycogen accumulation. Hepatomegaly and hepatic symptoms usually improve with age but progressive liver cirrhosis and failure may occur.

Hypoglycaemia occurs when there is not enough free glucose present in the blood circulation to be utilized for energy usage. This condition usually gives rise to hyperlipidaemia. In patients with progressive myopathy, it may not be apparent in infants or children but the severity increases after the third or fourth decade of life. Adult-onset myopathies may be distal or generalised. Atrophy of the leg or hand that leads to diagnosis of motor neurone disease may occur in patients with distal myopathy. The generalised myopathy tends to be more severe by affecting the respiratory muscles (Kishnani et al., 2010). However, despite cardiomyopathy is a common finding, cardiac failure is rare (Hendriksz & Gissen, 2014). Peripheral neuropathy had also been reported and this contributes to the weakness and the neurogenic features of some patients (Kishnani et al., 2010).

2.2.1.2 Diagnosis and Treatment of GSD III

GSD III is diagnosed upon demonstration of abnormal glycogen in the liver and/or muscle which contains glycogen with very short outer chains. Other than that, deficient debranching-enzyme activity in skin fibroblast or lymphocytes is important to diagnose GSD III. Gene mutation analysis can also be made.

Currently there is no treatment available for GSD III patients. However, dietary management is employed to maintain good blood glucose and correction of hyperlipidaemia. Frequent carbohydrate-rich meals with corn starch supplementation are the current strategy (Basit et al., 2014; Chen, 2011). High-protein diet may also be employed. In cases where cirrhosis and carcinoma occur, liver transplantation may be performed.

2.2.1.3 Molecular Genetics of GSD III

The gene that is central to the pathogenesis of GSD III is the amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase gene, also known as the *AGL* gene (MIM #610860). Reported molecular analysis of GSD III showed that the mutation spectrum of *AGL* gene in GSD III patients varies and this variation depends on ethnic group. There is high level of genetic heterogeneity involved with this disease in populations such as Japanese (Okubo et al., 2000), Italian and Caucasians (Shaiu et al., 2000). Meanwhile in populations where high rates of consanguinity occur, recurrent mutations have been reported such as the p.R408X in the Faroe Islands (Santer et al., 2001) and p.W1327X in the Egyptian and Turkish populations (Aoyama et al., 2009). Among these reported mutations, most are nonsense mutation where a truncated form of the protein is the expected outcome. Meanwhile, missense mutations of the enzyme may cause various alterations towards the protein. In vitro enzymatic activity may provide insight into the effect of the nonsense as well as missense mutations to the GDE function.

2.2.1.4 The *AGL* Gene and Glycogen Debranching Enzyme

Enzyme deficiencies occur largely due to mutations in the *AGL* gene (Figure 2.3). It comprises of 35 exons, spans a genomic sequence of 85 kb in length and is located on chromosome 1p21. There are six isoforms of the *AGL* gene as a result of alternative splicing in the 5' UTR of the gene (Lucchiari et al., 2007). The major isoform, isoform 1 (GenBank accession no. U84007.1) is widely expressed and encodes a 7 kb mRNA. Isoforms 2, 3 and 4 are muscle specific while minor isoforms 5 and 6 can be found both in liver and muscle (Bao et al., 1996).

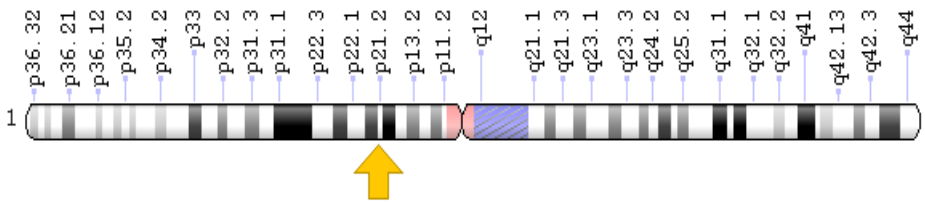


Figure 2.3: Schematic diagram of *AGL* gene on chromosome 1p21.
(Source: http://www.ncbi.nlm.nih.gov/genome/tools/gdp)

Glycogen debranching enzyme (GDE) is a large monomeric protein with a molecular weight of 175 kDa and it contains two catalytic activities on a single polypeptide chain (Yang et al., 1992). The two catalytic subunits are oligo 1,4-1,4 glucanotransferase (E.C.2.4.1.25) and amylo 1,6 glucosidase (E.C.3.2.1.33). Each enzymatic activity occurs at separate catalytic sites on a polypeptide chain and can function independently of each other (Ding et al., 1990). Full debranching enzyme activity however requires both catalytic activities. The enzyme is important for the breakdown of glycogen into glucose (Figure 2.4).

During glycogen breakdown, the glycogen phosphorylase hydrolyses the outer glucose residues leaving short outer branches. The transferase function of GDE transfer three glucose residues from one short branch to adjacent branch leaving only one glucose residue. The glucosidase function of GDE then release this branch-point glucose by hydrolysis. When the enzyme is defective, glycogen is not properly degraded causing accumulation of abnormally structured glycogen.

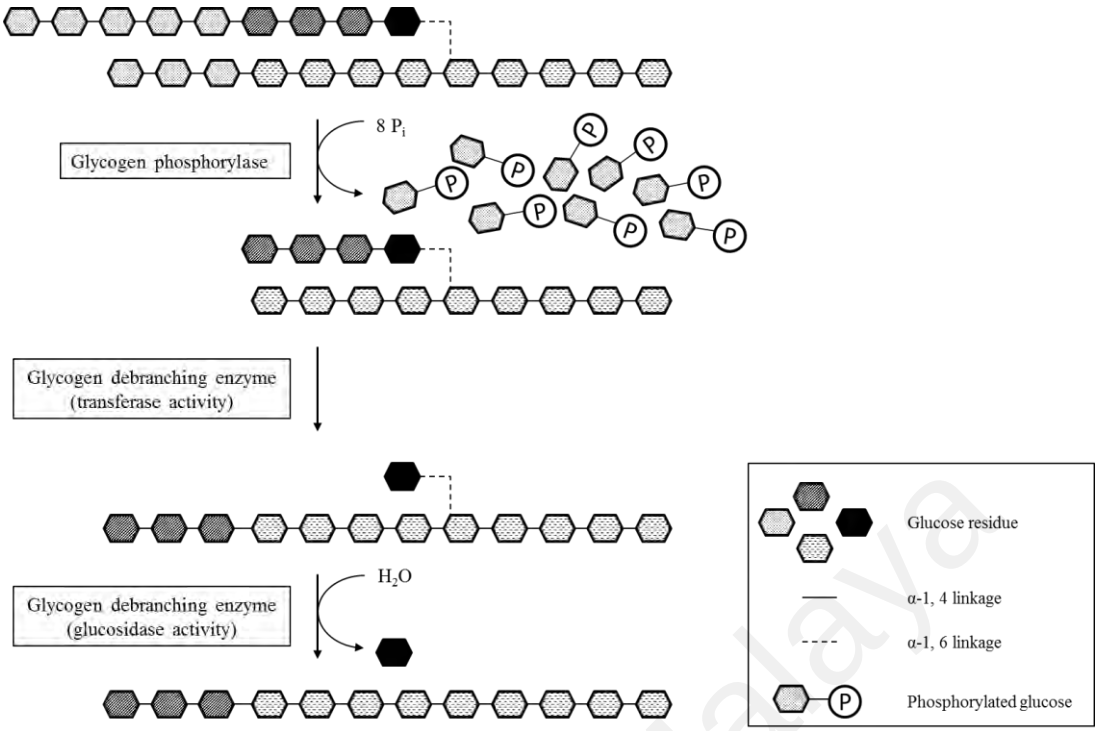


Figure 2.4: Schematic diagram of glycogen breakdown by glycogen debranching enzyme.

2.2.2 Glycogen Storage Disease Type 1 (GSD I)

Arion et al. (1975) postulated that at least two components of the ER participate in the process of glucose-6-phosphate (G6P) hydrolysis. The first component is a G6P specific translocase that shuttles G6P across the membrane. This was later identified as glucose-6-phosphate transporter which is an endoplasmic reticular trans-membrane protein responsible for transporting the glucose-6-phosphatase across the membrane. The second component is a relatively nonspecific phosphohydrolase, phosphotransferase that is located in the luminal surface of the membrane. This component is now identified as glucose-6-phosphatase (G6Pase); a transmembrane enzyme with catalytic activity on the luminal surface of the endoplasmic reticulum responsible in hydrolysing G6P to produce glucose (Figure 2.5). Defects in any of the two components of the enzymes cause glycogen storage disease type 1 (GSD 1; Von Gierke disease; MIM#232220). There are two major subgroups of the disease, GSD 1a and GSD 1b. The more prevalent form, GSD 1a is caused by deficiency of the glucose-

6-phosphatase activity (Chou & Mansfield, 1999). GSD 1b on the other hand is caused by deficiencies of the glucose-6-phosphatase transporter enzyme (Hiraiwa et al., 1999).

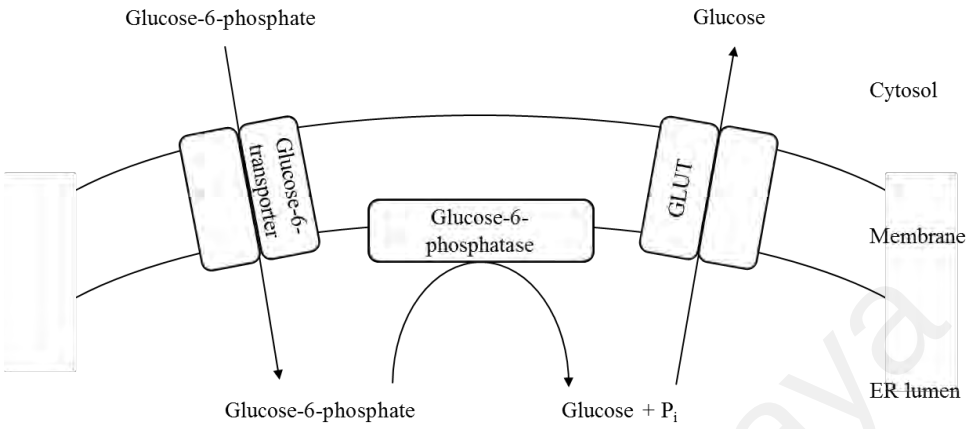


Figure 2.5: Glucose-6-transporter and glucose-6-phosphatase association with glucose formation.

2.2.2.1 Glycogen Storage Disease Type 1b (GSD 1b)

GSD 1b was described by Narisawa et al. in 1978 when he was explaining the paradox of four patients with GSD and normal glucose-6-phosphatase activity (Narisawa et al., 1978). He showed that these patients have a deficiency in their glucose-6-phosphate transporter enzyme, one of the two proteins that forms the complex needed for catalysing the final step of glycogenolysis and gluconeogenesis (Kishnani et al., 2014).

2.2.2.2 Clinical Symptoms of GSD 1b

Patients of GSD 1b are presented with the same general features as seen in GSD 1a cases which are low blood sugar (hypoglycaemia), high blood lipid, lactic academia, short stature, and delayed puberty (Chou et al., 2010). Symptomatic hypoglycaemia may appear soon after birth and typically appear only when the interval between feedings increases. Liver adenomas and high blood uric acid had been reported but occur very rarely (Kishnani et al., 2014). Some of GSD 1b patients had been reported to be presented with hypothyroidism and thyroid autoimmunity (Melis et al., 2007).

While most metabolic abnormalities of both GSD 1a and 1b are similar, patients of GSD 1b also suffer from infectious complications due to both heritable neutropenia and deficiencies of neutrophils and monocytes (Annabi et al., 1998). The severity varies from mild to recurrent bacterial infections. This contributes to increase severity of GSD 1b patients compared to GSD 1a patients. Neutrophil counts however may be normal during the first two years of life (Chou et al., 2010). Other effect of neutropenia manifestation among GSD 1b patients is the development of inflammatory bowel disease (IBD), which is similarly presented in idiopathic Crohn disease (Dieckgraefe et al., 2002; Visser et al., 2000). Children with GSD 1b are also prone to oral complications such as recurrent mucosal ulceration, gingivitis and rapidly progressive periodontal disease (Kishnani et al., 2014). Neutropenia and susceptibility to infections are now attributed to specific abnormalities in the neutrophil production and function. Glucose is required for neutrophil's metabolic burst, where mutation in the glucose-6-phospate transporter causes apoptosis of developing neutrophils, ineffective neutrophils production and neutropenia (Froissart et al., 2011).

2.2.2.3 Diagnosis and Treatment of GSD 1b

Diagnosis of GSD 1b is made upon presentation of hepatomegaly and hypoglycaemia. GSD I and GSD III have several features in common, but GSD 1 typically present earlier in life (first few months of life) and with severe fasting hypoglycaemia within 3-4 hours after feeding. Elevated uric acid and lactate levels as well as serum concentrations of hepatic transaminase (aspartate aminotransferase and alanine aminotransferase) are also increased.

Similar to most glycogen storage diseases, the main treatment of GSD 1b is to maintain normal blood glucose concentrations (Bali et al., 2016). This is achieved by a combination of continuous nasogastric tube feeding, uncooked corn starch and regular

oral feeds. Uncooked corn starch acts as a slow-release form of glucose. Lipid-lowering drugs such as HMG-CoA reductase inhibitors and fibrates are used to manage hyperlipidaemia. Granulocyte colony-stimulating factor (GC-SF) is used to correct the neutropenia and neutrophil function by increases both the proliferation of myeloid precursors and the functional capacity of neutrophils (Hendriksz & Gissen, 2014). It is also used to reduce infections (Chou et al., 2010). Despite the therapy improves neutropenia and decreases infections number or severity (Visser et al., 2002), development of splenomegaly had been reported in patients receiving the therapy (Calderwood et al., 2001). Periodic measurement of the spleen every three months are therefore required for those on GC-SF therapy (Bali et al., 2016). Complication that may arise from GC-SF therapy is acute myeloid leukaemia (AML) as had been reported in several studies. Patients receiving this therapy was therefore recommended to have regular bone marrow examinations (Chou et al., 2010).

2.2.2.4 Molecular Genetics of GSD 1b

Enzyme deficiencies usually occur due to mutations on the glucose-6-phosphatase transporter 1 gene (Narisawa et al., 1978), known as the *G6PT* gene or also referred to as *SLC37A4* gene (MIM # 602671).

Reported molecular analysis of GSD 1b showed that the mutation spectrum of *SLC37A4* gene in GSD 1b patients varies and ethnic variability exists (Chou et al., 2010). However, some ethnic group-specific common mutations account for ~90 % of known disease alleles. Several prevalent mutations had been reported in certain ethnic groups. These mutations include the c.1042_1043delCT and p.G339C seen among Caucasian GSD 1b patients (Veiga-Da-Cunha et al., 1999), or the p.W118R which is prevalent in Japanese patients (Chou et al., 2010; Kure et al., 1998).

2.2.2.5 The *SLC37A4* Gene and Glucose-6-Phosphate Transporter Enzyme

The human *SLC37A4* gene was first sequenced in 1997 by using bacterial transporter for phosphate ester as reference sequence. The sequenced human cDNA that encodes 46 kDa transmembrane protein is highly similar to its bacterial transporter homolog; glycerol-3-phosphate transporter (GlpT) and hexose-phosphate transporter (UhpT) (Chou & Mansfield, 2014). The putative human translocase hydrophobic profile is nearly superimpose with both GlpT and UhpT indicating *SLC37A4* as very hydrophobic, pointing to the fact that it is a membrane protein (Gerin et al., 1997).

Annabi reported the linkage of GSD 1b locus to genetic markers spanning a 3-cM region on chromosome 11q23 (Annabi et al., 1998). At about the same time, Marcolongo and his team also located the *SLC37A4* gene to chromosome 11 (Marcolongo et al., 1998) (Figure 2.6) and this was also supported by research data from Ihara's group. The latter group more precise localization was achieved by fluorescence in situ hybridization technique (Ihara et al., 1998).

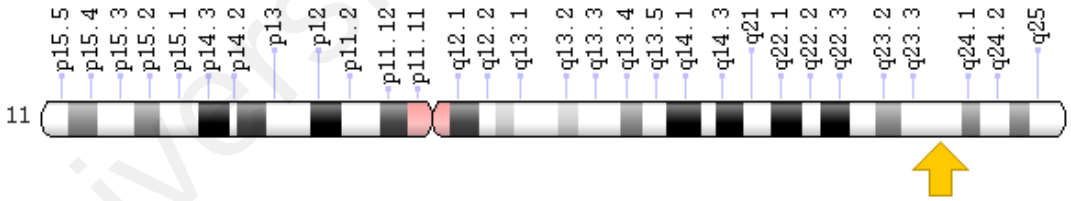


Figure 2.6: Schematic diagram of *SLC37A4* gene on chromosome 11q23. (Source http://www.ncbi.nlm.nih.gov/genome/tools/gdp)

The *SLC37A4* gene comprises nine exons (Veiga-Da-Cunha et al., 1998) and spans a genomic region of approximately 4 kb in length (Marcolongo et al., 1998). There are various transcript variants of the gene with variant one representing the longest transcript. Variant 1, 4 and 5 encode the same protein or isoform one (NCBI, http://www.ncbi.nlm.nih.gov). The variant that is expressed in the liver is a splicing variant missing exon seven. The small exon seven was however present in mRNAs of the brain (Gerin et al., 1997), heart and skeletal muscle (Lin et al., 2000).

The glucose-6-phosphatase transporter enzyme was studied intensively by Arion in 1971 (Arion et al., 1971). The study found that the enzyme works in tandem with glucose-6-phosphate enzyme for glucose homeostasis. These observations can be explained by a model for glucose-6-phosphatase in which the enzyme is positioned within the membrane of the endoplasmic reticulum. In order for substrate to reach the active site, transporters need to aid the reaction (Arion et al., 1972).

Later, Arion postulated that at least two components of the ER participate in the process of glucose-6-phosphate (G6P) hydrolysis based on the substrate-transport model (Arion et al., 1975). One is a G6P translocase that shuttles G6P across the membrane and two is a relatively nonspecific phosphohydrolase and phosphotransferase located in the luminal surface of the membrane (Narisawa et al., 1983). The G6P translocase was later termed as glucose-6-phosphatase transporter and found to be specific for glucose-6-phosphate as opposed to its 2-epimer mannose-6-phosphate. In Arion's substrate-transport model, he also postulated the involvement of another component, a glucose transporter. The glucose transporter (GLUT7) is however been now reported as artefact (Veiga-Da-Cunha et al., 1998).

Pan et al. in 1999 published a study that reported glucose-6-phosphate transporter enzyme as an enzyme that anchored to the ER by 10 transmembrane domains (Pan et al., 1999). Around the same time G6PT protein was also found to co-function with G6Pase-α for efficient G6P hydrolysis (Hiraiwa et al., 1999). Because of the functional G6PT/ G6Pase-α complex, absence of G6Pase-α activity causes low basal level of G6P uptake activity while mutation in the G6PT cause decrease or complete abolishment of G6P transport activity (Chen et al., 2002). This optimum activity of this complex is critical for the maintenance of interprandial glucose homeostasis (Chou et al., 2010).

2.2.3 Fanconi-Bickel Syndrome

Fanconi-Bickel syndrome (FBS) (MIM #227810) is a rare disease of autosomal recessive inheritance mode. In 1997, this disease was found to result from a defective monosaccharides transporter, GLUT2 protein (Santer et al., 1998) which will interfere with the exchange of glucose between several tissues such as the hepatocytes, pancreatic β cells, renal and intestinal epithelial cells and the bloodstream (Burwinkel et al., 1999). Contrasting to other GSDs, no enzyme defect was involved hence the use of term glycogen storage disease type XI (former name) was discouraged (Ekbote et al., 2012).

2.2.3.1 Clinical Symptoms of FBS

The first case of FBS was reported by Fanconi and Bickel in 1949 which described a patient named Claudio M. with failure to thrive, polydipsia and constipation at young age of 6 months. Subsequent follow up at infancy found that the patient had short stature, generalized osteopenia, protuberant abdomen, hyperlordosis, excessive hepatomegaly and facial obesity. Liver histology on the patient sample showed steatosis and excessive amounts of glycogen. Fasting hypoglycaemia and ketonuria as well as increased sensitivity to insulin were observed (Santer et al., 1997). These symptoms were later observed in other patients of FBS with several new symptoms were also presented. Similar clinical and chemical findings were observed by Rotthauwe in 1963 (Rotthauwe et al., 1963) with a new symptom of galactose intolerance and Odievre in 1966 reported a similar observation. Odievre later confirmed that glucose and galactose utilization were impaired by in vitro experiments. Those experiments also demonstrated that the monosaccharide utilization impairment was not generalized as erythrocytes were not affected (Odievre et al., 1966). This leads to the suggestion of the involvement of monosaccharide transporter across liver cells. A further supporting study of the

transporter involvement came about when a study by Fellers found that there were equal renal clearance of glucose and insulin in their patient (Fellers et al., 1967).

In general FBS is characterized by hepatic glycogen accumulation and renal Fanconi syndrome. Glycogen accumulation occurs due to intracellular trapping of glucose leads to the stimulation of glycogen synthesis and/or inhibition of glycogen degradation (Burwinkel et al., 1999). This leads to hyperglycaemia in the fed state or hypoglycaemia during fasting. Further consequences involve the renal glycogen accumulation. Proximal renal tubular dysfunction with glucosuria up to 325 g/day per $1.73m^2$ (Santer et al., 1998), mild proteinuria, phosphaturia, generalized aminoaciduria, bicarbonate wasting and hypophosphatemia may well be observed (Şimşek et al., 2009). Other symptoms that can appear earlier in life at age 3-10 months may include fever, vomiting, failure to thrive and hypophosphataemic rickets. Further clinical findings may include abdominal distension, X-bain deformity and hepatomegaly. A 'doll-like face' has also been associated with FBS patients (Dweikat et al., 2016). Symptoms that may appear much later in life such as rickets and osteoporosis that eventually lead to pathological fractures had been reported in few cases (Santer et al., 1998).

The size and growth of the kidney were found to be increased in relation to body height in most patients. Glomerular filtration rate is usually normal but a study reports that their patient had glomerular hyperfiltration, microalbuminuria and diffuse glomerular mesangial expansion resembling early diabetes and deficiency of glucose-6-phosphatase (Berry et al., 1995).

2.2.3.2 Diagnosis and Treatment of FBS

Definitive diagnosis of FBS is made upon identification of mutation in the *SLC2A2* gene by molecular diagnosis following presentation of the common clinical features. Diagnosis of FBS without the combination of hepatomegaly secondary to

glycogen storage and generalized tubulopathy is considered preliminary (Santer et al., 1998). Other supporting tests look for ketonuria in the fasting state as well as hyperglycaemia in the post absorptive state, liver histology to determine steatosis and storage of excessive amounts of glycogen, renal function tests to determine glucosuria, hyperphosphaturia, hyperuricaemia, hyperaminoaciduria and proteinuria.

Once diagnosis was made, symptomatic treatment includes replacement of water and electrolytes, alkalinisation with Shohl or bicarbonate solutions, supplementation of vitamin D and phosphate restriction of galactose (Santer et al., 1998). These treatment however have little impact on patient's growth so uncooked corn starch administered as slow-release glucose preparation (Al-Haggar et al., 2011). Transport of this monosaccharide by a specific transport protein, GLUT5 that is not affected in FBS allow its use as an alternative carbohydrate source (Blakemore et al., 1995).

2.2.3.3 Molecular Genetics of FBS

The first researcher to postulate a defect of the diffusion carrier facilitating the transport of glucose and galactose across liver was Manz et al. in 1987. He was also responsible in introducing the eponym Fanconi-Bickel syndrome (Santer et al., 1998). It was only in 1995 that Berry et al. speculated a defective GLUT2 as the possible explanation for FBS pathophysiology (Berry et al., 1995) and finally in 1997, Santer described mutations in the *SLC2A2* gene, the gene that is responsible for liver facilitative glucose transporter. He reached to this conclusion when homozygous mutations of the *SLC2A2* gene were found in the affected individuals and because all detected mutations were predicted be truncated, consequently no functional transport activity (Santer et al., 1997). The first evidence that support this conclusion is that the defect at codon 365 in one of the mutation reported (p.R365X) is within a highly conserved intracellular (R)XGRR motif, common to different facilitative glucose

transporters as well as to sugar transport superfamily (Ekbote et al., 2012). Second evidence highlight the importance of distal domain of glucose transport proteins as essential for monosaccharide transport which in truncating mutations would be lost (Oka et al., 1990). Meanwhile Thorens et al. (1996) detected carboxyl-terminal phosphorylation sites for a cAMP-dependent protein kinase, further enlighten the role of the carboxyl-terminal domain of the GLUT2 protein.

In summary, FBS occur due to defective glucose transporter 2 (GLUT2) protein unlike most other types of glycogen storage disease which is caused by enzymatic defect. Defects in the monosaccharide transporter is now known to be caused by more than 60 different mutations involving the whole coding sequence of *SLC2A2* gene as had been reported in Human Genome Mutation Database (HGMD, http://www.hgmd.cf.ac.uk/ac/all.php). The type of mutation varies with no apparent mutational hot spots, hence limiting efficient genetic testing. In one study conducted by Santer and colleagues, they reported 23 different novel mutations among 49 patients diagnosed with FBS (Santer et al., 2002). A similar observation was seen in a study with smaller cohort by Sakamoto where they reported four different novel mutations from three Japanese FBS patients. The mutations include a splice site mutation, a nonsense mutation and two missense mutations (Sakamoto et al., 2000). Furthermore, the mutations of *SLC2A2* were also detected in various ethnic groups with no specific mutations being prevalent. Most mutations are private as it is found to be diagnosed in single families (Santer et al., 1998).

2.2.3.4 The *SLC2A2* Gene and Glucose Transporter Protein

The human *SLC2A2* gene also known as *GLUT2* gene (MIM #138160) is a glucose-sensitive gene in the liver cells which is part of the Solute Carrier Family 2 (SLC2A) gene family that codes for a group of transporter proteins. These transporters

have the ability to process high sugar concentrations efficiently due to its high Vmax and Km for glucose (Gould et al., 1991). This gene is also expressed in the islet beta cells, intestine and kidney epithelium (Permutt et al., 1989). The successful cloning of this gene was first described in 1988, using transcripts that were isolated from adult human liver and kidney libraries. They then localized the gene to chromosome 3q26.1-q26.3 by somatic cell hybridization and in situ hybridization (Fukumoto et al., 1988) (Figure 2.7).

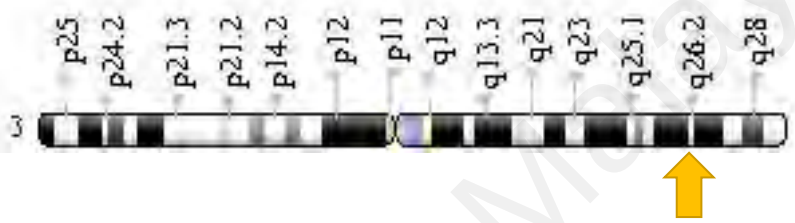


Figure 2.7: Schematic diagram of *SLC2A2* gene on chromosome 3q26. (Source: http://www.ncbi.nlm.nih.gov/genome/tools/gdp)

It was only in 1993 that Takeda reported the organization of the *SLC2A2* gene and that it consists of 11 exons, 10 introns and spans approximately 30 kb in length (Takeda et al., 1993). There exist different transcript variants of the gene with variant one representing the longest transcript (3439 bp mRNA) and codes for the longest isoform, isoform 1. The variants vary in the 5' UTR with the shorter variant lacking an alternate exon in the 5' coding region and translation is initiated at an alternate start codon (NCBI, http://www.ncbi.nlm.nih.gov).

The GLUT2 protein is a facilitative glucose transporter in the liver, kidney (Fukumoto et al., 1988), pancreas (Orci et al., 1989; Permutt et al., 1989) and small intestine (Brown, 2000). This protein belongs to the Class I transporter, one of the three classes of transporters that is grouped based on sequence similarity and to a lesser extend substrate specificity (Zhao & Keating, 2007). Class I transporters transport glucose as well as galactose. GLUT2 is the only low-affinity Class 1 glucose transporter

(Lachaal et al., 2000) but with the highest Km (~40mmol/L) hence making the protein a very efficient bidirectional transporter of glucose (Brown, 2000). This is particularly an important role in the liver as it allows for glucose uptake after a meal and glucose release when blood glucose is low (Santer et al., 1998). In the kidney, GLUT2 is involved in glucose re-absorption from urine into the blood whereas in the pancreatic β-cell GLUT2 is part of the glucose sensing mechanism where it coupled with glucokinase to stimulate glucose-induced insulin secretion (Mueckler et al., 1994). Besides that, this transporter is also involved in feeding response where GLUT2 is responsible for the stimulation of glucagon secretion or suppression during hypoglycaemic and hyperglycaemic conditions respectively (Burcelin & Thorens, 2001). GLUT2 is however a high-affinity transporter for glucosamine and it also transports fructose (Lachaal et al., 2000).

GLUT2 protein exists in several isoforms. Isoform one is the longest with 3439 bp mRNA, molecular weight of 57.5 kDa and contains 524 amino acids (Gupta et al., 2016). Like all members of the GLUT family, GLUT2 have 12 transmembrane α-helices segments with both the N- and C-termini are cytoplasmic and a single glycosylation site on one of the extracellular loops (Mueckler et al., 1985) and transport its substrate according to an alternating conformation model (Figure 2.8). The model suggests that the transport protein have alternative conformation with the substrate binding site directed either in the membrane or out of the membrane (Santer et al., 1998).

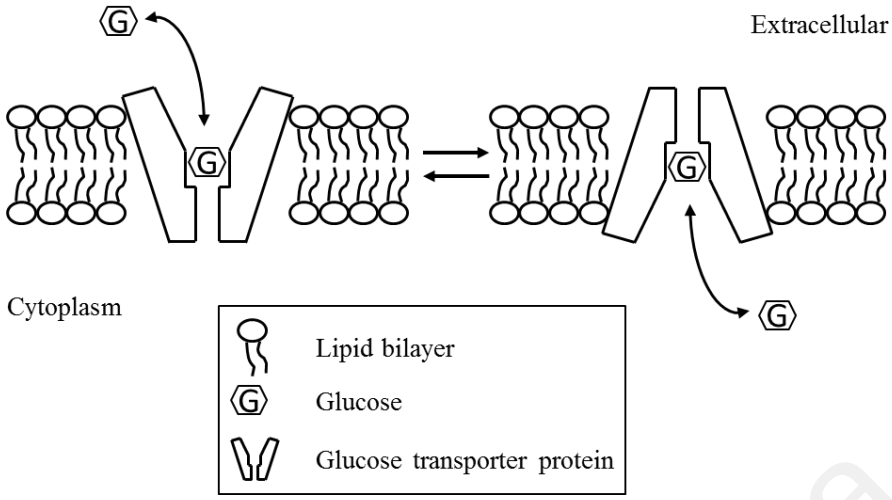


Figure 2.8: Diagram of alternating conformation model of glucose transport.

2.3 Lipids and Protein as Other Source of Energy

Lipids are hydrophobic molecules that carry various functions in the human body. They are important insulators in the subcutaneous layer to maintain body temperature and act as electrical insulators to the nerve fibres. They also act as signaling molecules, cellular metabolic regulators as well as important structural components of the membranes. While the major readily available source of energy in human is provided by carbohydrates, lipids primarily function as an energy reserve. Lipids are stored mostly in our fat cells (adipose cells) as well as other cells in the form of triglycerides.

Triglycerides are compound with glycerol backbone bonded to three fatty acids (Figure 2.9). The three fatty acids can be similar (forming simple triglycerides) or different types (forming mixed triglycerides).

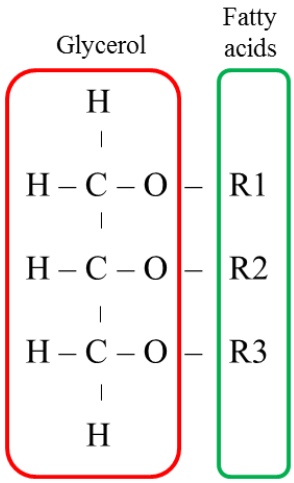


Figure 2.9: General structure of a triglyceride. Glycerol is linked to three fatty acids. H: hydrogen, C: carbon, O: oxygen, R1-R3: fatty acids.

During lipid metabolism, lipases cleave the triglycerides into glycerol and fatty acids. Glycerol is converted into dihydroxyacetone phosphate which is an intermediate in glycolysis. The fatty acids and other hydrocarbons are catabolized by beta-oxidation. The majority of lipids contain an even number of carbon atoms. The catabolism product of these even-numbered lipids is acetyl-CoA which will enter the TCA cycle. Some lipids are however containing odd numbers of carbon atoms and these yield acetyl-CoA units and propionyl-CoA. The propionyl-CoA will then be converted to succinyl-CoA by propionyl-CoA carboxylase and methylmalonyl-CoA mutase before it enters the TCA cycle for further oxidation to yield energy (Fowler et al., 2008).

The other source of energy comes from proteins, which is essentially a polymeric chain of amino acids linked together by peptide bonds. There are 20 different types of amino acids with the general structure comprising an α-carbon to which an amine group ($-NH_2$), a carboxyl group (-COOH) and a variable side chain (R) are bonded (Figure 2.10). The key elements are carbon, hydrogen, oxygen and nitrogen but the R chain may carry other elements.

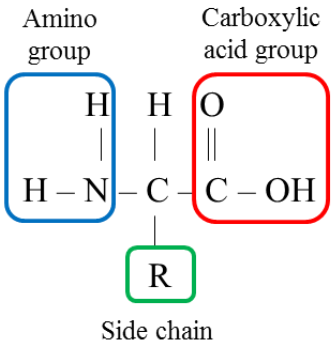


Figure 2.10: General structure of amino acids. H: hydrogen, N: nitrogen, C: carbon, O: oxygen, OH: hydroxide, R: side chain.

While some amino acids can be synthesized by the body, the others need to be obtained through diet. The amino acids that can be synthesized in the body are alanine, aspartic acid, asparagine, glutamic acid and serine. Meanwhile, the amino acids that must be obtained from diets are phenylalanine, valine, threonine, tryptophan, methionine, leucine, isoleucine, lysine and histidine. Other amino acids of which their synthesis can be limited under certain pathophysiological conditions are arginine, cysteine, glycine, glutamine, proline and tyrosine (Fenton et al., 2001).

Protein is important for body growth and maintenance. It is a major structural component of the cells such as the muscle, hair and skin. Other than that, proteins also function as glycoproteins that are important for cell signaling and cell-cell recognition. Furthermore, protein has also been shown to be involved in blood cells formation. The broken down product of protein, the amino acids are used as precursors to nucleic acid, hormones, immune response and other molecules essential for body function (Berg et al., 2002)

Protein catabolism starts in the intestines where proteins are digested by various peptidases to produce amino acids. In terms of catabolism, amino acids can be divided into ketogenic and glucogenic amino acids. Ketogenic amino acid (e.g: acetyl-CoA or acetoacetate) is an amino acid that can be degraded directly into acetyl-CoA which is the precursor of ketone bodies. Meanwhile glucogenic amino acid (e.g: succinyl-CoA,

fumarate, oxaloacetate, α-ketoglutarate) is an amino acid that is converted to alpha keto acids. This alpha keto acid is then converted into glucose by the action of enzymes propionyl-CoA carboxylase and methylmalonyl-CoA mutase enzyme during gluconeogenesis. Defect in either enzyme gives rise to propionyl-CoA carboxylase deficiency or methylmalonyl-CoA mutase deficiency (known as methylmalonic aciduria).

2.4 Methylmalonic Aciduria

Methylmalonic aciduria or methylmalonic acidemia (MMA) is an autosomal recessive disorder that is caused by deficient activity of the enzyme L-methylmalonyl-CoA mutase (MCM, EC 5.4.99.2) (Hörster & Hoffmann, 2004). The MCM enzyme catalyzes the isomerization of L-methylmalonyl-CoA to succinyl-CoA in the propionate catabolism, a process by which the metabolites from amino acids (valine, isoleucine, methionine, and threonine), odd fatty acid chains and cholesterol breakdown are feed into the tricarboxylic acid cycle (Fenton et al., 2001). Functional deficiency of MCM is caused by either defect in the MCM apoenzyme itself (designated as MMA mut-type, MIM# 251000) or defects in the enzymes required for providing the adenosylcobalamin (AdoCbl), a cofactor for MCM (designated as MMA cbl-type, MMA cblA-type; MIM#251100 and MMA cblB-type; MIM#251110) (Ledley, 1990; Willard et al., 1978). These diseases are called isolated methylmalonic acidurias. However there are other forms of MMA that occur with homocystinuria, namely the MMA cblC-type (MIM#277400), MMA cblD-type (MIM#277410) and MMA cblF-type (MIM#277380) (Willard et al., 1978). Collectively, the estimated incidence of methylmalonic aciduria worldwide ranges between 1:48,000 and 1:250,000 (Wang et al., 2010). MMA mut-type and MMA with homocystinuria, cblC-type will be discussed in this study.

2.4.1 MMA Caused by Defective Methylmalonyl-CoA Mutase (MCM)

MMA mut-type (MIM #251000) that arises from defective MCM is caused by mutations in the *MUT* gene (Jansen & Ledley, 1990). The MCM deficiency can be further subdivided into two categories which are mut^0 and mut^-. When fibroblasts exhibit no [^{14}C] propionate incorporation under any conditions or no detectable apoenzyme activity in vitro, it is termed as mut^0. Meanwhile when fibroblasts exhibit [^{14}C] propionate incorporation when stimulated by high concentration of hydroxycobalamin, it is termed as mut^- . The difference in clinical expression of MMA is attributed to this variation in biochemical activity (Imtiaz et al., 2015).

2.4.1.1 Clinical Symptoms of MMA Mut-Type

Onset of manifestation of MMA ranges from neonatal period to adulthood. Patients suffering from methylmalonic aciduria usually presented with episodes of illness called metabolic crises. This include poor appetite, recurring vomiting, dehydration, lethargy, hepatomegaly, respiratory distress, muscular hypotonia and in severe cases with overlapping illnesses such as progressive renal failure, metabolic stroke or cardiomyopathy, deep coma and death may occur (Fenton et al., 2001; Hörster & Hoffmann, 2004). Other common biochemical feature that may be observed is severe ketoacidosis, hyperammonemia as well as large amounts of methylmalonic acid in the urine (Lempp et al., 2007).

2.4.1.2 Diagnosis and Treatment of MMA Mut-Type

The diagnosis of MMA mut-type is confirmed in fibroblasts by demonstrating an undetectable MCM activity (mut^0) or low MCM activity (mut^-) which increases in response to hydroxycobalamin (Berger et al., 2001). Tandem mass spectrometry analysis shows high elevation of propionylcarnitine that is indicative of either MMA or propionic academia however an urine organic analysis of abnormally high excretion of

methylmalonic acids is definitive of MMA (Keyfi et al., 2016). Other confirmatory method to diagnose MMA is by direct sequencing which may reveal mutation in the *MUT* gene (Ghoraba et al., 2015).

Patients of MMA usually receive treatments that may include protein restriction diet, daily intramuscular hydroxycobalamin injections, L-carnitine and metronidazole syrup, correction of metabolic acidosis, infection and electrolyte imbalance (Ghoraba et al., 2015). Oral administration of vitamin B12 that may reduce urinary excretion of methylmalonic acid may be employed (Sakamoto et al., 2000).

2.4.1.3 Molecular Genetics of MMA Mut-Type

Various studies have identified different mutations in the human *MUT* gene. Most mutations found to be either missense or nonsense mutations (HGMD, http://www.hgmd.cf.ac.uk/ac/). Despite mutations of the *MUT* gene was found to be highly heterogenous, some mutations however have been reported to be associated with certain populations such as the c.655A>T in black patients (Adjalla et al., 1998) and the c.655A>T in Caucasian (Acquaviva et al., 2001). A particular mutation, the c.1181C>T had been found in three nations which are Korea, Japan and America (Jung et al., 2005). While few mutations had been associated with certain population, other study had found that most disease-causing mutation of the *MUT* gene occurs in the exon 2 (Acquaviva et al., 2001; Ghoraba et al., 2015; Jung et al., 2005; Mikami et al., 1999). The study conducted by Ghoraba alone found 26 different allelic variants, occurring within the intronic region as well as the coding region of exon 2 (Ghoraba et al., 2015). Furthermore, some mutations found within exon 2 were also found to be common in certain ethnicities such as the mutation c.322C>T (p.R108C) in Hispanics (Worgan et al., 2006) and p.E117X in Japanese (Ghoraba et al., 2015).

2.4.1.4 The *MUT* Gene and Methylmalonyl-CoA Mutase (MCM)

The *MUT* gene (MIM# 609058) contains 13 exons, including the untranslated regions and spanning over 35 kb of the genome. In situ hybridization, analysis of somatic cell hybrids (Ledley et al., 1988), and genetic linkage had localized the gene to chromosome 6p21 (Ledley et al., 1990) (Figure 2.11) . cDNA clones for this gene were first described in 1988 which contains 2.7 kb mRNA and transcribed in the nucleus (Fuchshuber et al., 2000).

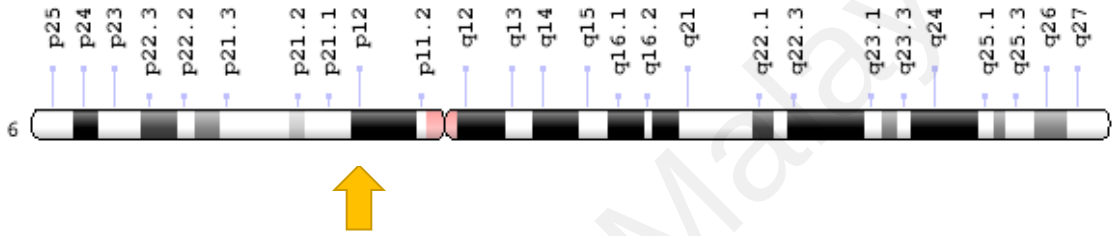


Figure 2.11: Schematic diagram of *MUT* gene on chromosome 6p12.21.1. (Source: http://www.ncbi.nlm.nih.gov/genome/tools/gdp)

The *MUT* gene codes for mitochondrial enzyme methylmalonyl-CoA mutase that is important for the propionate metabolism (Ledley et al., 1990). The enzyme converts methylmalonyl-CoA into succinyl-CoA (Figure 2.12) which will later be used in the Krebs cycle (Jansen & Ledley, 1990).

In the process, amino acid valine, isoleucine, methionine and threonine as well as odd-fatty acids and cholesterol are degraded into propionyl-CoA, a precursor for the propionate catabolism reaction. There are three enzymatic reactions involved in the conversion of propionyl-CoA to succinyl-CoA. The enzyme propionyl-CoA carboxylase converts propionyl-CoA into D-methylmalonyl-CoA which is then racemized into L-methylmalonyl-CoA by enzyme methylmalonyl-CoA mutase. The third reaction, which converts the L-methylmalonyl-CoA to succinyl-CoA is carried out by the enzyme methylmalonyl-CoA mutase (Fenton et al., 2001). This enzyme requires adenosylcobalamin (AdoCbl), an activated form of vitamin B_{12} as a cofactor (Hörster &

Hoffmann, 2004). Defective MCM that led to MMA were unable to perform the third reaction leading to excretion of not metabolized methylmalonic acid in the urine and other organic acids in the blood and tissues (Ledley et al., 1990).

The MCM is an 85 kDa protein composed of 742 amino acids with an initial mitochondrial leader sequence of 32 amino acids that is truncated after entering the mitochondria (Jansen & Ledley, 1990). The leader sequence that is strongly positively charged (4 Arg, 2 Lys, 1 Glu) directs the precursor to the mitochondria where it will be recognized, translocated across the membrane and finally cleaved by matrix endoprotease to form the mature subunit (Ledley et al., 1990). The mature subunit is composed of the remaining amino acid chain which will assemble into a homodimer that binds with 2 AdoCbl molecules to form the activated enzyme form (Thoma & Leadlay, 1996). The cloning and sequence analysis of both human MCM (Ledley et al., 1988) and the MCM from *P. shermanii* (Marsh et al., 1989) revealed 65 % identity between the mature human enzyme and the α subunit of the *P. shermanii* that also had been shown to bind AdoCbl and substrate, despite having different quaternary structure (Thoma & Leadlay, 1996).

The three dimensional model of the human MCM structure has also been deduced from the crystal structure of *P. shermanii* MCM (Mancia et al., 1996). The model proposed that each subunit of MCM has two major domains (Figure 2.13) which consist of an N-terminal eight-stranded β/α_8 barrel (residues 88-422) that is involved in the substrate binding and a C-terminal cobalamin-binding β/α_5 domain (residues 578-750). Both domains are linked by a linker region that stretches across residues 423-577. Besides that, the model suggest leader sequence to occupy residues 1-32 followed by N-terminal extended segment that is involved in the subunit interactions to occupy residues 33-87 (Thoma & Leadlay, 1996).

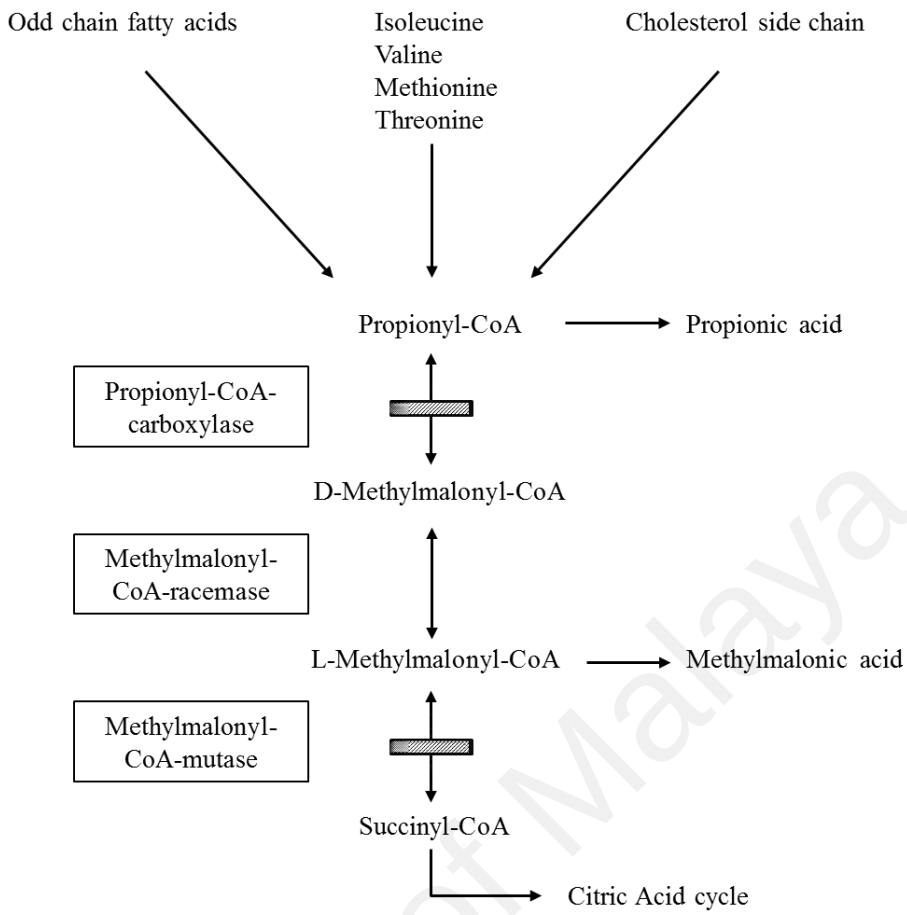


Figure 2.12: Diagram of propionate metabolism.

2.4.2 Methylmalonic Aciduria with Homocystinuria, CblC Type

MMA cblC-type (MIM #277400) is one form of combined MMA and homocystinuria. This disease is caused by impaired biosynthesis of methylcobalamin (MeCbl) and adenosylcobalamin (AdoCbl) (Rosenblatt et al., 1997) that results from mutation in the *MMACHC* gene (Lerner-Ellis et al., 2006). MeCbl and AdoCbl are essential for the activities of both methionine synthase enzyme (MIM#156570) and methylmalonyl-CoA mutase (discussed in this study) where decreased level of the coenzymes subsequently reduces the activity of the aforementioned enzymes. This consequently leads to increased concentrations of methylmalonic acid and homocysteine in plasma and urine, while the concentration of methionine in the plasma decreases (Wang et al., 2010). Among the three distinct complementary groups of

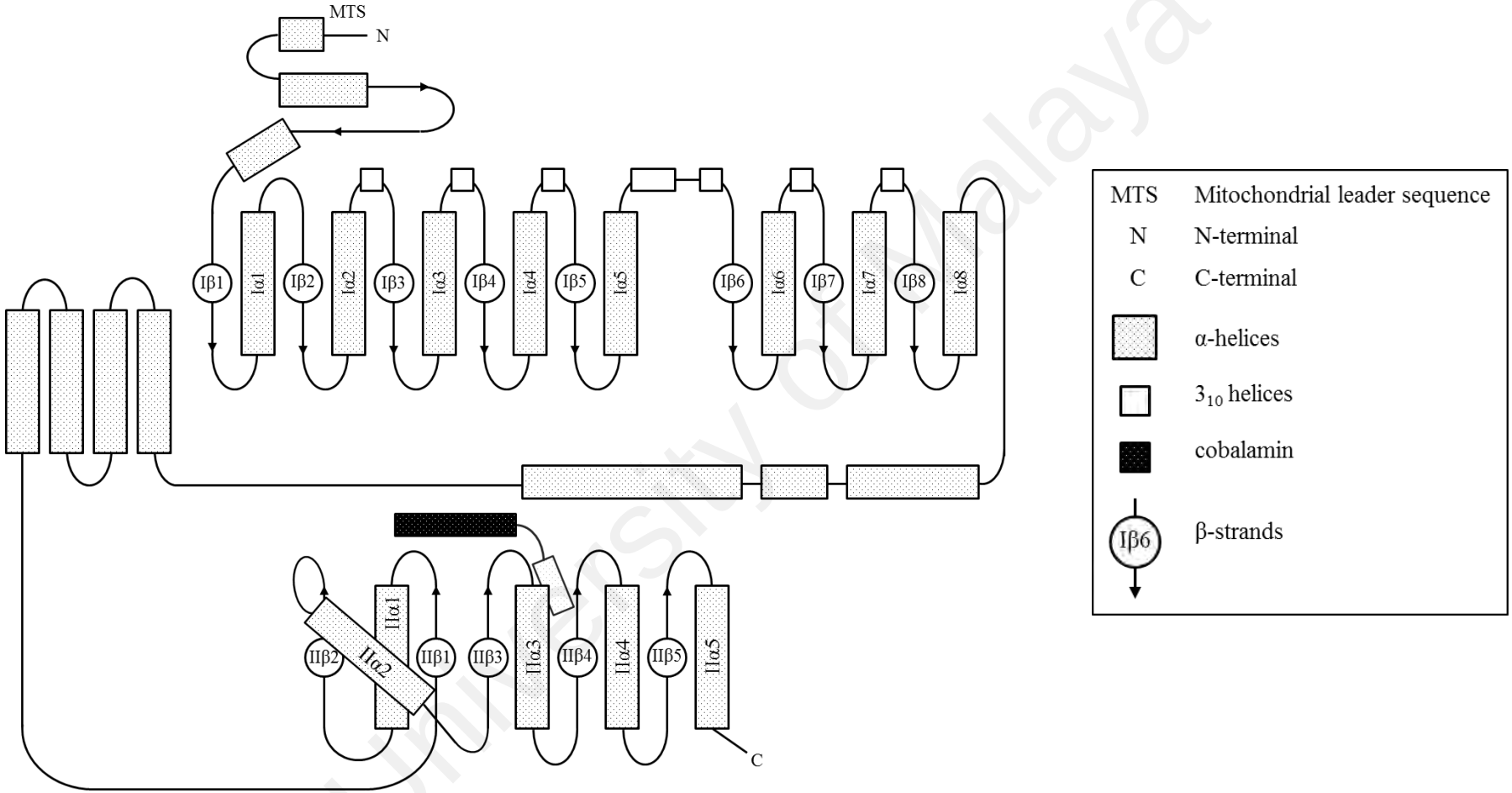


Figure 2.13: Topology diagram of human methylmalonyl-CoA mutase. (Adapted from Thoma & Leadlay, 1996).

combined methylmalonic academia and homocystinuria (described in section 2.2), cblC type appear to be the most prevalent form (Morel et al., 2006) with reported incidence of approximately 1:100,000 (Weisfeld-Adams et al., 2010).

2.4.2.1 Clinical Symptoms of MMA CblC-Type

Patients suffering from methylmalonyl aciduria and homocystinuria are presented with neurological, developmental, haematological and ophthalmologic problems. Patients can also be broadly divided into early-onset (onset in the first year of life) and late-onset groups (symptoms appear at age four years old) (Carrillo-Carrasco et al., 2012). Early-onset patients have a more severe clinical presentation and less favourable outcome with one quarter of patients do not survive (Rosenblatt et al., 1997). They are presented with feeding difficulties, hypotonia, developmental delay, seizures, anaemia and ocular diseases (Carrillo-Carrasco et al., 2012; Fowler et al., 2008). The ocular disease involvements in early-onset patients are significantly more common than the late-onset (Gerth et al., 2008; Weisfeld-Adams et al., 2013). Pigmentary retinopathy, macular disease, optic atrophy, strabismus as well as refractive error appear to be the common findings (Gizicki et al., 2014; Huemer et al., 2014). Young death as well as mental retardation among early-onset patients are also common findings (Huemer et al., 2014).

The late-onset patients have less severe symptoms and are associated with better response to treatment. They are presented with acute onset of neurological symptoms (Ben-Omran et al., 2007) such as ataxia, dementia, psychosis, confusion, cognitive decline and other neurologic symptoms after 4 years old (Carrillo-Carrasco et al., 2012; Fowler et al., 2008). Unsteady gait, impaired speech has also been reported (Rosenblatt et al., 1997).

There are reports of MMA cblC patients presented with minor facial abnormalities. Several patients had been found to have high forehead, long face, large low-set ears and flat intranasal depression (Cerone et al., 1997). One late-onset patient was found with features such as triangular face, large prominent nose, micrognathia and down-slanting eyes (Ben-Omran et al., 2007). However, due to different features observed the authors concluded that facial dysmorphism may not be a consistent feature of cblC disease, particularly in the case of late onset form.

The MMA cblC-type disease symptoms can progress into severe complications despite treatment (Wang et al., 2015). There are three factors that may contribute to the complications; increased homocysteine concentrations, impaired methyl group metabolism or oxidative stress (Rosenblatt et al., 1997).

2.4.2.2 Diagnosis and Treatment of MMA CblC-Type

Early diagnosis and treatment have greatly reduced the mortality and morbidity in children with MMA. Prenatal diagnosis that include measuring the activity of methylmalonyl-CoA mutase in the amniotic fluid, amniotic fluid cells and chorionic cells have been successful. Quantification of cbl metabolitis in amniotic fluid cells may also be carried out. Current development of acylcarnitine analysis by electroscopy tandem mass spectrometry (ESI/MS/MS) had also been included in prenatal diagnosis (Zong et al., 2015). In more recent years, with the availability of genomic information regarding the *MMACHC* gene, combination of biochemical analysis and sequencing has been the method of choice for diagnosing MMA cblC-type.

Treatment of cblC patients is focusing to improve the biochemical parameters. Both hydroxocobalamin (OHCbl) and oral betaine have been established as the treatment. Betaine has been shown to be effective at reducing the homocysteine levels and increasing methionine levels. However, there are various levels of responsiveness

towards OHCbl treatment and this may be dependent on the mutations present (Froes et al., 2009). Other than that, dietary restriction and daily intake of folinic acid and carnitine are used to control the symptoms (Weisfeld-Adams et al., 2010). However, there is a different response towards medication noted between early-onset and late-onset patients. The early-onset patients respond poorly while in contrast, the late-onset patients respond with significant improvement towards treatment for both the biochemical and clinical Manifestation (Baumgartner Et al., 2014; Huemer Et al., 2014).

2.4.2.3 Molecular Genetics of MMA CblC-Type

Since its discovery, more than 80 different mutations of the *MMACHC* gene have been identified (http://www.hgmd.cf.ac.uk). Several mutations appear to cluster by ethnicity. Mutation c.609G>A for example is frequently found in the Chinese population (Wang et al., 2009), while the c.328_331delAACC mutation is frequently found in Hispanics. In Italians however, several mutations were noted to be frequently encountered such as the c.457C>T, c.468_469delCT and the c.666C>A (Lerner-Ellis 2009).

Genotype-phenotype correlations have been studied and there appear to be mutations that is associated with the early-onset or the late-onset form of the disease (Wang et al., 2010). The mutation c.271dupA (p.R91KfsX14) for example has been associated with early-onset (Lerner-Ellis et al., 2006) while nonsense mutation c.394C>T (p.R132X) has been associated with late-onset form (Morel et al., 2006; Nogueira et al., 2008).

2.4.2.4 The *MMACHC* Gene and MMACHC Protein

Mapping of the gene showed that it is located at chromosome region 1p34.1 (Figure 2.14), consists of 5 exons and spans a region of approximately 5.2 kb of mRNA

(Lerner-Ellis et al., 2006). Exons 1-4 are coding exons while exon 5 is noncoding (Wang et al., 2010).

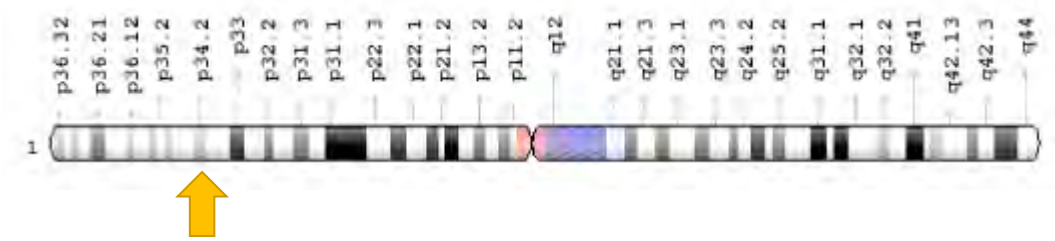


Figure 2.14: Schematic diagram of *MMACHC* gene on chromosome 1p34.1. (Source: http://www.ncbi.nlm.nih.gov/genome/tools/gdp)

The MMACHC protein comprises 282 amino acids with a molecular weight of 31.7 kDa. Despite the protein found to be conserved only within mammals, motifs homologues to bacterial genes with cobalamin-related functions were identified. The residues 118-138 share 52 % amino acid similarity to cobalamin-binding region residues of *S.avermitilis*. Furthermore, residues 181-282 at the C-terminus showing similarity with TonB, a protein derived from Gram negative bacterium involves in energy-coupled transport of vitamin B_{12}. This is supported by three-dimensional modelling where despite low sequence identity (14 %) between them, the protein models superimposed closely (Lerner-Ellis et al., 2006). Meanwhile, the residue 122-156 also contains the motif (122-HXXGX$_{126\text{-}154}$GG-156), a cobalamin binding site found in methionine synthase and methylmalonyl-CoA mutase enzymes (Froese et al., 2009).

The protein involves in converting vitamin B12 (also called cobalamin) into either the adenosylcobalamin (AdoCbl) or methylcobalamin (MeCbl). Specifically, the protein released the R group from the cobalamin (cbl-R) and formed the cobalamin (II) product. This cobalamin (II) is a substrate used for assimilation into the active cofactor forms MeCbl and AdoCbl that is needed by methionine synthase and methylmalobyl-CoA respectively (Kim et al., 2008). Besides that, the protein has also been found to be

involved in the decyanation of cyanocobalamin (CNCbl) to yield cobalamin (II) and cyanide. The protein utilizes electron reduction using NADPH (Kim et al., 2009). Another reported function of MMACHC protein is to catalyse the dealkylation of newly synthesized MeCbl and AdoCbl (Hannibal et al., 2009).

2.5 Mutation Screening of Inherited Diseases

The mutation screening on human inherited diseases allows for identification of the mutation that occurs in a particular gene. This knowledge is important to allow for better understanding on the molecular events that leads to disease pathogenesis. Effective medication or treatment can then be developed, and genetic counselling can be offered to expecting parents or parents with children suffering from inherited disease.

There are various techniques that can be adopted to identify disease-causing mutations. Methods such as the Single strand conformational polymorphism (SSCP), Denaturing gradient gel electrophoresis (DGGE) and Heteroduplex analysis were previously widely used (Mahdieh & Rabbani, 2013). In recent years, advances in molecular technique allow a more rapid and efficient analyses. This includes direct sequencing, TaqMan assay and next-generation sequencing. While next-generation sequencing allows large amounts of DNA that includes the whole exome and genome to be sequenced, direct sequencing is still the benchmark technique to be used for mutation screening (Ferrari et al., 2008). In studies involving diseases with known causative gene, direct sequencing provides as a cheaper option for mutation screening.

Focusing on genetic studies conducted on IMDs, literatures reported that there is high level of genetic heterogeneity involved in each type of the disease (Crushell et al., 2010; Gizicki et al., 2014; Ledley & Rosenblatt, 1997). Some of the mutations also show high prevalence in specific ethnic groups (Ben-Omran et al., 2007; Choi et al., 2017; Mili et al., 2011). Many studies also show that some of these mutations are

recurrent mutation (Han et al., 2015; Parvari et al., 1997). For example, more than 89 mutations of the glucose-6-phosphatase gene (GSD I) have been identified with p.R83C as common in Jewish population (Parvari et al., 1997) and also found to be recurrent in Caucasian population(Carlin et al., 2013) . Thus, mutation pattern for each type of IMD in a specific population provide useful information to allow further understanding of the disease progression and consequently future development of diagnostic method and treatment.

2.5.1 Population Screening of Mutated Allele

In cases where unreported point mutations were identified, excluding the possibility of it being a single nucleotide polymorphism and not disease-causing mutation is of immediate interest. Single nucleotide polymorphisms (SNPs) are common genetic variations which occur on average one per 1000 bases (Wang & Moult, 2001). While most SNPs are not disease causing and may instead act as biological markers, others can have direct role in disease by disrupting gene's function (Wang & Moult, 2001). A mutation is considered as being a polymorphism if the rate of its occurrence is $\geq$ 1 % in a given population (Brookes, 1999). Therefore, to exclude an unreported mutation as being a mere polymorphism, screening of the population for the mutated allele needs to be carried out. While DNA-based sequencing method is available, alternative methods are usually employed due to large number of samples needed to be tested. The alternative methods are listed below.

i) Restriction enzyme (RE) digestion assay

The principle of this assay is similar to RFLP where a specific restriction enzyme is employed to cut a specific DNA sequence termed as restriction site. A mutation might alter this restriction site sequence preventing the RE to cleave the DNA. Therefore, an enzyme can be used to differentiate sample carrying the mutant and

normal allele. Region of interest will first be amplified by PCR and then subjected to incubation with the RE. The cleaved DNA fragment then can be analysed by agarose gel electrophoresis (Jenkins et al., 1999).

ii) TaqMan® SNP Genotyping Assay

This method uses a labelled probe which is hybridized to the template between the primers during a PCR reaction. The probe carries a fluorescent reporter molecule at the 5' end and a quencher molecule at the 3' end. Probes designed are complementary to the alleles of interest and labelled with different reporter molecules. Upon cleavage by Taq polymerase, the reporter dye will fluoresce as it is no longer quenched and the intensity of the emitted light is measured by fluorometer (Malkki & Petersdorf, 2012).

2.5.2 Functional Analysis of Mutated Gene

In order to investigate the pathogenicity level of unreported mutations, functional analyses are carried out to determine their effect on normal protein function. There are two approaches that can be adopted to carry out functional analysis, as listed below.

i) Bioinformatics

Functional analysis carried out with this approach is predictive that uses known data, statistical algorithms and machine learning techniques to identify the likelihood an event to occur. There are various software that can be used to predict the outcome of mutations and these include PolyPhen-2, SIFT, and MutationTaster. Other than that, modelling of the mutant protein to predict its structural changes can also be carried out. Protein prediction software construct the target protein model from its amino acid sequence and uses information on three-dimensional structure of related protein(s) (Frousios et al., 2013).

ii) Recombinant gene expression

Another approach to functional analysis is by doing recombinant gene expression of the mutant gene. Firstly, gene of interest is cloned into the expression vector and then transfected into the cell line of choice. Gene expression product is then tested for functional analysis (Brown, 2016).

CHAPTER 3: MATERIALS AND METHODS

This chapter describes the study cohort and the methods used throughout this work. In general, this study can be divided into three parts which are mutation screening, population screening for potentially novel mutations and functional studies of the effect of mutation on normal gene/protein function.

Section 3.1 outlines the methods used to detect mutations among the patients recruited in this study. Section 3.2 described the methods used to screen mutant *AGL* alleles in the population while section 3.3 described the software used to predict the effect of unreported mutations. Experimental work to determine the effect of mutation on mRNA splicing and in vitro expression work were described in section 3.4 and section 3.5 respectively. Lastly, section 3.6 described the methods used to analyse mutant proteins.

3.1 Mutation Screening

Initially, samples were screened for mutations within genes that is associated with the disease diagnosis. These inherited metabolic disorders include glycogen storage disease type 1b; GSD 1b, glycogen storage disease type III; GSD III, Fanconi-Bickel syndrome; FBS, methylmalonic aciduria; MMA mut-type and methylmalonic aciduria with homocystinuria cblC type; MMA cblC-type . Peripheral bood samples from patients were used as source of DNA and were screened for mutations. The mutation screening was carried out across the exons and immediate intronic regions of 100 bp. Parents and/or siblings were also tested when samples were made available. The following subsections described the experimentation in greater detail.

3.1.1 Samples

Blood samples from patients diagnosed with GSD Ib (n=5), GSD III (n=15), FBS (n=4), MMA mut-type (n=1) and MMA cblC-type (n=2) attending University Malaya Medical Center (UMMC) and Kuala Lumpur Hospital (HKL) were investigated. The blood samples were received in EDTA-containing blood collection tubes and stored in -20 ºC until further use. Use of the samples included in this study has received the approval of Medical Ethics Committee University Malaya Medical Centre reference number: UMMC MEC 2012/908.27 and written informed consent was obtained for all partcipants included in this study.

3.1.2 DNA Extraction

Genomic DNA was extracted from the peripheral blood leucocytes using QIAgen DNA extraction kit (QIAgen, USA) according to the manufacturer's instruction. Briefly, 20 µL of Proteinase K was added into a 1.5 mL microcentrifuge tube containing 200 µL of sample. Next, a total of 4 µL of RNAse A solution (100 mg/mL) was added into the tube to remove RNA. A total of 200 µL of Buffer AL was added to the mixture and it was homogenized by pulse-vortexing for 15 sec. After incubation at 56 ºC for 10 min, the tube was briefly centrifuged before addition of 200 µL of absolute ethanol to the sample. Later, the mixture was carefully transferred into QIAamp Spin Column and centrifuged at 6000 rcf for 1 min. Buffer AW1 (500 µL) was added into the column and further centrifuged to facilitate the washing step. Next, 500 µL of Buffer AW2 was added into the column and centrifuged at 20000 rcf for 3 min. Once the supernatant was discarded, the column was further centrifuged for 1 min to remove residual ethanol. To collect the DNA, QIAamp Spin Column was placed inside a clean 1.5 mL microcentrifuge tube and 200 µL of Buffer AE or distiiled water was added into the column. The column was let to stand for 1 min before

centrifuged at 6000 rcf for 1 min. Flowthrough containing the extracted DNA sample was kept in -20 ºC until future use.

3.1.3 DNA Quantification and Quality Estimation

The quantity and quality of extracted DNA was estimated using a spectrophotometer (Implen, Germany). Briefly, one microliter of sample was placed into the cuvette and absorbance ratios reading of A_{260}/A_{280} and A_{260}/A_{230} were obtained. The quality of extracted DNA was also estimated by subjecting the sample to agarose gel electrophoresis.

3.1.4 Agarose Gel Electrophoresis (AGE)

One percent of agarose gel powder (Promega, USA) was dissolved in 1 X TBE buffer (445 mM Tris base, 445 mM Borate and 10 mM EDTA) and heated. The molten gel was left to cool before it was stained with 10 mg/mL EtBr. While waiting for the gel to cool to a suitable temperature, a casting tray was fixed with a comb to be use for making the gel. Subsequently, the molten gel was poured into the casting tray and left to solidify. A total of 5 µL of PCR product mixed with 6 X loading dye (0.25 % Bromophenol blue, 0.25 % Xylene cyanole FF and 30 % glycerol) was loaded into the wells of the solidified gel. One microlitre of either 100 bp DNA ladder (Seegene, Korea) or 1 kb DNA ladder (Promega, USA) was used as molecular weight marker. The gel was then electrophoresed at 120 V for 25min. Viewing of the gel was carried out under 302 nm UV light using a gel documentation system (Alpha Innotech Corp, USA).

3.1.5 Primer Design

Primers used for the mutation screening were designed to amplify the genes related to each disease included in this study. Table 3.1 shows the disease name, the related gene and the gene reference sequence number. Primers were designed according to the specific gene reference sequence and the full primer sequences are listed in the

Appendix A. The designed primers were evaluated *in silico* using Sequence Manipulation Suite software (http://www.bioinformatics.org).

Table 3.1: Inherited metabolic disorders and its associated genes.

Disease	Gene	Gene reference sequence	Number of exons
Glycogen storage disease type 1b; GSD 1b	*SLC37A4*	NG_013331.1	9
Glycogen storage disease type III; GSD III	*AGL*	NG_012865.1	35
Fanconi-Bickel syndrome; FBS	*SLC2A2*	NG_008108.1	10
Methylmalonic aciduria; MMA mut-type	*MUT*	NG_007100.1	13
Methylmalonic aciduria with homocystinuria cblC type; MMA cblC-type	*MMACHC*	NG_013378.1	4

3.1.6 Optimization of Polymerase Chain Reaction (PCR)

Optimization was performed for each primer pair to determine the best PCR condition. Control DNA sample was used as template for the optimization experiment. All primer pairs were subjected to gradient PCR to determine the suitable annealing temperature. Briefly, eight identical tubes containing PCR reaction mixtures (described in section 3.1.7) were subjected to different annealing temperatures ranging from 50 ºC to 65 ºC. A total of 5 µL from the PCR product was then subjected to AGE as described in section 3.1.4. Products of gradient PCR that produced unspecific fragments were further subjected to $MgCl_2$ titration PCR. For this procedure, PCR reactions containing all the components as described in 3.1.7 but with varying concentration of $MgCl_2$ were prepared. The $MgCl_2$ concentration ranges from 0.5 mM to 2.5 mM. Similarly, the PCR reaction was analysed by subjecting the product to AGE. Amplification reaction that produced a single fragment with correct expected size and good band intensity were chosen as the optimized condition for a specific primer pair.

3.1.7 Polymerase Chain Reaction

A 50 µL PCR reaction containing 1 X PCR buffer A, 1.5 mM $MgCl_2$, 200 µM of each dNTP, 1 U *Taq* DNA polymerase (EURx, Poland), 10 µM of each primer and 50-100 ng of DNA template was prepared. DNA template was replaced with sdH_2O in the negative control reaction tube. In general, the PCR reaction was conducted with an initial denaturation step at 95 ºC for 3 min, 32 cycles of 95 ºC for 45 sec, specific annealing temperature for each primer set for 45 sec and an elongation step at 72 ºC for 1 min followed by a final extension at 72 ºC for 5 min. The reaction was carried out in an Arktik thermocycler (Thermo Scientific, USA). A total of 5 µL from the PCR reaction product was then subjected to AGE as described in section 3.1.4.

3.1.8 PCR Product Purification Method

PCR products were purified either by using QIAquick PCR purification kit (QIAgen, Netherlands) or using QIAquick gel extraction kit (QIAgen, Netherlands).

PCR products with no unspecific bands were purified using QIAquick PCR purification kit according to the manufacturer's recommendation. Briefly, replicates of the same PCR reaction were pooled into a sterile 1.5 mL microcentrifuge tube. Next, five volumes of Buffer PB was added to one volume of PCR reaction and mixed. The mixture was then transferred into QIAquick column and centrifuged at 17900 rcf for 60 sec. Once the supernatant was discarded, 750 µL of Buffer PE was added and centrifuged to wash the column. To remove any residual wash buffer, the column was centrifuged again once the supernatant was discarded. Elution of DNA was obtained by placing the column into a clean 1.5 mL microcentrifuge tube and 30 µL of Buffer EB was added. The column was left to stand for 1 min and then centrifuged. Agarose gel electrophoresis was carried out afterwards to analyzed the purified DNA. The product was stored at -20 ºC until future use.

PCR products that contain unspecific bands or primer dimer were purified using QIAquick gel extraction kit. Initially, PCR replicates were pooled and subjected to agarose gel electrophoresis. Subsequently, a sterile 1.5 mL microcentrifuge tube was weighed. Next, the band of interest was viewed under UV light and later excised from the gel using a clean scalpel. The excised gel was then transferred into the weighed 1.5 mL microcentrifuge tube and the weight of gel-containing tube was recorded. Next, three volume of Buffer QG were added to one volume of excised gel (100 mg = 100 µL). The excised gel was dissolved by heating at 50 ºC for 10 min. Once the gel had completely dissolved, one gel volume of isopropanol was added to the sample and the mixture was transferred into QIAquick spin column. Subsequently the column was centrifuged at 11300 rcf for 1 min and the flow-through was then discarded. Next, a total of 500 µL of Buffer QG were added into the spin column and centrifuged again. After discarding the flow-through, 750 µL of Buffer PE was added into the spin column and further centrifuged. An additional centrifugation step was carried out after the flow-through was discarded to remove traces of ethanol residues in the sample. Finally, the spin column was placed into a sterile 1.5 mL microcentrifuge tube and 30 µL of sdH_2O were added to the centre of the membrane in QIAquick spin column and let to stand for 1 min. The DNA was eluted by centrifugation at 11300 rcf for 1 min and later analysed by AGE, while the remaining product was stored at -20 ºC until future use.

3.1.9 Sequencing

Purified samples were subjected to sequencing using BigDye™ Terminator Cycle Sequencing kit version 3.1 (Applied Biosystem, USA) and electrophoresed in 3130*xl* Genetic Analyzer (Applied Biosystem, USA). To prepare the samples for sequencing, initial reaction of cycle sequencing was carried out to amplify the region of interest. Reaction was carried out in a 96-well plate with total volume of 20 µL per well containing 1 X BigDye Sequencing buffer, 3.2 pmol of primer and 50 ng of DNA. The

primers used for DNA sequencing were similar as the ones used for PCR amplification. However, nested primers were used for DNA sequencing of the *AGL* gene (as listed in Appendix B). The thermocycling profile used was denaturation at 96 ºC for 1 min, 25 cycles of 96 ºC for 10 sec, annealing at 50 ºC for 5 sec and extension at 60 ºC for 4 min followed by 4 ºC for 1 min. Subsequently, the PCR product was purified by ethanol/EDTA precipitation. Briefly, the plate containing the PCR product was centrifuged before five microlitres of 125 mM EDTA were added into each reaction well. Next, 60 µL of 100 % ethanol was added and the plate was sealed and vortexed. The mixture was then left to incubate at room temperature for 15 min. Once incubated, the plate was centrifuged at 2250 rcf for 45 min. After the supernatant was discarded, 60 µL of 70 % ethanol was added into each well and further centrifuged at 1650 rcf for 15 min at 4 ºC. Pellet formed was dried using a thermocycler. The program used was 90 ºC for 1 min followed by 50 ºC for 5 min. Next, a volume of 10 µL of Hi-Di formamide was added into each well and denatured at 95 ºC for 5 min. Once denatured, the plate was placed on ice immediately before electrophoresis on the 3130*xl* Genetic Analyzer.

3.1.10 Sequencing Analysis

The sequencing results were analyzed using Applied Biosystem Sequence Analysis software v5.3 and Chromas software (Technelysium Ptd Ltd). The sequences of each exon for each gene was analysed by comparing with the gene reference sequence as listed in Table 3.1. Mutations found were designated according to standard nomenclature rules and further cross-checked with the gene mutation list in the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php) and 1000 Genomes database (http://www.1000genomes.org).

3.2 Determination of Mutant Allele Frequency in the Population

To obtain support that the unreported mutations found were causative of the diseases in question and not neutral polymorphisms, a total of 150 healthy control samples (Malay *n*=50, Chinese *n*=50 and Indian *n*=50) were screened for their presence. Where possible, an assay based on restriction enzyme digestion was used to facilitate screening as this would be the cheaper alternative. In cases where a restriction enzyme digestion assay cannot be used, a TaqMan® SNP Genotyping assay was designed for genotyping purposes. The control samples were kindly donated by Assoc. Prof. Dr Ng Ching Ching (Faculty of Science, UM) and Dr Azlina Ahmad Annuar (Faculty of Medicine, UM). Informed consent from control individuals were obtained for the purpose of this study. Patients' samples that carry the unreported mutations were also included in the experiment. This part of the study was only carried out for the *AGL* gene due to time limitations.

3.2.1 Restriction Enzyme (RE) Digestion Assay

Restriction enzymes used for the digestion assay were determined through restriction enzyme mapping softwares that is available online at http://nc2.neb.com/NEBcutter2/ and Sequence Manipulation Suite (http://www.bioinformatics.org/sms2/rest_summary.html). The desired fragments were initially amplified via standard PCR (refer to section 3.1.7) for each specific fragment. PCR products were then digested with the appropriate restriction enzyme as listed in Table 3.2 and carried out according to the manufacturer's instruction. In general, the RE digestion assay was conducted in a final volume of 20 µL containing 10 X buffer, 1 U of appropriate restriction enzyme and 1 µg of PCR product. For positive control reactions, 1 µg of the PCR product carrying the mutated allele was used while for the negative control reaction, restriction enzyme was substituted with sdH_2O. The reaction tube was then incubated at 37 ºC for 16 hours and the inactivation was carried out at

appropriate conditions depending on the restriction enzyme used. Later, the restriction digests product were electrophoresed on 1 % - 3 % gel (depending on expected fragment size) at 120 V for 25 min as described in 3.1.4. Banding patterns were observed between the positive control, the negative control and the other samples.

3.2.2 TaqMan® SNP Genotyping Assay

Two unreported mutations of the *AGL* gene, c.1423+1G>T and c.3814_3815delAG were subjected to TaqMan® SNP genotyping assay as there were no suitable restriction enzyme assay that could be designed to detect these mutations. Mutation c.4490G>A was also subjected to this assay as the RE results was not convincing. Primers and probes' sequences were designed based on *AGL* gene reference sequence (GenBank accession number: NG_012865.1) using Primer Express software (Applied Biosystem, USA), as listed in Table 3.3. All PCR reactions were carried out using the TaqMan® GTXpress Master Mix (ABI, Applied Biosystems) and according to the manufacturer's instructions. Briefly, reaction tube with a total volume of 10 µL containing 2 X TaqMan® GTXpress Master Mix, 20 X TaqMan® Genotyping assay mix, and five nanograms of genomic DNA template was prepared. No-template tube was also prepared as negative control and tube containing patient's DNA sample with known genotype was included as positive control. Genotyping was then performed using *QuantStudio 12K Flex Real Time PCR System* (Life Technologies, USA). The PCR conditions used were activation at 95 ºC for 10 min, followed by 40 cycles of denaturation at 95 ºC for 15 sec and annealing/extension at 60 ºC for 60 sec. Fluorescence signal generated by the different fluorophore-labelled probe was analysed as clustered patterns in a bivariate plot using the QuantStudio™ 12K Flex software.

Table 3.2: Restriction enzymes used to digest PCR fragment carrying mutated sequence.

Mutation	Restriction enzyme	Assay conditions		Inactivation conditions	
		Temperature (°C)	Incubation time (hours)	Temperature (°C)	Incubation time (min)
c.2914_2915delAA	Cfr10I (BsrFI)	37	16	-	-
c.4333T>G	Sau3AI (Bsp1431I/ BfuCI)	37	16	65	20
c.4490G>A	StyI	37	16	65	5
c.4531_4534delTGTC	HpyCH4III	37	16	80	20

Table 3.3: TaqMan® assay primers and probes sequences to screen unreported mutations in the population.

TaqMan assay	Primers and probes' sequences	Amplicon size (bp)
c.1423+1G>T	Primers Forward: 5'-GGGAGATGATCCTCTTCGAAACTTT-3' Reverse: 5'-CTTTCCCCATCCACAGAGAAGTTAA-3' Probes Normal allele: 5'VIC- CTGAACCGGGTATGTAA-3' Mutant allele: 5'FAM- CTGAACCGG**T**TATGTAA-3'	68
c.3814_3815delAG	Primers Forward: 5'-GTGGCACATGGATGGATAAAATGG-3' Reverse: 5'-CATTTACACTACCTTGGTGTGGCT-3' Probes Normal allele: 5'VIC- TAGAAACAGAGGAATCC-3' Mutant allele: 5'FAM- CTAGAAAC^AGGAATCC-3'	82
c.4490G>A	Primers Forward: 5'-GGTCTTGCCTATTTTGTTAATATGAATTAAATTATGTCT-3' Reverse: 5'- GCATTCTCATTGGTCAGTTCTGGAA-3' Probes Normal allele: 5'VIC- CAGATCCCCTTGGAAAG-3' Mutant allele: 5'FAM- TTCAGATCCCCTT**A**GAAAG-3'	108

* bold and underlined letter indicates mutation
^ indicates deletion

3.3 *In Silico* Prediction of the Unreported Mutations

Prediction on whether a mutation results in mRNA splicing error was performed using the Human Splice Finder (HSF) version 3, freely available at http://www.umd.be/HSF3/. The HSF system combines various algorithms based on Position Weight Matrices and Maximum Entropy (MaxEnt) method to identify and predict the effect of mutations on the functionality of splicing motifs including the acceptor and donor splice sites. For HSF calculations, consensus values range from 0 to 100, and the threshold is defined at 65. Scores above the threshold are considered as indicators of the presence of splice sites. If the known splice site (WT) score is above the threshold and the score variation (between WT and Mutant) is under -10 % for HSF, this is considered to indicate that the splice site has been disrupted. For MaxEnt calculations, consensus values range from -20 to +20, and the threshold is defined at 3. Similar to HSF, scores above the threshold are considered as indicators for the presence of splice sites. If the known splice site (WT) score is above the threshold and the score variation (between WT and mutant) is under -30 % for MaxEnt, this is indicative of the splice site has been disrupted.

Prediction of the consequence of an amino acid substitution on the structure and function of a human protein was performed using PolyPhen-2. This freely available software (http://genetics.bwh.harvard.edu/pph2) predicts the effects of amino acid changes based on a number of features comprising the sequence, phylogenetic and structural information characterizing the substitution. PolyPhen-2 calculates the probability a mutation is damaging using supervised machined-learning trained Naïve Bayes classifier. The analysis is reported as estimates of false positive rate (FPR) and true positive rate (TPR). FPR is defined as 'the chance that the mutation is classified as damaging when it is in fact non-damaging' while TPR is defined as 'the chance that the mutation is classified as damaging when it is indeed damaging'. A mutation is also

analysed qualitatively based on FPR threshold defined at 10 % / 20 % and reported as 'benign', 'possibly damaging' or 'probably damaging'. Mutations with posterior probability value below 10% are predicted to be probably damaging, while mutations at 10 % or below 20 % are predicted to be possibly damaging. Mutations with values higher than 20 % are predicted to be benign. A mutation will be reported as unknown if there is lack of data available to carry out the prediction.

The effect of deletion, insertion and nonsense mutation on the frameshift alteration of gene sequence was investigated using Sequence Manipulation Suite (http://www.bioinformatics.org/sms2/trans_map.html). The software accepts a DNA sequence and returns a textual map displaying the translated protein.

3.4 Determination of Mutation Effect on mRNA Splicing

One unreported mutation (c.1423+1G>T) occur at the exon intron boundaries. To investigate the effect of the mutation on splicing regulation of the gene, mRNA splicing assay was carried out. The c.2681+1G>A mutation which had been previously shown to cause abnormal splicing was used as a reference/control assay (Hadjigeorgiou et al., 1999).

3.4.1 RNA Extraction

Blood samples were obtained from the patients carrying the c.1423+1G>T and c.2681+1G>A mutations. Total RNA was extracted from the peripheral blood diluted in Roche RNA/DNA Stabilization reagent, ratio 1:10 (Roche, USA) using Trizol-LS reagent (Life Technologies, USA) according to the manufacturers' instruction. Briefly, 750 µL of Trizol-LS was added to 250 µL of sample (ratio3:1). Sample was homogenized by pipeting and let to incubate for 5 min. A total of 200 µL of chloroform was added and vigourously shaken by hand for 15 sec. Next, the mixture was incubated for five minutes at room temperature before centrifugation at speed 12000 rcf for

15 min at 4 ºC. Colourless aqeuos phase containing the RNA was carefully removed and transfered into new sterile microcentrifuge tube. 500 µL of 100 % isopropanol was added to aqeuos phase and mixed gently before incubation for 10 min at room temperature. RNA was pelleted by centrifugation at 12000 rcf for 10 min at 4 ºC. Supernatant was discarded and pellet was washed by exerting vortex with 1000 µL of 75 % ethanol. The mixture was then centrifuged for 5 min, 4 ºC at speed 7500 rcf. The supernatant was discarded and RNA pellet was resuspended in 50 µL RNase-free water. Suspension was incubated in water bath at 55-60 ºC for 10 - 15 min. The quantity and quality of extracted RNA was estimated using a spectrophotometer and AGE as described in section 3.1.3/ 3.1.4.

3.4.2 Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

RT-PCR was performed using SuperScript® III One-Step RT-PCR System with Platinum® *Taq* DNA polymerase according to the manufacturer's instruction (Invitrogen, USA). Primers used in this assay were either newly designed or adapted from Shaiu et al. (2000) (Table 3.4). The RT-PCR cycling parameters used were 45 ºC for 30 min, 94 ºC for 5 min, 30 cycles of 94 ºC for one minute, 58 ºC for one minute and 72 ºC for one minute, followed by final extension at 72 ºC for 5 min. Electrophoresis on 1 % gel was performed at 100 V for 30 min to visualize RT-PCR product. The product was then purified using QIAquick gel extraction kit and sequenced. Primers used for sequencing was similar as used during RT-PCR.

Table 3.4: Primers used for RT-PCR of intronic mutations.

Primer	Sequence
Demm2(1)	5'-TACAAAGAAAAGGGAATACC-3'
DE1560(2)	5'-CTCAGCTACGTGAAGAGGTG-3'
aglRNA_F3F	5'-AAACTAGGACCTGTCACTAGAAAGC-3'
DE1655(2)	5'-GTAACAAAGACATTGTCCAG-3'
DEmm3(1)	5'-TACAAAGAAAAGGGAATACC-3'
DE3083(2)	5'-ACTGAACCCAATGAAAGGTG-3'

3.5 In Vitro Expression of Mutant Gene

To observe the effect of other unreported mutations (c.2914_2914delAA, c.3814_3815delAG, c.4333T>G, c.4490G>A and c.4531_4534delTGTC) on the *AGL* gene, plasmids carrying mutant *AGL* gene were constructed and expressed in human liver carcinoma cell line HepG2. The following explains the experimental design and procedures involved.

3.5.1 *AGL* Gene Construct

Wild type *AGL* gene construct was purchased from Origene (USA) (Figure 3.1). Initially, the lyophilized clone stock was resuspended with sdH_2O to a concentration of 1 µg/ µl. A total concentration of 250 ng of the suspended clone was used for subsequent analysis. The construct was screened by sequencing to confirm the insert.

3.5.2 Site-Directed Mutagenesis (SDM)

Five *AGL* gene constructs carrying the desired mutations were created using Phusion site-directed mutagenesis kit (Thermo Scientific, USA) as recommended by the manufacturer. The mutagenesis was achieved by preparing a reaction tube with a total volume of 50 µL consisting of 1X Phusion HF Buffer, 10 mM dNTPs, 0.5 µM forward and reverse primers, 10 ng of DNA template and 0.02 U/µL Phusion Hot start DNA polymerase. Phosphorylated primers used were listed in Table 3.5. The mixture was

subjected to PCR and the cycling setting were as followed: initial denaturation at 98°C for 30 s, denaturation at 98°C for 10 s, annealing at 65-72°C for 30 s, extension at 72°C for 30 s. The denaturation, annealing and extension was carried out for 25 cycles and the final extension was at 72°C for 10 min. The PCR product was sequenced (as described in section 3.1.8) to confirm successful mutagenesis. Successful product was used for the ligation step.

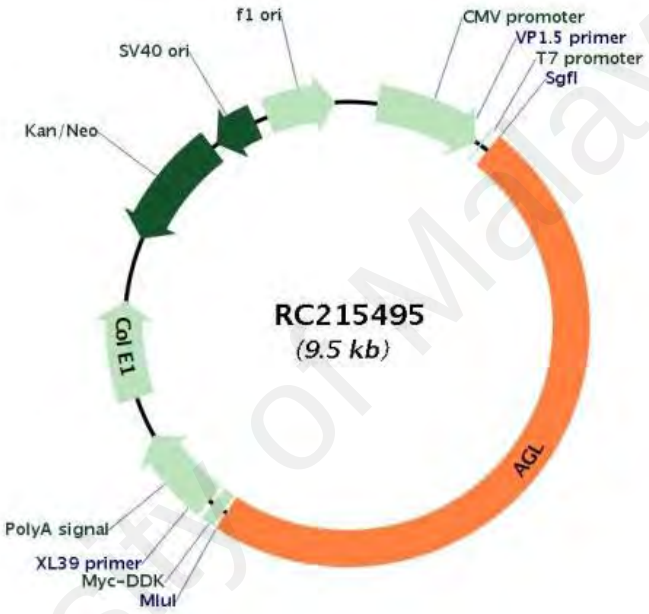


Figure 3.1: Plasmid map of *AGL* gene in pCMV6-Entry vector (RC215495). (Source: https://www.origene.com)

3.5.3 Ligation

To circularized the PCR product, ligation was carried out using T4 DNA Ligase kit (New England Biolabs, UK). Two microliter of 5X Rapid Ligation Buffer was added into a tube containing 5 µL (20 ng) of site-directed mutagenesis product. The reaction volume was adjusted to 9.5 µL using sdH_2O and mixed thoroughly. A volume of 0.5 µL T4 DNA ligase was added to the reaction mixture to make a total reaction volume of 10 µL. The mixture was then centrifuged briefly and incubated at 4 °C overnight before used for transformation of competent cells.

Table 3.5: Phosphorylated primers used for site-directed mutagenesis.

Mutation	Primer	Sequence
c.2914_2915delAA	sdm_agl_972_f	5'-CCGGCTTATTTCACGATCAGGAAC-3'
	sdm_agl_972_r	5'-ACTGACATAGTCAATCATCCAATCTC-3'
c.3814_3815delAG	sdm_agl_1272_f	5'-GAATCCCAGCCACACCAAGAGATG-3'
	sdm_agl_1272_r	5'-CTGTTTCTAGCTCTGTCACTTTCTCC-3'
c.4333T>G	sdm_agl_1445_f	5'-CTAAAGGTTTCAATGATCACCAAGGAC-3'
	sdm_agl_1445_r	5'-CAAGATTGTAGTTGTCATTGTCTAATGC-3'
c.4490G>A	sdm_agl_1497_f	5'-AGATCCCCTTAGAAAGGACTTCC-3'
	sdm_agl_1497_r	5'-CTCAAGATGAACATAATGTCGGGAAAG-3'
c.4531_4534delTGTC	sdm_agl_1511_f	5'-CTTTCAGCTGTGAAACACAAGCCTGGTC-3'
	sdm_agl_1511_r	5'-GTACTGGGCATTCTCATTGGTCAG-3'
	sdm_agl_1511b_f	5'-TGCAATTACTTGTATTATAGACGCGTACGCGGCCGCTCGAG-3'
	sdm_agl_1511b_r	5'- TACTTAATATCTGTAATAAACTATAAATCATAAAGTGTCTC-3'

3.5.4 Transformation

E. coli competent cells strain TOP10 (Invitrogen, USA) were used to transform the ligation product. About 10 µL of ligation product was briefly centrifuge and transferred into tube containing 100 µL of competent TOP10 cells. The mixture was mixed gently by flicking and subsequently incubated on ice for 30 min. Later, the mixture was subjected to a heat-shock at 42 °C for 50 sec and was then immediately placed on ice for two minutes. Next, 900 µL of Luria-Bertoni (LB) broth (Promega, USA) was added to the sample and incubated for 3 hours at 37 °C with shaking. Once incubated, the tube was centrifuged for two minutes at 3000 rcf and 600 µL of the supernatant was discarded. The remaining supernatant was used to resuspend the pellet. A total of 100 µL of the mixture was then transferred into LB plate containing 25 µg/mL of Kanamycin. Lastly, the sample was spread evenly on the plate and incubated overnight at 37 °C.

3.5.5 Colonies Selection

White colonies were observed on plate after overnight incubation. The colonies may carry either the transformed or untransformed construct. To carry out the selection, a mini library was prepared by transferring the selected colony onto 6 x 6 grid LB-Kan plate containing 25 µg/mL Kanamycin using an inoculate loop. The same colony was also transferred into a tube containing 100 µL of sdH_2O. The mixture was then boiled at 99 °C for 10 min. Subsequently, the mixture was used as template for colony PCR (protocol as described in section 3.1.7) and the PCR product was later subjected to both AGE and sequencing. Primers used for PCR and sequencing of the plasmid were listed in Table 3.6.

Table 3.6: Primers used for colony PCR.

Primers	Sequences
VP1.5	5'-GGACTTTCCAAAATGTCG-3'
aglRNA_F1R	5'-CAGGCAGGTTTTAAGTGTGGAG-3'
aglRNA_F2F	5'-TGGAATGATGTTGGACAGCTAGTG-3'
aglRNA_F2R	5'-CATTGTGTGCCATCAGAAAACAAGC-3'
aglRNA_F3F	5'-AAACTAGGACCTGTCACTAGAAAGC-3'
aglRNA_F3R	5'-ACCTGTGTTTGAAGGCAATGCTTC-3'
aglRNA_17F	5'- TTATTGCAGCCAGGTGTGC-3'
aglRNA_20R	5'- CTTTTGTGGCAACTCCA-3'
DE1996(1)	5'-GAAGCATTGCCTTCAAAC-3'
DEmm3(2)	5'-TACATCCATTAGGGCAGGTG-3'
aglRNA_F5F	5'-TAAATCTGGCAGCCTAGCTGTTG-3'
DE3083(2)	5'-ACTGAACCCAATGAAAGGTG-3'
DE2913(1)	5'-GCAGGCTATGTTCTTCTACC-3'
DE3799(2)	5'- ATATTAAAACCTTCGTCCTTCAT-3'
aglRNA_28F	5'- CACATGCAGGGCATAC-3'
aglRNA_32R	5'- CTAAAGTTTTCATGCCAAGG-3'
aglRNA_F7F	5'-GAGCTTCAAGTCCTTGGTGTGAC-3'
XL39	5'-ATTAGGACAAGGCTGGTGGG-3'

3.5.6 Plasmid Extraction

Bacterial colony that was confirmed to carry the mutant gene was cultured overnight at 37 °C in 10 mL of LB broth containing 25 µg/mL of Kanamycin. A volume of 850 µL of the overnight culture was subsequently transferred into a sterile 1.5 mL microcentrifuge tube containing 150 µL of sterile 100 % glycerol and kept in the -80 °C as a stock. The rest of the overnight culture was transferred into a 15 mL falcon tube and centrifuged at 13000 rcf for 5 min at room temperature. Supernatant was discarded, and bacteria pellet was resuspended with 200 µL of Solution I (50 mM glucose, 10 mM EDTA and 25 mM Tris-HCl pH 8) by vortexing. Once the pellet was dissolved, the suspension was transferred into a sterile 1.5 mL tube and 200 µL of Solution II (0.2 N NaOH and 1 % SDS) was added into the suspension. The mixture was mixed gently for

4 min at room temperature. Next, 200 µL of Solution III (3 M KoAc and 10 % acetic acid) were added and mixed gently for 15 min at 0 °C. The mixture was then centrifuged at 11300 rcf for 10 min and the resulting supernatant was transferred into a sterile 1.5 mL microcentrifuge tube. Five microliters of RNAse A (stock 50 mg/mL) were added into the tube and mixed gently before incubation for 3 hours at 37 °C. Subsequently, 600 µL of phenol solution were added to the incubated solution and centrifuged at 11300 rcf for 3 min. The aqueous phase was transferred into a sterile 1.5 mL microcentrifuge tube and 600 µL of chloroform was added and mixed. It was further centrifuged at 11300 rcf for 3 min and the aqueous phase was again transferred into a new 1.5 mL microcentrifuge tube. Subsequently, 0.1 V of 5 M NaCl and 2.5 V of isopropanol were added to the aqueous phase and incubated on ice for 20 min. Later, the solution was centrifuged at 11300 rcf for 15 min and the supernatant was discarded. The resulting pellet was washed with one millilitre of 70 % ethanol and dried at 50 °C for 10 min. Finally, the pellet was resuspended in 30 µL sdH_2O and placed in 4 °C overnight. The quantity and quality of the extracted plasmid DNA was estimated the following day as described in section 3.1.3.

3.5.7 Transfection

Human liver carcinoma (HepG2) cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % Fetal Bovine Serum (FBS) and 1 % Penicillin/Streptomycin) (GIBCO, USA) at 37 °C in a 5 % CO_2:95 % air-humidified atmosphere. The cells were sub-cultured every two days to maintain the cells confluency at 70-80 %. For transfection, the cells were seeded at density of $25x10^4$ in a 6-well cell culture plate one day before transfection. Plasmid containing the construct was transfected into the cell using calcium-phosphate transfection method. Briefly, plasmid containing desired construct was vortexed and homogenized by centrifugation at 11300 rcf for 15 sec. Next, one microgram of the plasmid was transferred into a

sterile 1.5 mL microcentrifuge tube and the volume was adjusted to 93.7 µL with sdH_2O. Next, 31.3 µL of $CaCl_2$ was added into the tube and mixed. After incubation of 5 min at room temperature, 125 µL of 2X HBS (50 mM HEPES pH7, 1.5 mM Na_2HPO_4, 140 mM NaCl, 10 mM KCl, 12 mM Dextrose) were added into the tube and further incubated for 5 min. The mixture was gently pipetted into the cell culture plate seeded with HepG2 and subsequently incubated overnight at 37 °C. The transfection efficiency was assessed the following day using 1X73 Olympus inverted microscope (Olympus, Japan). The mutant construct did not carry any fluorescence marker gene. To assess the transfection efficiency, HepG2 cells was transfected with plasmid carrying the cyan fluorescence marker (cfp). It was assumed that the *AGL* gene construct will have the same transfection efficiency as the 6-well culture plate transfected with cfp plasmid.

3.5.8 Stable Transfection

Due to low transfection efficiency, stable transfection approach was carried out. Briefly, transfection was carried out as described in section 3.4.7 and the cells were subsequently grown in DMEM containing 10 % FBS and 800 µg/ml of Geneticin (G418, GIBCO, USA) at 37 °C. The cells were gradually transferred into a 10 cm and 15 cm culture plate once the confluency reached 75 %. Fresh DMEM media was supplemented to the cells every three days until cell colonies formed. A total of 12 cell colonies were selected for each construct and transferred into a 24-well culture plate containing the culture media. The cells were grown and gradually transferred into a 6-well culture plate.

3.5.9 Transfection Confirmation

To confirm successful transfection, PCR was carried out on the harvested HepG2 cells. DNA extraction was carried out using QIAgen DNA extraction kit

(QIAgen, Germany) as recommended by the manufacturer. Initially, stably transfected HepG2 cells grown in 6-well culture plate were trypsinized and counted to determine the cell number count. Appropriate volume of suspended cells were transferred into a sterile 1.5 mL microcentrifuge tube and centrifuged at 1500 rcf for 5 min at 4 °C to pellet the cells. Resulting supernatant was discarded and 200 µL of 1 X PBS buffer (GIBCO, USA) were added into the tube. Subsequently, 20 µL of Proteinase K were added and followed by 200 µL of Buffer AL. The mixture was immediately pulse-vortexed for 15 sec and incubated at 56 °C for 10 min. Next, 200 µL of 100 % ethanol was added and mixed by pulse-vortex for 15 sec. The tube was briefly centrifuged afterward to remove drips from inside the lid. Later, 620 µL of lysate were transferred into the QIAamp spin column and centrifuged at 6000 rcf for 1 min. The subsequent steps were carried out as described in 3.1.2. Extracted DNA was later used for PCR reaction as described in 3.1.5 while the primers used were listed in Table 3.7.

Table 3.7: Primer sequences used to detect successful transfection.

Mutation	Primers	Sequences
c.2914_2915delAA	AGL972F AGL972R	5'-GCTTCTCTTGCCTCCAGATTAAC-3' 5'-CCAATGAAAGGTGTTTCACAAAGG-3’
c.3814_3815delAG	AGL1272F AGL1272R	5'-AAGTCATACAGGAAGCAATGC-3' 5'-GCCACGTTTGTGAACCAG-3'
c.4333T>G	AGL1445F AGL1445R	5'-GTCCTTGGTGTGACTATCAGC-3' 5'-GGTCAGTTCTGGAAGTCCTTTC-3'
c.4490G>A	AGL1497F AGL1497R	5'-CAATTATCACCAAGGACCTG-3' 5'-GTTCAGGAAACAGCTATGAC-3'
c.4531_4534delTGTC	AGL1497F AGL1497R	Same as above

3.5.10 Expression Confirmation

To confirm the expression of the mutant construct in the cell, RT-PCR was carried out on the harvested RNA from HepG2 cells. Initially, RNA was extracted from

the cultured cells using Easy spin™ Total RNA Extraction Kit (Intron Biotechnology, Korea) as suggested by the manufacturer. To extract RNA, cells were trypsinized and a total number of approximately $100x10^4$ cells were counted and transferred into a 1.5 mL tube. The tube containing the cells was centrifuged at 11300 rcf for 10 sec to remove the media. Next, 1 mL of Lysis buffer was added into the tube and vigorously vortexed in room temperature for 10 sec. A total of 200 µL of chloroform was added into tube, vortexed and centrifuged for 10 sec at temperature 4 °C and 11300 rcf. Next, the upper aqueous layer was transferred into a sterile 1.5 mL microcentrifuge tube and 400 µL binding buffer was added and mixed well by pipetting or gentle inversion. The mixture was incubated for 1 min at room temperature and then loaded into spin column. Centrifugation was carried out at 11300 rcf for 30 sec and the resulting supernatant was discarded. A total volume of 700 µL of Washing buffer A was added into the column and centrifuged before addition of 700 µL of Washing buffer B into the column. The mixture was centrifuged again and supernatant was discarded. Following that, the tube was centrifuged to remove ethanol residues and the tube was air dried. Finally, the column was placed inside a sterile 1.5 mL microcentrifuge tube and 50 µL of elution buffer was added to collect RNA. The concentration of the extracted RNA was quantified using nanospectrophotometer (Implen, Germany) and the RNA integrity was confirmed by gel electrophoresis. The RNA product was then used for RT-PCR (as described in section 3.3.2) and later sequenced (as described in section 3.1.9). All primers used were as tabulated in Table 3.7.

3.6 Protein Analysis

Mutant protein was generated and collected for analysis. The harvested total protein lysate was subjected to ultracentrifugation to separate glycogen-enriched fraction and later used for further analysis.

3.6.1 Harvesting of Protein Lysate

Expressed protein was harvested using nuclear lysis buffer (1 M Tris pH7.5, 5 M NaCl, 1 M $MgCl_2$, 0.5 M EDTA, 100 % glycerol, 1 M DTT, 100 mM PMSF and 1 X Protease inhibitor). All reagents used were purchased from Promega, USA. Briefly, culture media was removed from the culture plate containing the cells. The plate was washed with 1 X phosphate-buffered saline (PBS) before the cells were trypsinized. Fresh media was then added into the plate to neutralize the trypsin and resuspended before being transferred into a 15 mL falcon tube. The tube was centrifuged at 11300 rcf for 5 min. Once the supernatant was discarded, the cells were further washed twice with 1 X PBS and centrifuged. During the last wash, 1mL of the supernatant was left in the tube and used to resuspend the pellet. Ten microliters of the suspension was placed on a haemocytometer for cell counting. Once the cell number was determined, the cell suspension was centrifuged at 11300 rcf for 1 min to remove the PBS. Cell pellet was then resuspended in lysate buffer and incubated on ice for 10-15 min. New clean 1.5 mL microcentrifuge tube was also incubated on ice. Later, the incubated cells were centrifuged at 20000 rcf for 10 min at 4 °C. Supernatant containing crude cell lysate was then transferred into a pre-chilled 1.5 mL microcentrifuge tube and stored at -80 °C until later use.

3.6.2 Isolation of Glycogen-Enriched Fraction

Media were removed from the cultured cells and washed twice with 1 X PBS. Next, the cells were scraped in 50 mM Tris (pH 7.5) and 150 mM NaCl supplemented with complete EDTA-free protease inhibitor (Roche,Germany) using rubber cell scrapper (SPL, Korea). Suspended cells were then transferred into a clean 50 mL falcon tube and further sonicated at 30 % maximal setting, 10 pulses (1 sec/pulse). Cell extracts were then centrifuged at 3000 rcf for 10 min to collect cell debris. The supernatant was then transferred into another clean 50 mL tube and centrifuged at

14000 rcf for 10 min and then at 100,000 rcf for 1 hour to obtain the cytosol (supernatant) and glycogen-enriched fraction (pellet). The pellet was then solubilized in lysis buffer containing 50 mM Tris pH7.5, 150 mM NaCl, 1 % Triton X-100, complete EDTA-free protease inhibitor by rotation for 15 min. Lysate was centrifuged once more at 14000 rcf for 5 min to remove unwanted debris. The lysate was then stored at -80 °C until further use.

3.6.3 Concentrating Protein by Trichloroacetic Acid (TCA) Method

Protein lysate were concentrated using the trichloroacetic acid method. A starting volume of 500 µL protein lysate was transferred into a clean 1.5 mL microcentrifuge tube. Subsequently, 50 µL of trichloroacetic acid were added into the tube and incubated on ice for 30 min. Pellet was formed by centrifugation at 18000 rcf for 15 min at 4 °C and supernatant was discarded. Washing was carried out twice by adding 500 µL of ice-cold acetone into the tube and centrifuged at 18000 rcf for 5 min at 4 °C. Once the supernatant was discarded, tube containing protein pellet was air-dried for 10 min. Pellet was dissolved in one volume PBS/5 % SDS and stored at -80°C.

3.6.4 Sodium-Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE was carried out to observe the presence of desired protein in the cell extract. The gel cassette (Bio-Rad, USA) was assembled according to manufacturer's instruction. A 10 % resolving gel (1.5 M Tris-HCl pH 8.8, 30 % acrylamide/0.8 % bisacrylamide, 10 % SDS, 10 % APS, TEMED and sdH_20) and 4.5 % stacking gel (0.5 M Tris-HCl pH 6.8, 30 % acrylamide/0.8 % bisacrylamide, 10 % SDS, 10 % APS, TEMED and sdH_20) were prepared. The gel was let to solidify for 30-45 min before being transferred into the electrophoresis tank filled with 1 X electrophoresis buffer (25 mM Tris, 200 mM glycine and 3.5 mM SDS). Sample that

was mixed with 2 X SDS-PAGE sample buffers (125 mM Tris-Cl pH 6.8, 20 % glycerol, 10 % β-mercaptoethanol, 4 % SDS and 0.05 % bromophenol blue) were heated at 99 °C for 4 min before being electrophoresed at 180 V for 60 min. Precision Plus Protein Standards (Bio-Rad, USA) was used as protein molecular weight reference. Following electrophoresis, the gel was fixed with fixing solution (40 % methanol and 10 % acetic acid) for 30 min with gentle shaking. Next, the gel was stained with Coomassie Blue solution (50 % methanol, 0.05 % Coomassie Brilliant Blue R250 & 10 % acetic acid) for one hour with gentle shaking followed by washing with destaining solution (20 % ethanol and 7 % acetic acid) for several minutes until bands clearly appeared.

3.6.5 Western Blotting

Western blotting was carried out to detect the wild type and mutant proteins. Initially, the gel was prepared and electrophoresed as described in 3.5.3. Subsequently, Hybond™-C extra membrane (GE Healthcare, USA), spans and Whatman No 1 filter paper (Sigma Aldrich, USA) were soaked in Transfer buffer (Appendix C) at 4 °C. Next, gel was taken out from the cassette and rinsed with distilled water before being assembled with the membrane into a sandwich clamp. The sandwich cassette was placed into the tank with ice block and magnetic bar. Transfer buffer was added into the tank until the sandwich cassette was submerged and electrophoresis was carried out at 100 V for 1 hour. Next, the membrane was soaked overnight in Blocking buffer (Appendix C) at 4°C. In the following day, the membrane was soaked in Probing solution (Appendix C) with anti-*AGL* antibody (Santa Cruz, USA) at dilution 1:500 for 2.5 hours with gentle shaking. Afterward, the membrane was rinsed with 1 X TBS (Appendix C) and washed five times with 1 X TBST (Appendix C) for 5 min. Next, the membrane was rinsed again with 1 X TBS and later soaked in probing solution containing anti-goat antibody (Sigma Aldrich, USA) for 2 hours with gentle shaking.

Subsequently, the membrane was rinsed again with 1 X TBS, washed three times with 1 X TBST and another rinsed with 1 X TBS. Four millilitres of alkaline-phosphatase conjugate substrate was added onto the membrane and kept in the dark. Once the bands developed, the reaction was stop by addition of distilled water.

CHAPTER 4: RESULTS

This chapter details the results from the experimental work conducted in this study. DNA samples from a total of 27 patients and 16 parents (when made available) were analysed for five different types of inherited metabolic disorders (IMDs). The results are presented in the following sections. Section 4.1 provides the results for mutation screening of the GSD III patients. Section 4.2 provides the results for mutation screening of other GSD types (GSD 1b and FBS). Section 4.3 provides the results for mutation screening of selected IMDs namely the MMA mut-type and the MMA with homocystinuria cblC-type. Section 4.4 provides the results obtained from experimental works conducted to support the pathogenicity of all potentially novel/unreported mutations found in this study. For the purpose of describing mutations, a standard mutation nomenclature system as suggested by the HGVS (https://varnomen.hgvs.org/) is used. Nucleotide and amino acid numbering for the *AGL* gene follows the system that had been described by Bao et al. (1997). An earlier nucleotide reference numbering system was originally proposed by Yang et al. in 1992, but the most current literature has adopted the newer numbering system.

4.1 Mutation Screening Results of the *AGL* Gene (GSD III)

A total of 15 GSD III patients (and eight parents/siblings) were recruited for mutation screening of the *AGL* gene. The whole coding sequence of the *AGL* gene that spanned 35 exons were amplified in 21 fragments (some exons were amplified together as a single fragment due to short exons and/or short introns). Section 4.1.1 presents the results for PCR optimization and section 4.1.2 presents the results for mutation screening of the *AGL* gene.

4.1.1 PCR Optimization for *AGL* Gene Amplification

Twenty-one pairs of primers were used to amplify the 35 exons of the *AGL* gene. The primers were designed according to gene reference sequence (GenBank accession no. NG_012865.1) with minor modifications from previous reports (Goldstein et al., 2010; Lam et al., 2004). *In silico* analysis carried out on the modified primers showed that all primers passed the PCR suitability test for most pre-set criteria (Selected examples are shown in Appendix D).

Optimization of the annealing temperature was carried out for all primer pairs. The temperature that yielded fragment with correct expected size was chosen as the optimized annealing temperature (T_A). Different primer pairs were found to have T_A value ranging from 47 ºC to 65 ºC. Amplification with primer pairs that yielded multiple unspecific bands as well as desired fragment were further subjected to $MgCl_2$ titration PCR. Magnesium concentration ($[Mg^{2+}]$) that produced a single band at the expected size was chosen as the optimized concentration. Different primer pairs were found to have optimized $[Mg^{2+}]$ of either 1.0 mM or 1.5 mM. A summary of the PCR conditions for the amplification of all 21 fragments representing the entire *AGL* gene is presented in Table 4.1.

Table 4.1: Primers and optimized PCR condition for *AGL* gene amplification.

Fragment	PCR primer	Exon	Expected size (bp)	Annealing temp. (˚C)	$[Mg^{2+}]$ (mM)
1	1-3 F 2 R	1, 2	761	65	1.5
2	Seq3 1-3 R	3	324	55	1.0
3	4a F 4a R	4A	395	55	1.0
4	4-5 F 4-5 R	4, 5	1546	65	1.5
5	6 F 6 R	6	505	55	1.5

Table 4.1, continued.

Fragment	PCR primer	Exon	Expected size (bp)	Annealing temp. (˚C)	$[Mg^{2+}]$ (mM)
6	7-8 F 7-8 R	7, 8	1368	60	1.5
7	9-11 F 9-11 R	9, 10, 11	1289	63	1.5
8	12 F 12 R	12	402	57	1.5
9	13 F 13 R	13	410	55	1.5
10	14 F 14 R	14	458	55	1.5
11	15-18 F 15-18 R	15, 16, 17, 18	1429	57	1.5
12	19-21 F 19-21 R	19, 20, 21	1154	55	1.5
11	15-18 F 15-18 R	15, 16, 17, 18	1429	57	1.5
12	19-21 F 19-21 R	19, 20, 21	1154	55	1.5
13	22 F 22 R	22	600	55	1.5
14	23-25 F 23-25 R	23, 24, 25	1669	60	1.5
15	26 F 26 R	26	517	55	1.5
16	27 F 27 R	27	477	55	1.5
17	Seq28 E29R	28	255	63	1.5
18	29 F 29 R	29	382	63	1.5
19	30 F 31 R	30, 31	1490	47	1.5
20	32 F 33-34 R	32, 33, 34	1531	63	1.5
21	35 F 35 R	35	334	55	1.5

Results of successful amplification for each fragment are shown as agarose gel electrophoresis images in Figure 4.1 below.

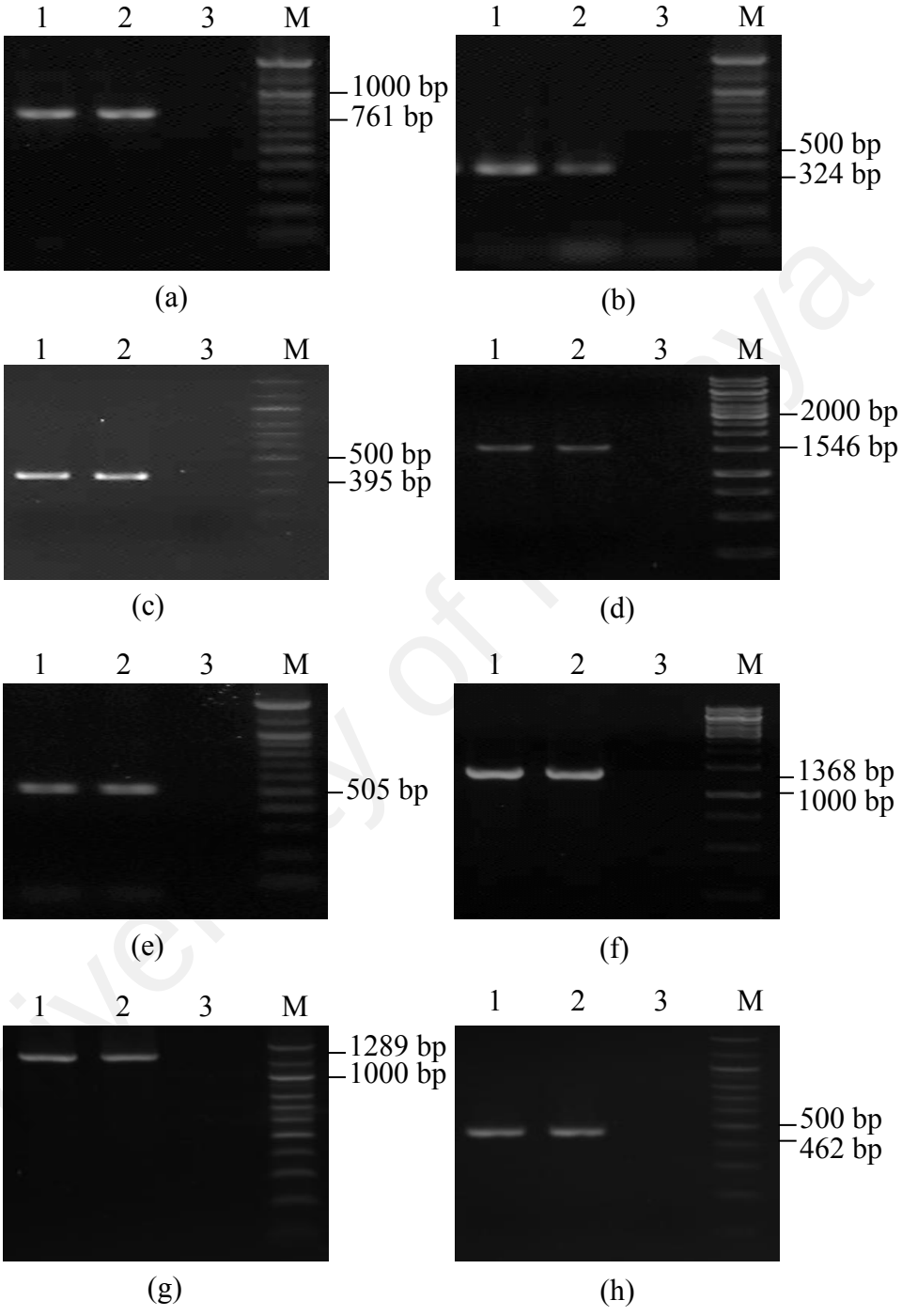


Figure 4.1: Example of agarose gel electrophoresis for *AGL* gene amplification. (a) Fragment 1, 761 bp (b) Fragment 2, 324 bp (c) Fragment 3, 395 bp (d) Fragment 4, 1546 bp (e) Fragment 5, 505 bp (f) Fragment 6, 1368 bp (g) Fragment 7, 1289 bp (h) Fragment 8, 462 bp. Lanes 1-2: PCR product from patients' sample, Lane 3: Negative control (no DNA template), Lane M: Molecular weight marker (100 bp/1 kb).

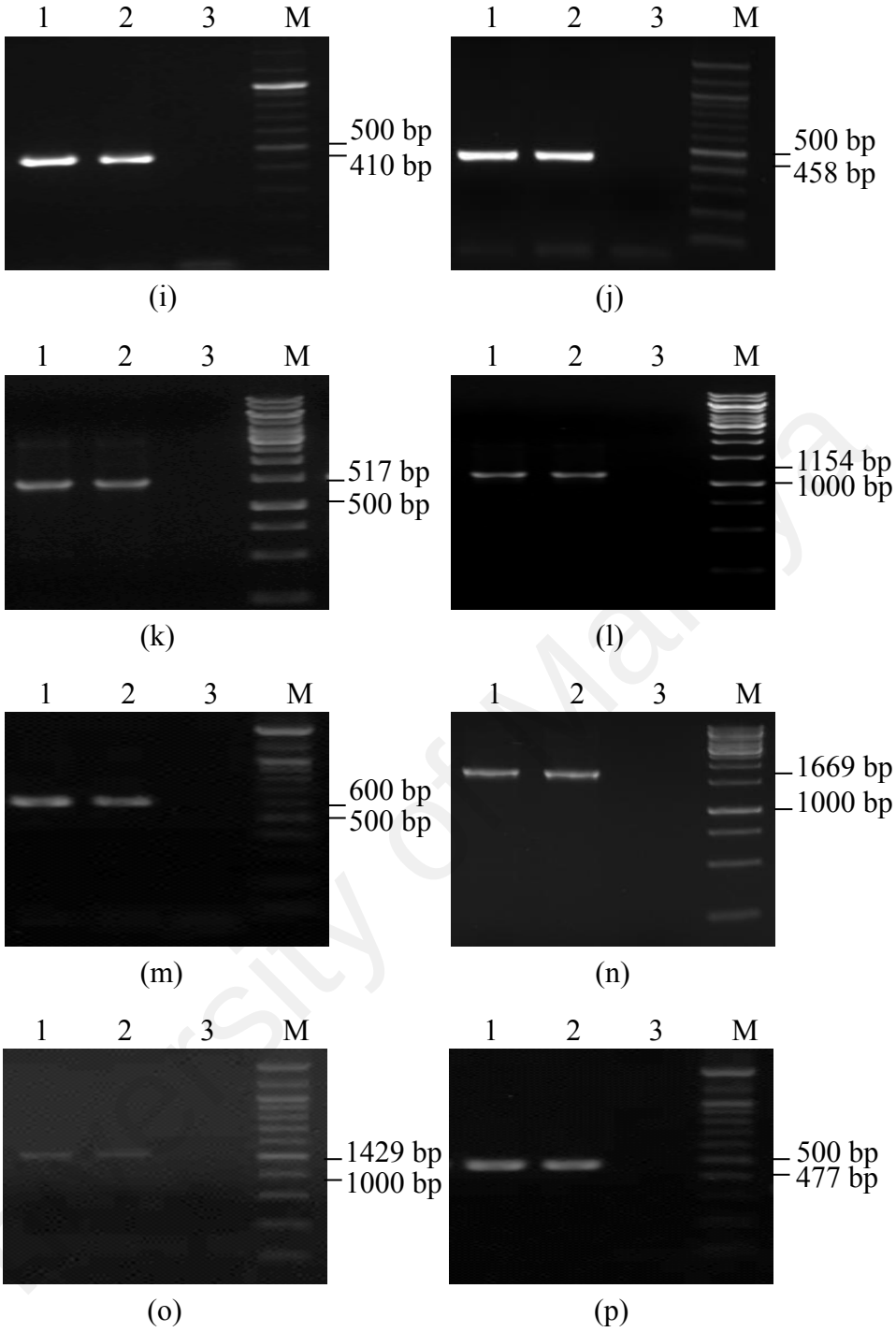


Figure 4.1, continued: Example of agarose gel electrophoresis for *AGL* gene amplification. (i) Fragment 9, 410 bp (j) Fragment 10, 458 bp (k) Fragment 11, 1429 bp (l) Fragment 12, 1154 bp (m) Fragment 13, 600 bp (n) Fragment 14, 1669 bp (o) Fragment 15, 517 bp (p) Fragment 16, 477 bp. Lanes 1-2: PCR product from patients' sample, Lane 3: Negative control (no DNA template), Lane M: Molecular weight marker (100 bp/1 kb ladder).

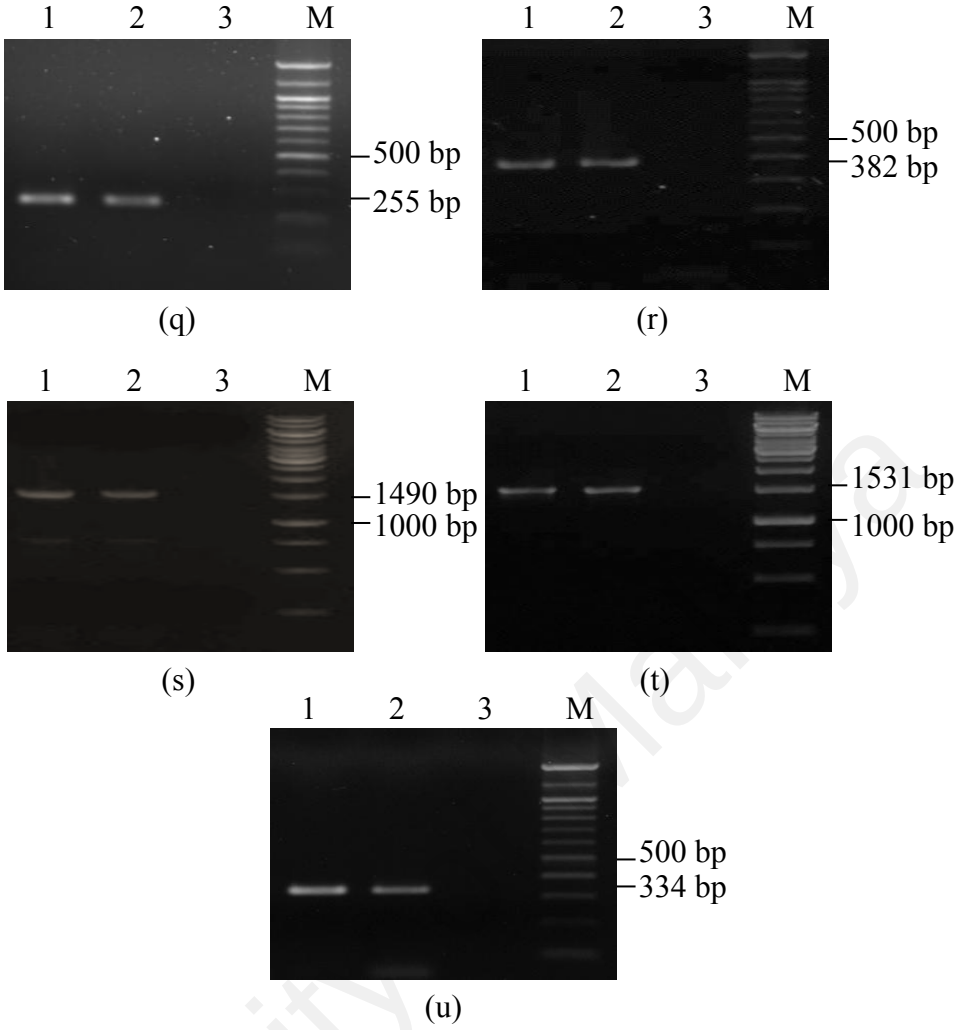


Figure 4.1, continued: Example of agarose gel electrophoresis for *AGL* gene amplification. (q) Fragment 17, 255 bp (r) Fragment 18, 382 bp (s) Fragment 19, 1490 bp (t) Fragment 20, 1531 bp (u) Fragment 21, 334 bp. Lanes 1-2: PCR product from patients' sample, Lane 3: Negative control (no DNA template), Lane M: Molecular weight marker (100 bp/1 kb ladder).

4.1.2 Mutations Found in the *AGL* Gene

Purified PCR products were subjected to sequencing analysis using the same primers used for PCR. When amplification specificity problems were encountered, nested sequencing primers were designed (listed in Appendix B). Sequencing chromatograms were analysed and edited (to remove ambiguity, where appropriate) using the Chromas sequence analysis software. Sequences generated from the sequencing reactions were compared with the *AGL* reference sequence. As previously mentioned, any variation detected was compared against mutation databases to confirm

their involvement in the pathogenesis of GSD III in these patients. In cases where no match was found, the probable involvement of suspected pathogenic mutations is supported by additional analyses (e.g *in silico* prediction, population distribution and/or functional analysis as described in Section 4.4). A summary of the mutations found were listed in Table 4.2 and the sequence analysis result for each fragment is presented below: -

i) Fragment 1

Exon 1 and exon 2 were amplified as a single fragment with an expected size of 761 bp. No sequence variations were found in both exons for all patients.

ii) Fragment 2

Exon 3 was amplified in fragment 2 with an expected size of 324 bp. Sequencing analysis revealed one sequence variation within the intronic region of exon 3 (Figure 4.2). This variation, c.83-33C>T was detected in patient 2, 3, 8 and 10. However, c.83-33C>T has been reported to be polymorphism in the mutation database (rs2307129).

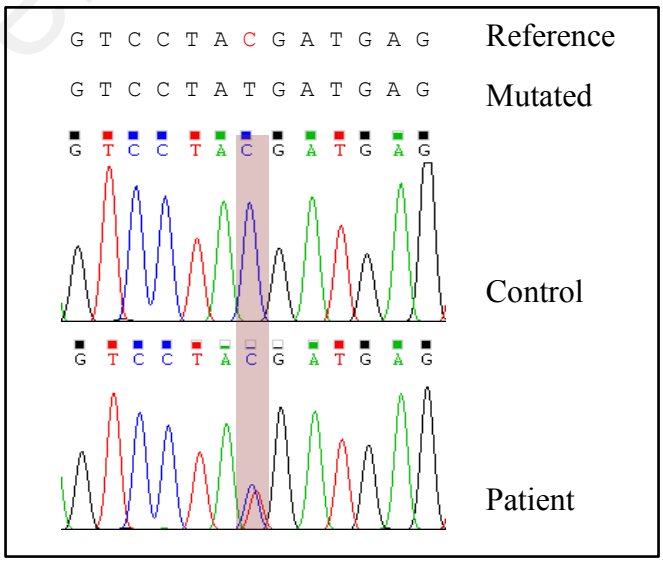


Figure 4.2: Partial sequencing chromatograms of intron 3 showing a C>T substitution for mutation c.83-33C>T. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

iii) Fragment 3

Exon 4a (alternative variant) was amplified in fragment 3 with an expected size of 395 bp. No sequence variation was found in exon 4a for all patients.

iv) Fragment 4

Exon 4 and exon 5 were amplified as a single fragment 4 with an expected size of 1546 bp. Sequence analysis revealed one sequence variation within exon 4 of Patient 8. He was found to be homozygous for a c.99C>T mutation (Figure 4.3). The mutation changes an arginine to a stop codon (p.R34X) and has been previously reported to be pathogenic of GSD III (Lucchiari et al., 2002a).

v) Fragment 5

Exon 6 was amplified in fragment 5 with an expected size of 505 bp. No sequence variation was found in exon 6 for all patients.

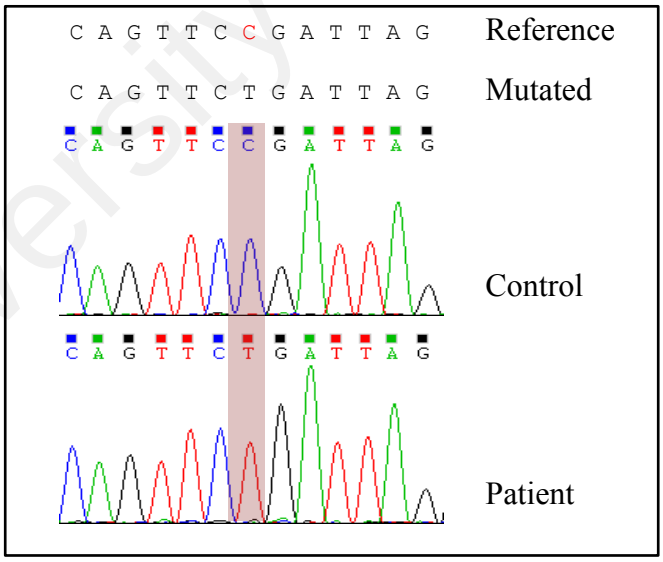


Figure 4.3: Partial sequencing chromatograms of exon 4 showing a C>T substitution for mutation c.99C>T. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

vi) Fragment 6

Exon 7 and exon 8 were amplified as a single fragment 6 with an expected size of 1368 bp. Sequence analysis revealed two sequence variations within this fragment, both are substitution mutations.

- Mutation c.894C>T (Figure 4.4a) was detected in exon 8 of Patient 8 and 10. Both patients are found to be homozygous for the mutation. c.894C>T is a silent mutation that maintains a leucine residue at codon 298 (p.L298L) and has been previously reported to be a polymorphism (rs2230306).
- Mutation c.959-18G>A (Figure 4.4b) was detected in intron 8 region of Patient 1, 2, 3, 4 and 6. All affected patients are found to be heterozygous for the mutation except Patient 4. This mutation occurs at the non-coding region and has been previously reported to be a polymorphism (rs634880).

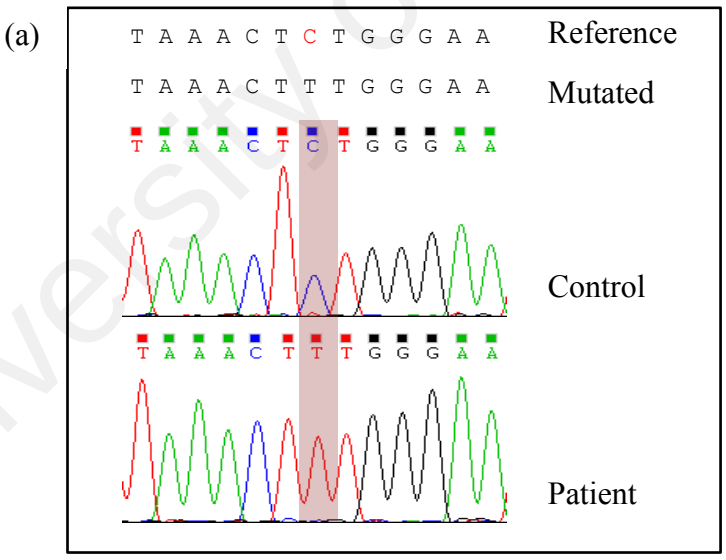


Figure 4.4: Partial sequencing chromatograms for fragment 6 of the *AGL* gene. The chromatograms of (a) exon 8 showing a C>T substitution for mutation c.894C>T. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

(b)

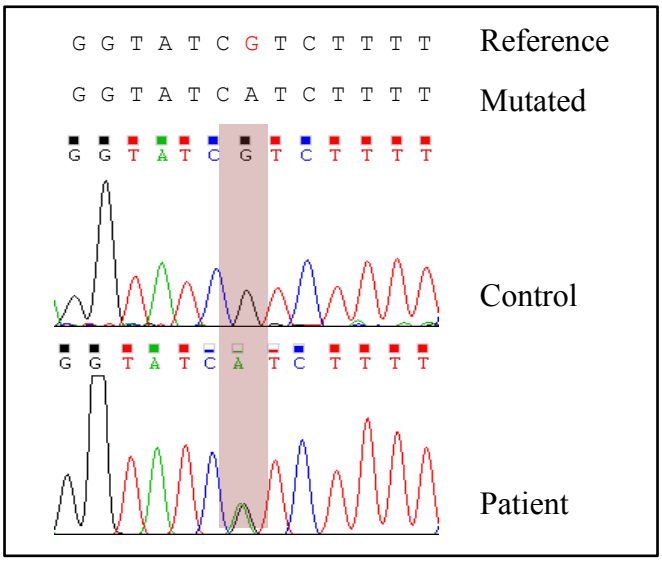


Figure 4.4, continued**:** Partial sequencing chromatograms for fragment 6 of the *AGL* gene. The chromatograms of (b) intron 8 showing a G>A substitution for mutation c.959-18G>A. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

vii) Fragment 7

Exons 9, 10 and 11 were amplified as a 1289 bp fragment. Sequence analysis on the fragment found no sequence variation at the exonic regions however a single sequence variation was found in intron 10. The variation, a c.1185+15T>C mutation (Figure 4.5) was detected in Patient 1 and 4. Patient 1 was found to be heterozygous for the mutation while patient 7 was a homozygote. This mutation has not been previously reported. However, it occurs at the non-coding region of the *AGL* gene.

viii) Fragment 8

Exon 12 was amplified in fragment 8 with an expected size of 462 bp. Sequence analysis revealed no sequence variation at the exonic region however a single sequence variation was found at the exon-intron 12 border. The variation, a c.1423+G>T mutation (Figure 4.6) was detected in heterozygous form in Patient 7. This mutation has not been previously reported as pathogenic and would require further analysis to confirm its role in the pathogenesis of GSD III in this patient (see section 4.4).

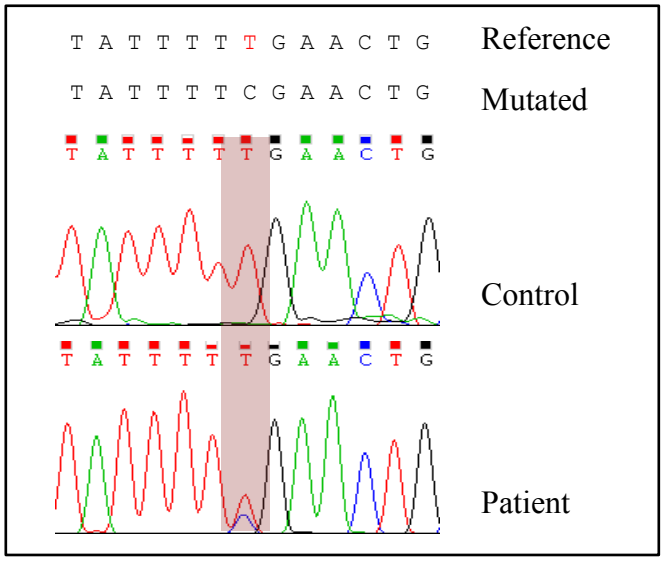


Figure 4.5: Partial sequencing chromatograms of intron 10 sequence showing a T>C substitution for mutation c.1185+15T>C. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

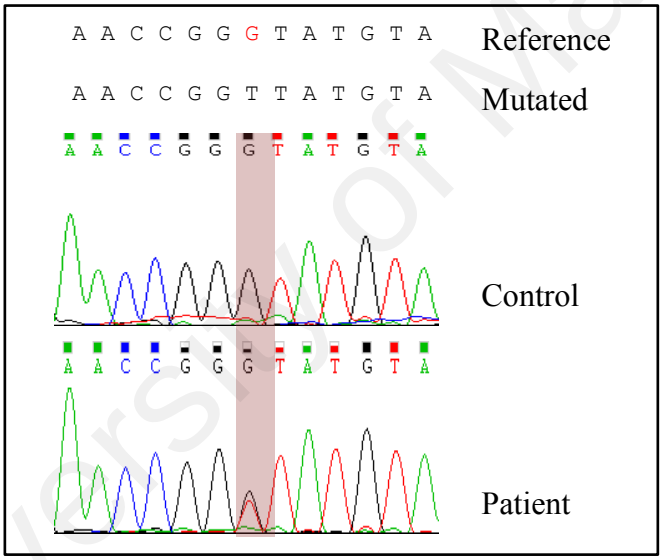


Figure 4.6: Partial sequencing chromatograms of exon-intron 12 showing a G>T substitution for mutation c.1423+1G>T. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

ix) Fragment 9

Exon 13 was amplified in fragment 9 with an expected size of 410 bp. No sequence variation was detected in all patients.

x) Fragment 10

Exon 14 was amplified in fragment 10 with an expected size of 458 bp. No sequence variation was detected in all patients.

xi) Fragment 11

Exons 15, 16, 17 and 18 were amplified as a single fragment 11 with an expected size of 1429 bp. Sequence analysis revealed two sequence variations within this fragment, one being a nonsense mutation and the other is a substitution mutation.

- Mutation c.1783C>T (Figure 4.7a) was detected in exon 15 of Patient 11. He was found to be homozygous for the mutation. Further analysis on both the mother and the father samples revealed that they are carriers for the mutation. Mutation c.1783C>T changes an arginine to a stop codon (p.R595X) and has been previously reported to be pathogenic of GSD III (Aoyama et al., 2009).
- Mutation c.2001+8T>C (Figure 4.7b) was detected in intron 16 of Patient 1, 2, 3, 4, 7 and 11. Patient 1, 7 and 11 were found to be homozygous for the mutation. Patient 2, 3, and 4 were heterozygous for the mutation. This mutation occurs at the non-coding region of the *AGL* gene and has been previously reported to be a polymorphism (rs3736296).

xii) Fragment 12

Exon 19, 20 and 21 were amplified as a single fragment 12 with an expected size of 1154 bp. Sequence analysis revealed one variation at the exon-intron 21 border. The variation, a c.2681+1G>A mutation (Figure 4.8) was detected in heterozygous form in Patient 1, 6 and 7. Patient 5 and 9 were found to be homozygous for the mutation. Mother and father samples for Patient 1 and 7 were available for sequencing. Mother of both Patient 1 and 7 were found to be the carrier for the mutation. This mutation has been previously reported to be pathogenic as it causes skipping of exon 21 (Hadjigeorgiou et al., 1999).

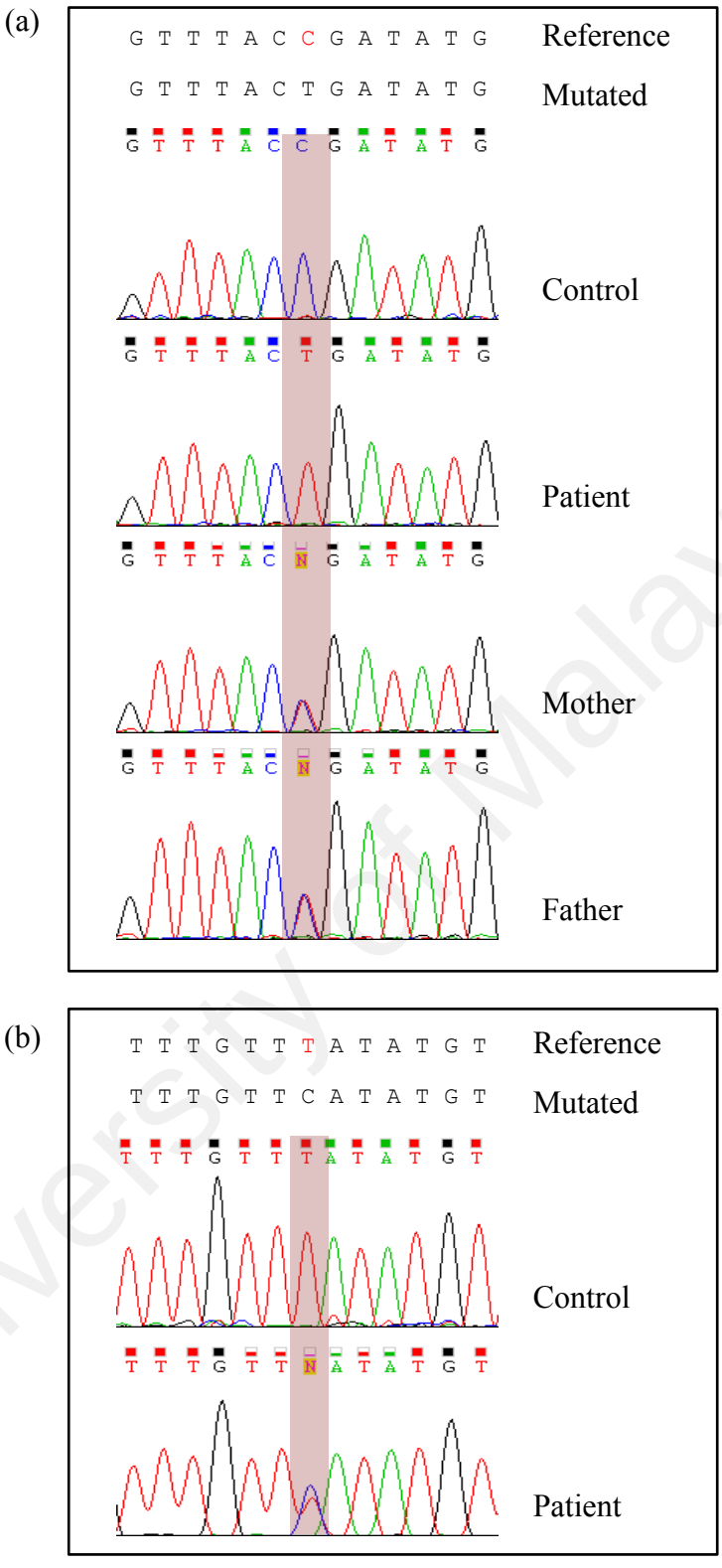


Figure 4.7: Partial sequencing chromatograms for fragment 11 of the *AGL* gene. The chromatograms of (a) exon 15 showing a C>T substitution for mutation c.1783C>T (b) intron 16 showing a T>C substitution for mutation c.2001+8T>C. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

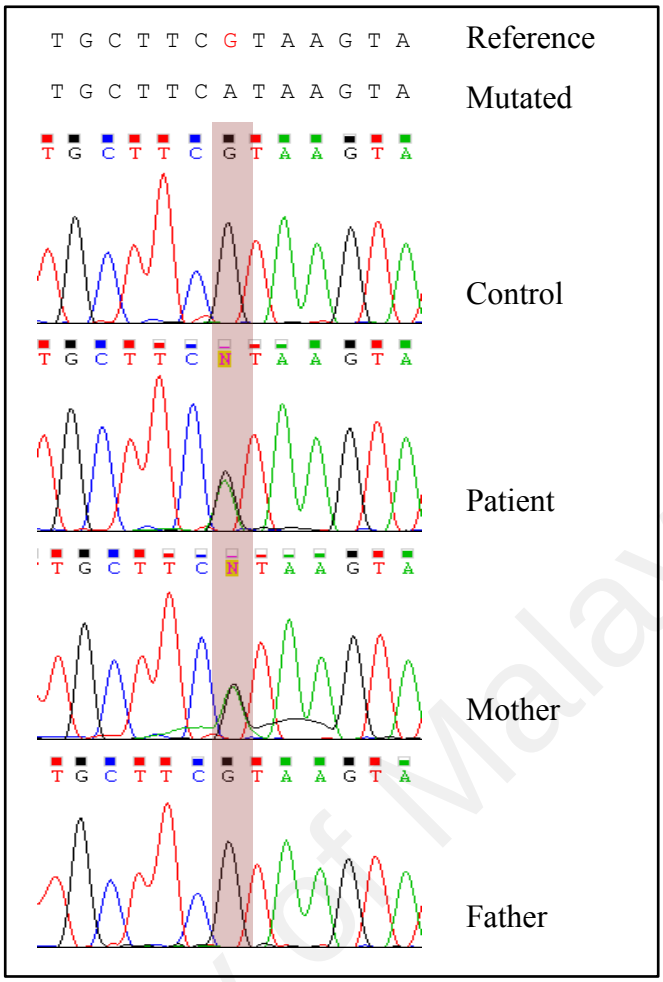


Figure 4.8: Partial sequencing chromatograms of exon-intron 21 showing a G>A substitution for mutation c.2681+1G>A. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

xiii) Fragment 13

Exon 22 was amplified in fragment 13 with an expected size of 600 bp. Sequence analysis revealed no mutation within the exonic region however a single mutation was detected in intron 22. The c.2812+11G>A mutation (Figure 4.9) was found in Patient 1, 2, 3, 7, 10 and 11. Patient 1, 7, 10 and 11 were found to be heterozygous for the mutation. Patient 2 and 3 were homozygous for the mutation. This mutation that occurs at the non-coding region of the *AGL* gene has been previously reported to be a polymorphism (rs555929).

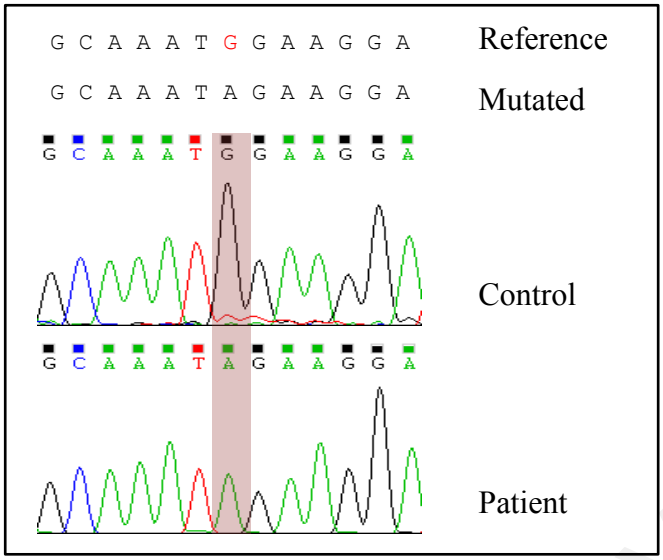


Figure 4.9: Partial sequencing chromatograms of intron 22 showing a G>A substitution for mutation c.2812+11G>A. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

xiv) Fragment 14

Exons 23, 24 and 25 were amplified as a single fragment 14 with an expected size of 1669 bp. Sequence analysis revealed four sequence variations, which includes one deletion and three substitutions.

- Mutation c.2914_2915delAA (Figure 4.10a) was detected in exon 23 of Patient 1. He was found to be heterozygous for the mutation. Mother and father samples were available for sequencing and analysis revealed the father as the carrier for mutation c.2914_2915delAA. This mutation has not been previously reported as pathogenic. However, preliminary analysis suggested that this mutation will lead to a frameshift, and subsequently generate premature stop codon 10 amino acids downstream (See section 4.4 for detailed explanation).
- Mutation c.2950-21T>A (Figure 4.10b) was detected in intron 23 of Patient 1, 2, 3, and 7. Patient 1 and 7 were found to be heterozygous for the mutation. Patient 2 and 3 were homozygous for the mutation. This mutation

occurs at the non-coding region of the *AGL* gene and has been previously reported to be a polymorphism (rs2035961).

- Mutation c.3199C>T (Figure 4.10c) was detected in exon 25 of Patient 6 and 10. Both patients were found to be heterozygous for the mutation. This mutation occurs at the non-coding region of the *AGL* gene and has been previously reported to be a polymorphism (rs3753494).
- Mutation c.3259+37G>A (Figure 4.10d) was detected in intron 25 of Patient 2, 3, 4, 7 and 11. All patients were found to be homozygous for the mutation. This mutation occurs at the non-coding region of the *AGL* gene and has been previously reported to be a polymorphism (rs594249).

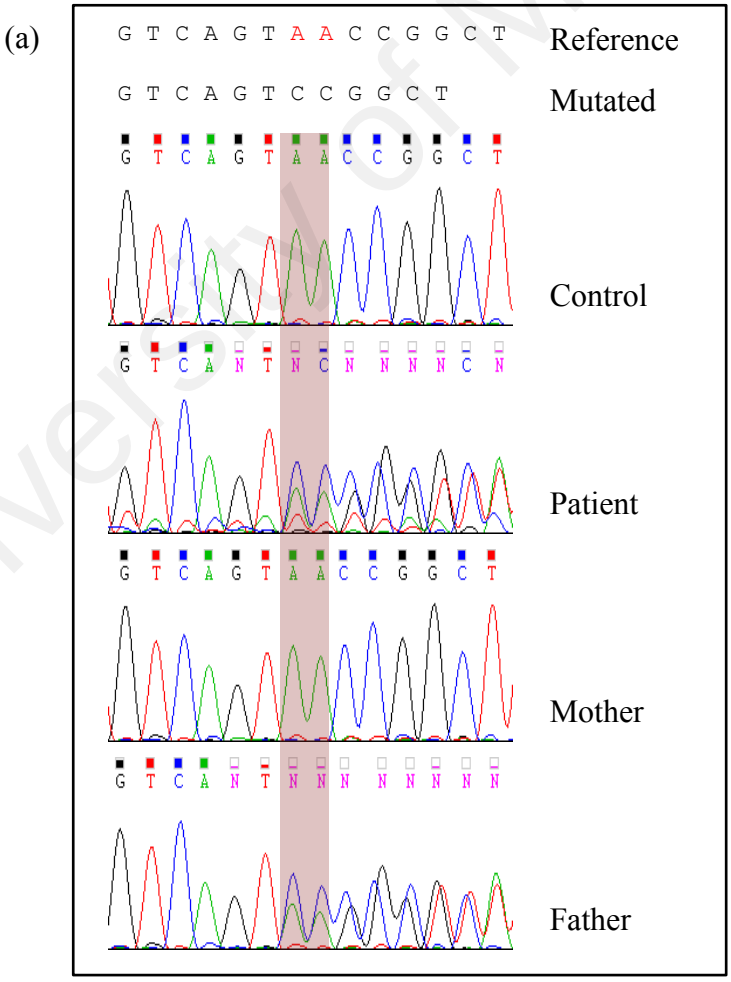


Figure 4.10: Partial sequencing chromatograms showing mutations found within fragment 14 of the *AGL* gene. The chromatograms showing a (a) 2-bp deletion for mutation c.2914_2915delAA. Red letters in the reference sequence and the coloured box highlight the deleted nucleotides.

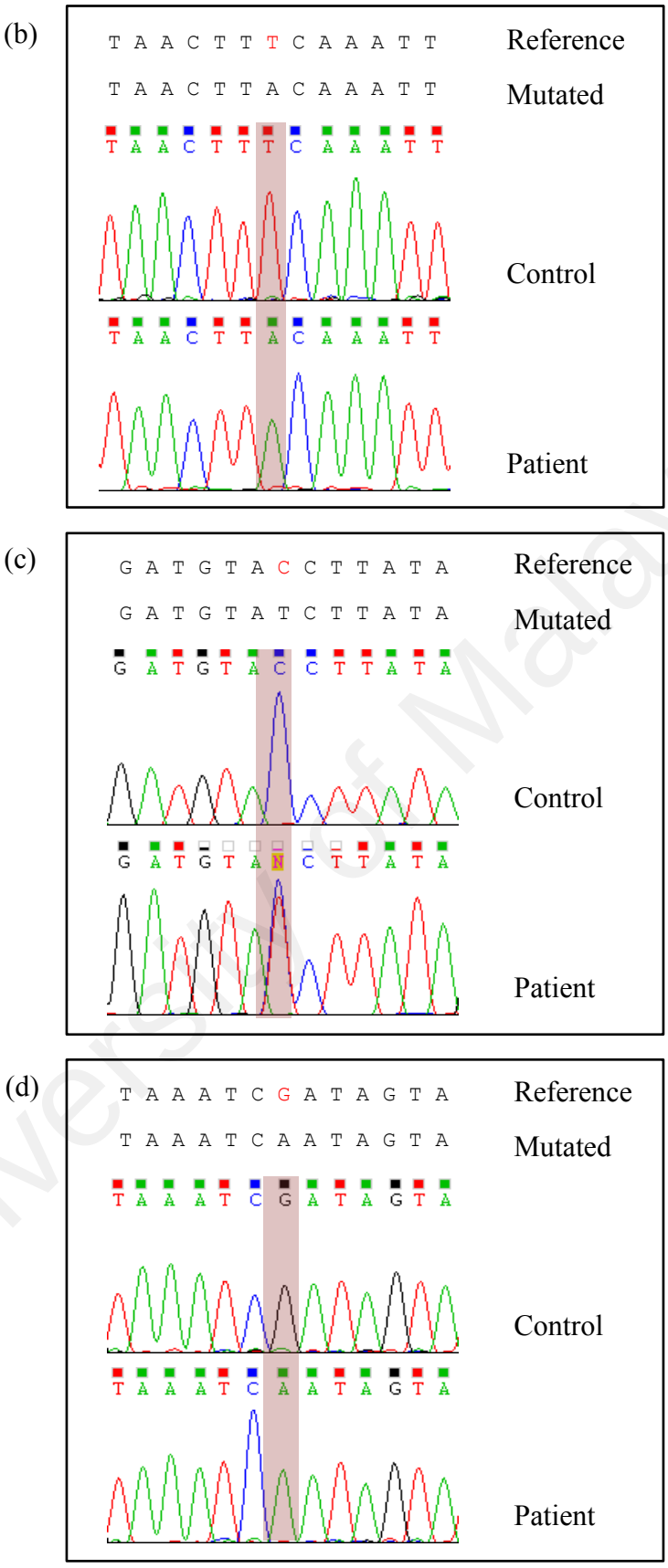


Figure 4.10, continued**:** Partial sequencing chromatograms showing mutations found within fragment 14 of the *AGL* gene. The chromatograms showing a (b) T>A substitution for mutation c.2950-21T>A (c) C>T substitution for mutation c.3199C>T (d) G>A substitution for mutation c.3259+37G>A. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

xv) Fragment 15

Exon 26 was amplified in fragment 15 with an expected size of 517 bp. Sequence analysis revealed a single mutation that was found in Patient 2, 3 and 7. All patients were found to be homozygous for the mutation c.3343G>A (Figure 4.11). This mutation that changes a glycine residue with arginine (p.G1115R) has been previously reported to be polymorphism (rs7536086).

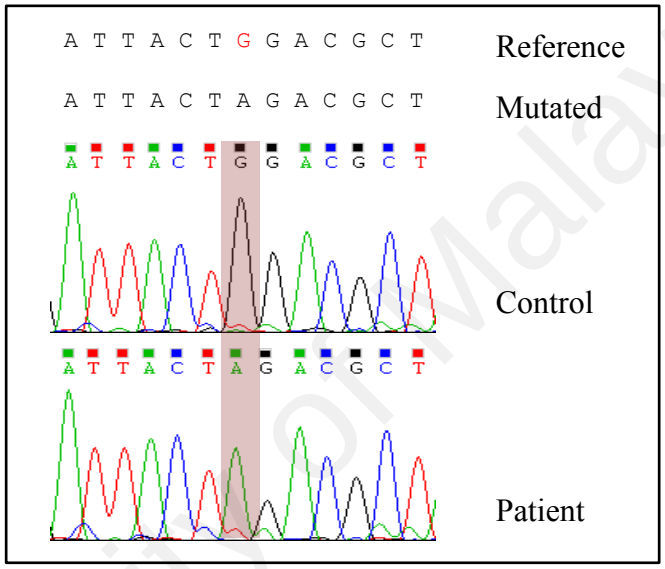


Figure 4.11: Partial sequencing chromatograms of exon 26 showing a G>A substitution for mutation c.3343G>A. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

xvi) Fragment 16

Exon 27 was amplified in fragment 16 with an expected size of 477 bp. No sequence variation was found in the exon for all patients.

xvii) Fragment 17

Exon 28 was amplified in fragment 17 with an expected size of 255 bp. No sequence variation was found in the exon for all patients.

xviii) Fragment 18

Exon 29 was amplified in fragment 18 with an expected size of 382 bp. Sequence analysis revealed three different mutations within this fragment which includes one deletion mutation and two substitution mutations.

- Mutation, c.3814_3815delAG (Figure 4.12a) was detected in exon 29 of Patient 6 and 10. Patient 6 was found to be heterozygous while Patient 10 was homozygous for the mutation. This mutation has not been previously reported as pathogenic. However, preliminary analysis suggested that this mutation will lead to a frameshift and subsequently generate premature stop codon 18 amino acids downstream (see section 4.4 for detailed explanation).
- Mutation c.3836+45G>A (Figure 4.12b) was detected in intron 29 region of Patient 7. She was found to be heterozygous for the mutation. This mutation occurs at the non-coding region of the *AGL* gene and has been previously reported to be a polymorphism (rs2274570).
- Mutation c.3836+53T>A (Figure 4.12c) was detected in intron 29 of Patient 2, 3 and 7. All patients were found to be heterozygous for the mutation. This mutation has not been previously reported. However, it occurs at the non-coding region of the *AGL* gene and therefore not expected to be pathogenic.

xix) Fragment 19

Exons 30 and 31 were amplified as a single fragment 19 with an expected size of 1490 bp. No sequence variations were found in both exons for all patients.

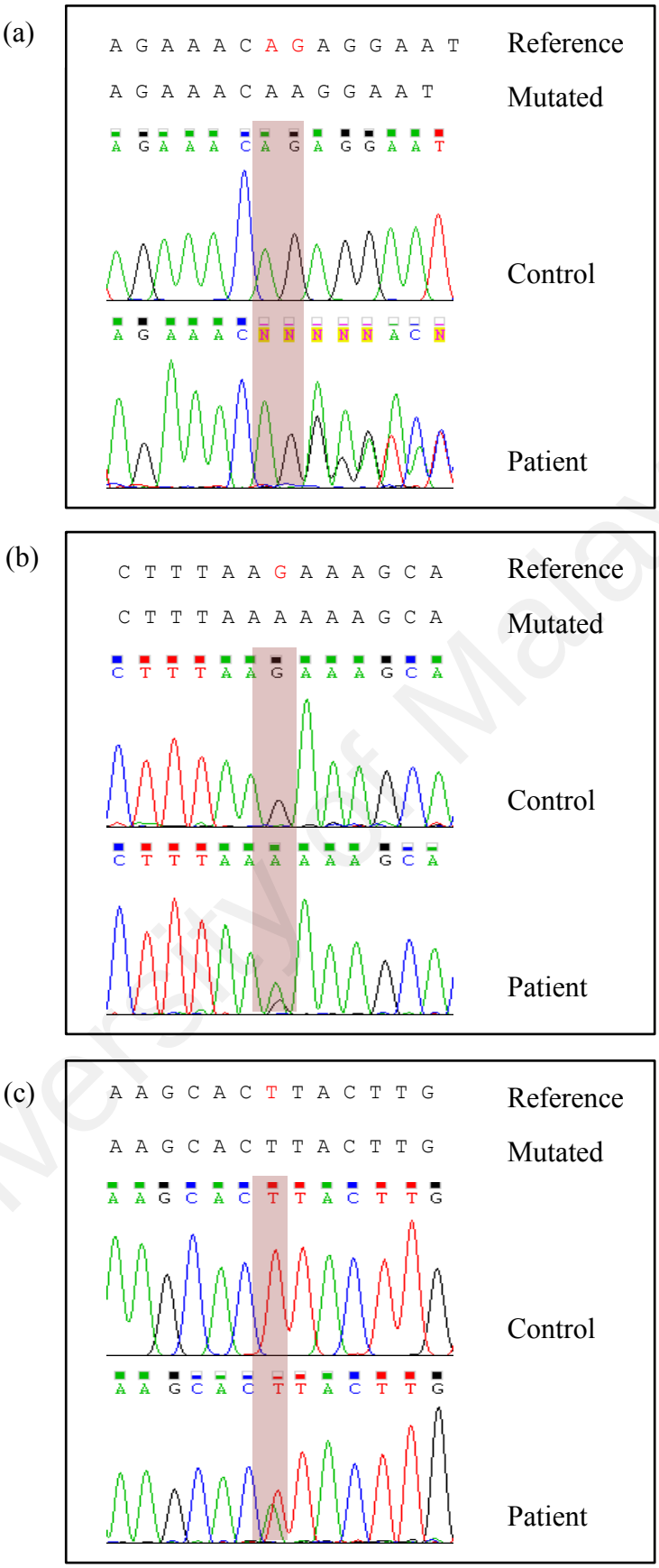


Figure 4.12: Partial sequencing chromatograms showing mutations found within fragment 18 of the *AGL* gene. The chromatograms showing a (a) 2 bp deletion for mutation c.3814_3815delAG (b) G>A substitution for mutation c.3836+45G>A (c) T>A substitution for mutation c.3836+53T>A. Red letter in the reference sequence and the coloured box highlights the affected nucleotides.

xx) Fragment 20

Exons 32, 33 and 34 were amplified as a single fragment 20 with an expected size of 1531 bp. Sequence analysis revealed a single c.4333T>G mutation (Figure 4.13) that was detected in exon 33 of Patient 2 and 3 (siblings). Both patients were found to be heterozygous for the mutation. Mother and father samples were available for sequencing and analysis revealed that the mother is the carrier. This mutation has not been previously reported as pathogenic. However, the mutation changes a tyrosine residue with an aspartic acid at codon 1445 (p.Y1445D), and would require further analysis to confirm its role in the pathogenesis of GSD III in this patient (see section 4.4).

xxi) Fragment 21

Exon 35 was amplified in fragment 21 with an expected size of 334 bp. Sequence analysis revealed two different sequence variations which include a nonsense mutation and a deletion mutation.

- Mutation c.4490G>A (Figure 4.14a) was detected in exon 35 of Patient 4. He was found to be homozygous for the mutation. This mutation has not been previously reported as pathogenic. However, this mutation will cause tryptophan replacement with a stop codon (p.W1497X) which most likely to be pathogenic.
- Mutation c.4531_4534delTGTC (Figure 4.14b) was detected in Patient 2 and Patient 3 who are siblings. They are both found to be heterozygous for the mutation. Mother and father samples were available for sequencing and analysis revealed the mother as the carrier. This mutation has not been previously reported to be pathogenic. However, preliminary analysis suggested a translational readthrough may occur and would require further

analysis to confirm its role in the pathogenesis of GSD III in the patient (See section 4.4).

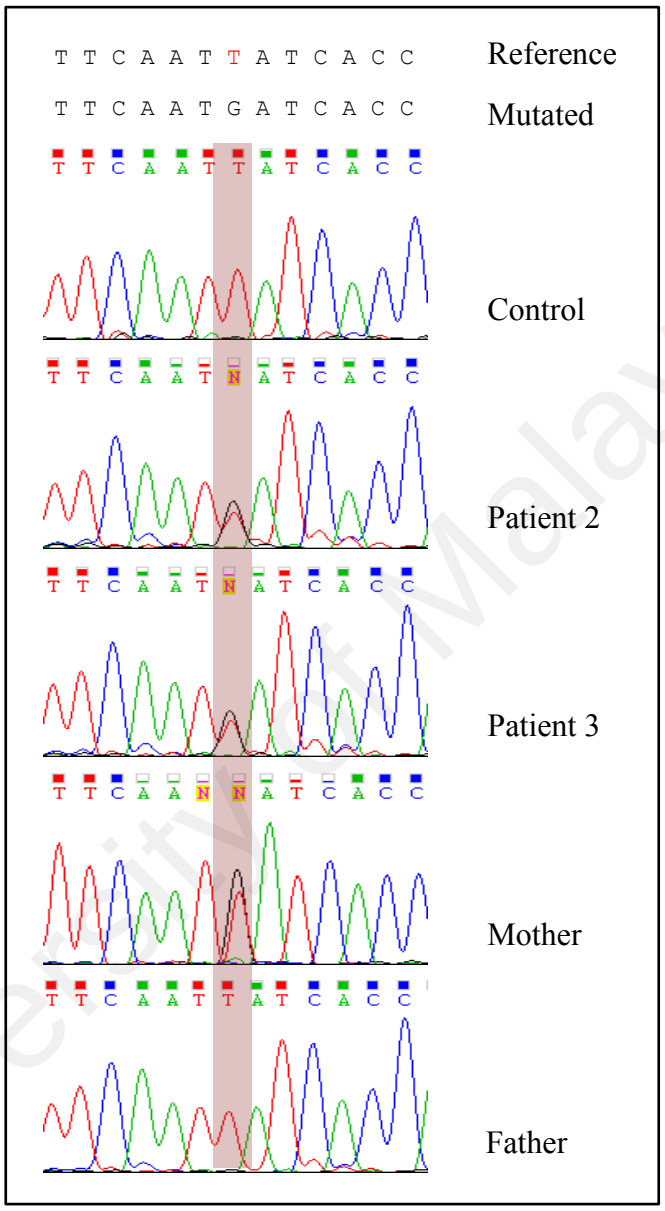


Figure 4.13: Partial sequencing chromatograms of fragment 20 showing a T>G substitution for mutation c.4333T>G. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

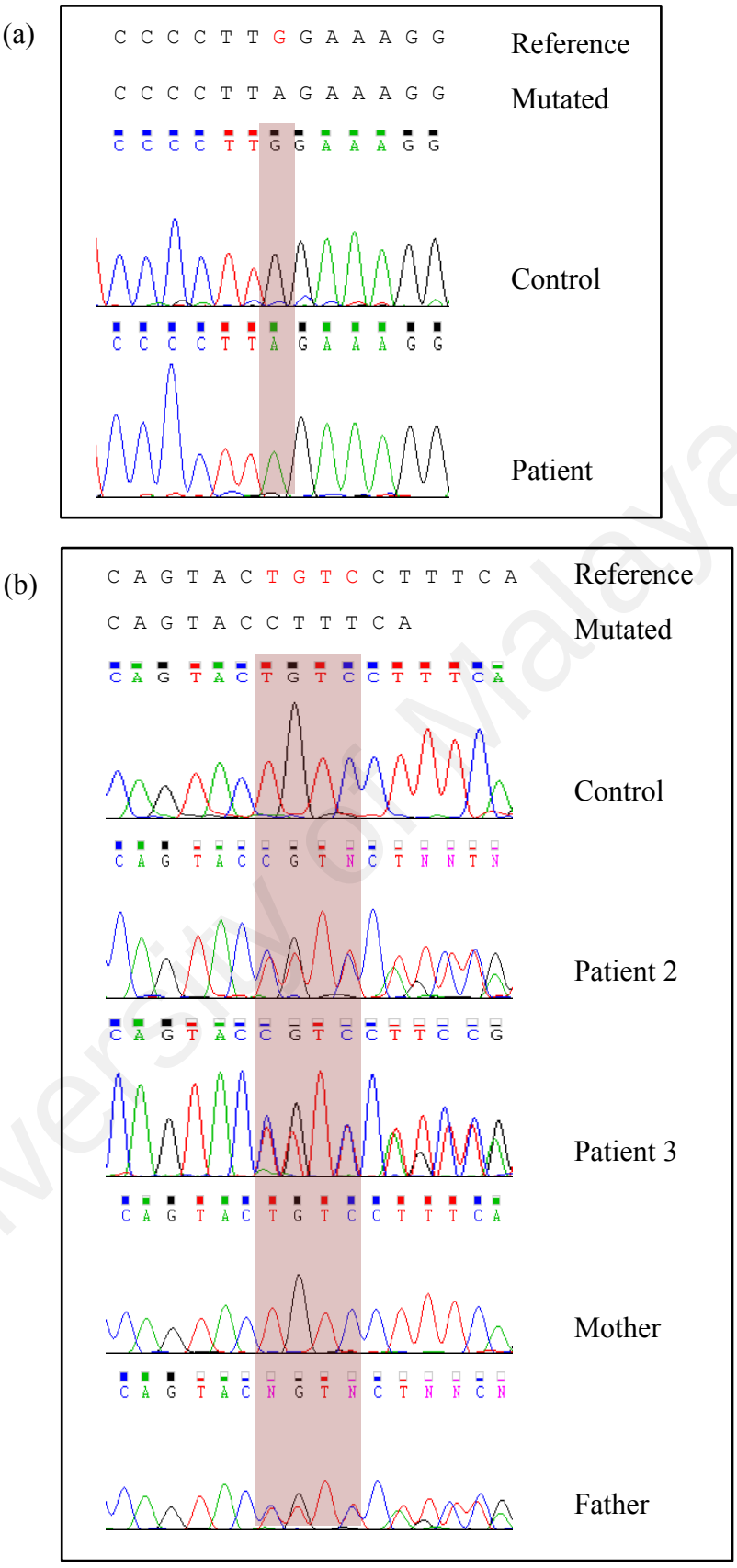


Figure 4.14: Partial sequencing chromatograms showing mutations found in fragment 21 of the *AGL* gene. The chromatograms showing a (a) G>A substitution for mutation c.4490G>A (b) 4-bp deletion for mutation c.4531_4534delTGTC. Red letters in the reference sequence and the coloured box highlight the affected nucleotides.

4.1.3 Unreported Polymorphism

In the course of this study, two unreported variations have been identified in which their involvements in the pathogenesis of GSD III in these patients are thought to be unlikely. This conclusion was made due to three justifications, as stated below:

i. The mutations occurred outside the exonic region of the *AGL* gene

ii. The other mutations found in the patient has been reported as pathogenic or

iii. The other mutations found predicted to be pathogenic due to its expected effect (nonsense, missense, frameshift or splicing mutation).

Table 4.2: Genotypes of fifteen Malaysian GSD III patients.

Patient	Race	Source	Genotype	*AGL* mutation	Effect	Exon	Intron	Reference
1[a]	Malay	UMMC	Heterozygous	c.2681+1G>A	Exon skipping		21	Hadjigeorgiou et al., 1999
				c.2914_2915delAA	p.N972Pfs*10	23		Unreported
				# c.959 -18G>A	Non pathogenic		8	rs634880
				c.1185 +15T>C	Non pathogenic		10	Unreported
				c.2001 +8T>C	Non pathogenic		16	rs3736296
				c.2812 +11G>A	Non pathogenic		22	rs555929
				c.2950-21T>A	Non pathogenic		23	rs2035961
2*, [a]	Chinese	UMMC	Heterozygous	c.4333T>G	p.Y1445D	33		Unreported
				c.4531_4534delTGTC	p.C1511Lfs*35	35		Unreported
				# c.83 -33C>T	Non pathogenic		3	rs2307129
				c.959 -18G>A	Non pathogenic		8	rs634880
				c.2001 +8T>C	Non pathogenic		16	rs3736296
				c.2812 +11G>A	Non pathogenic		22	rs555929
				c.2950-21T>A	Non pathogenic		23	rs2035961
				c.3259 +37G>A	Non pathogenic		25	rs594249
3*, [a]	Chinese	UMMC	Heterozygous	c.4333T>G	p.Y1445D	33		Unreported
				c.4531_4534delTGTC	p.C1511Lfs*35	35		Unreported
				# c.83 -33C>T	Non pathogenic		3	rs2307129
				c.959 -18G>A	Non pathogenic		8	rs634880
				c.2001 +8T>C	Non pathogenic		16	rs3736296
				c.2812 +11G>A	Non pathogenic		22	rs555929
				c.2950-21T>A	Non pathogenic		23	rs2035961
				c.3259 +37G>A	Non pathogenic		25	rs594249

Table 4.2, continued.

Patient	Race	Source	Genotype	*AGL* mutation	Effect	Exon	Intron	Remark
4	Malay	UMMC	Homozygous	c.4490G>A	p.W1497X	35		Unreported
				c.4490G>A	p.W1497X	35		Unreported
				# c.959 -18G>A	Non pathogenic		8	rs634880
				c.1185 +15T>C	Non pathogenic		10	Unreported
				c.2001 +8T>C	Non pathogenic		16	rs3736296
				c.3259 +37G>A	Non pathogenic		25	rs594249
5	Malay	HKL	Homozygous	c.2681+1G>A	Exon skipping		21	Hadjigeorgiou et al., 1999
				c.2681+1G>A	Exon skipping		21	Hadjigeorgiou et al., 1999
6	Malay	HKL	Heterozygous	c.2681+1G>A	Exon skipping		21	Hadjigeorgiou et al., 1999
				c.3814_3815delAG	p.R1272Rfs*18	29		Unreported
				# c.959 -18G>A	Non pathogenic		8	rs634880
				c.3199 C>T	Non pathogenic		25	rs3753494
7[a]	Malay	HKL	Heterozygous	c.1423+1G>T	Intron retention		12	Unreported
				c.2681+1G>A	Exon skipping		21	Hadjigeorgiou et al., 1999
				# c.2001 +8T>C	Non pathogenic		16	rs3736296
				c.2812 +11G>A	Non pathogenic		22	rs555929
				c.2950-21 T>A	Non pathogenic		23	rs2035961
				c.3259+37 G>A	Non pathogenic		25	rs594249
				c.3343 G>A	Non pathogenic		26	rs7536086
				c.3836+45 G>A	Non pathogenic		29	rs2274570
				c.3836+53T>A	Non pathogenic		29	Unreported

Table 4.2, continued.

Patient	Race	Source	Genotype	*AGL* mutation	Effect	Exon	Intron	Remark
8	Malay	HKL	Homozygous	c.99C>T	p.R34X	4		Lucchiari et al., 2002
				c.99C>T	p.R34X	4		Lucchiari et al., 2002
				# c.83-33 C>T	Non pathogenic		3	rs2307129
				c.894 C>T	Non pathogenic	8		rs2230306
9	Malay	HKL	Homozygous	c.2681+1G>A	Exon skipping		21	Hadjigeorgiou et al., 1999
				c.2681+1G>A	Exon skipping		21	Hadjigeorgiou et al., 1999
10	Malay	HKL	Homozygous	c.3814_3815delAG	p.R1272Rfs*18	29		Unreported
				c.3814_3815delAG	p.R1272Rfs*18	29		Unreported
				# c.83-33C>T	Non pathogenic		3	rs2307129
				c.894C>T	Non pathogenic	8		rs2230306
				c.2812+11G>A	Non pathogenic		22	rs555929
				c.3199C>T	Non pathogenic		25	rs3753494
11	Bajau	HKL	Homozygous	c.1783C>T	p.R595X	15		Aoyama et al., 2009
				c.1783C>T	p.R595X	15		Aoyama et al., 2009
				# c.2001 +8T>C	Non pathogenic		16	rs3736296
				c.2812+11G>A	Non pathogenic		22	rs555929
				c.3259+37 G>A	Non pathogenic		25	rs594249
12	Malay	HKL		No mutation				
13	Chinese	HKL		No mutation				
14	Malay	HKL		No mutation				
15	Malay	HKL		No mutation				

*Patient 2 and patient 3 are siblings

[a] Provided with parents' blood samples

Other variations identified, but they have been reported/ predicted to be polymorphisms (not pathogenic)

4.2 Mutation Screening Results of the *SLC37A4* Gene (GSD 1b) and the *SLC2A2* Gene (FBS)

Five GSD 1b patients and four FBS patients were recruited for mutation screening of the *SLC37A4* gene and the *SLC2A2* gene respectively. The results for PCR optimization and mutation screening of the *SLC37A4* gene are presented in section 4.2.1 and section 4.2.2 respectively. The results for PCR optimization and mutation screening of the *SLC2A2* gene are presented in section 4.2.3 and section 4.2.4 respectively.

4.2.1 PCR Optimization for *SLC37A4* Gene Amplification

Nine pairs of primers were used to amplify nine exons of the *SLC37A4* gene. The primers were designed according to gene reference sequence with GenBank accession no. NG_013331.1). *In silico* analysis carried out on the designed primers showed that all primers passed the PCR suitability test for most pre-set criteria (Selected examples are shown in Appendix E).

Optimization of the annealing temperature was carried out for all primer pairs. The temperature that yielded fragment with correct expected size was chosen as the optimized annealing temperature (T_A). Different primer pairs were found to have T_A value ranging from 55 ºC to 65 ºC. Subsequent optimization on the magnesium concentration [Mg^{2+}] found that magnesium concentration of 1.5 mM were suitable to generate fragment with expected size for all primer pairs. A summary of the PCR conditions for *SLC37A4* gene amplification is presented in Table 4.3.

Table 4.3: Primers and optimized PCR condition for *SLC37A4* gene amplification.

PCR primer	Exon	Expected size (bp)	Annealing temp. (ºC)	[Mg^{2+}] (mM)
G6PT-1F & G6PT-1R	1	232	62	1.5
G6PT-2F & G6PT-2R	2	337	62	1.5
G6PT-3F & G6PT-3R	3	341	65	1.5
G6PT-4F & G6PT-4R	4	249	62	1.5
G6PT-5F & G6PT-5R	5	143	62	1.5
G6PT-6F & G6PT-6R	6	191	62	1.5
G6PT-7F & G6PT-7R	7	310	55	1.5
G6PT-8F & G6PT-8R	8	360	55	1.5
G6PT-9F & G6PT-9R	9	251	65	1.5

Results of successful PCR amplification for each exon are presented as agarose gel electrophoresis images in Figure 4.15.

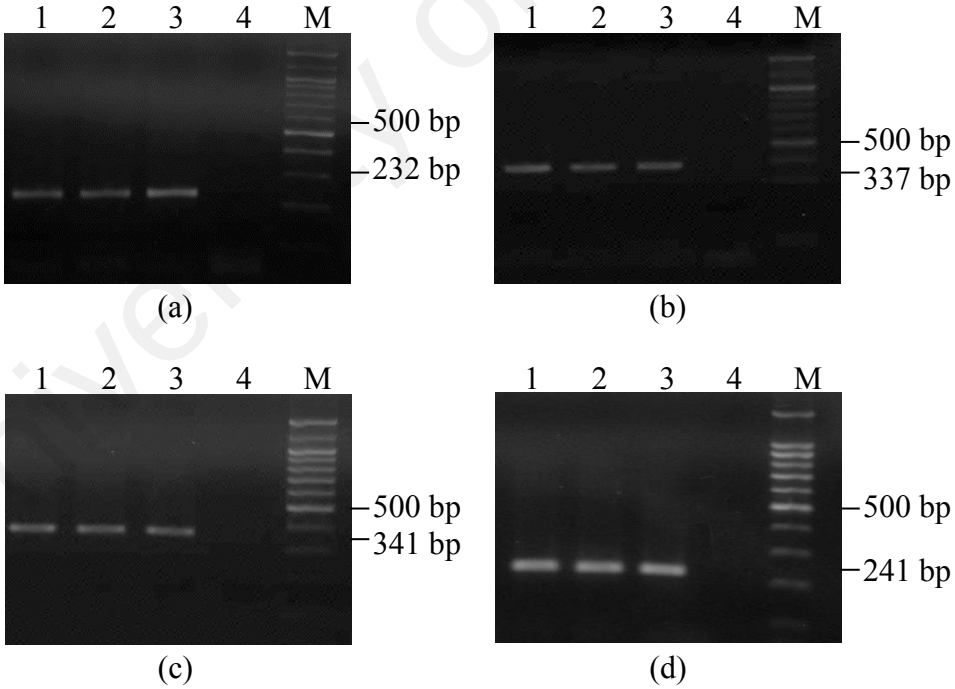


Figure 4.15: Example of agarose gel electrophoresis for *SLC37A4* gene amplification. (a) Exon 1, 232 bp (b) Exon 2, 337 bp (c) Exon 3, 341 bp (d) Exon 4, 249 bp. Lanes 1-3: PCR product from patients' sample, Lane 4: Negative control (no DNA template), Lane M: Molecular weight marker (100 bp ladder).

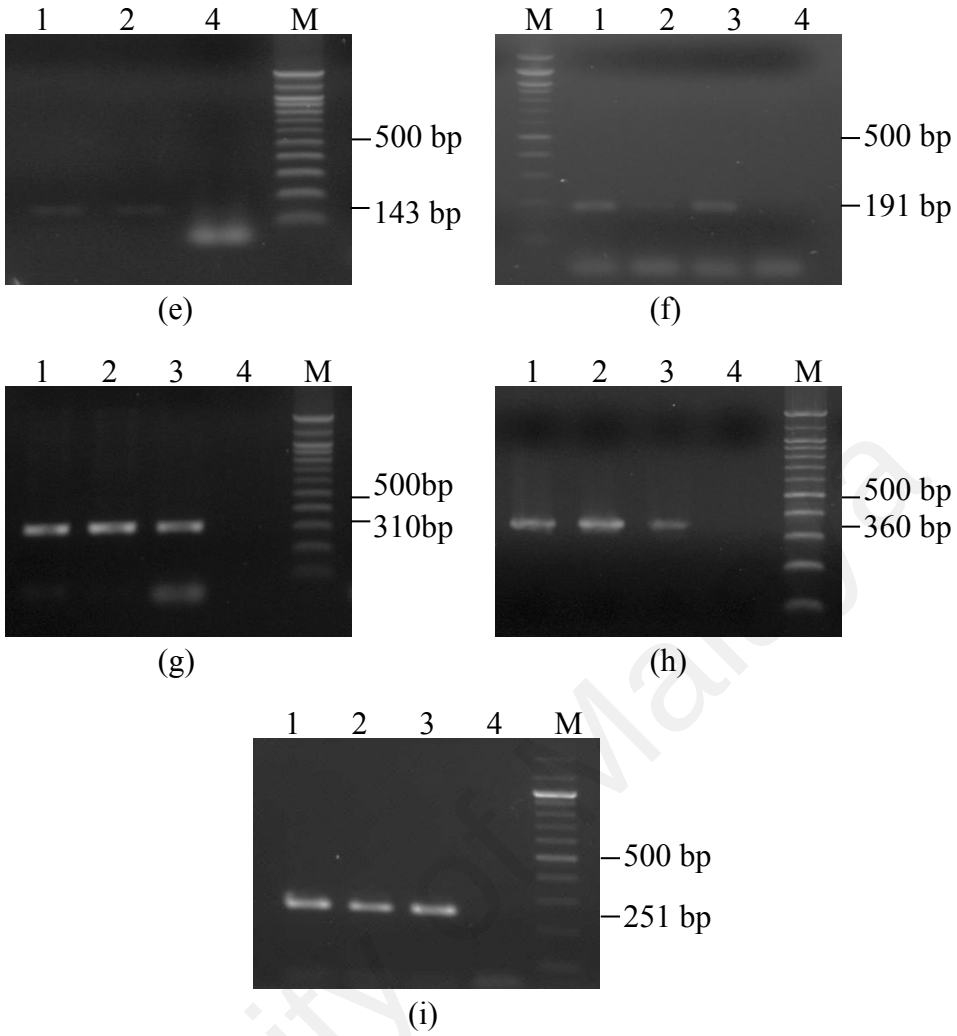


Figure 4.15, continued: Example of agarose gel electrophoresis for *SLC37A4* gene amplification. (e) Exon 5, 143 bp (f) Exon 6, 191 bp (g) Exon 7, 310 bp (h) Exon 8, 360 bp (i) Exon 9, 251 bp. Lanes 1-3: PCR product from patients' sample, Lane 4: Negative control (no DNA template), Lane M: Molecular weight marker (100 bp ladder).

4.2.2 Mutations Found in the *SLC37A4* Gene

Purified PCR products were subjected to sequencing analysis using the same primers used for PCR. Sequencing chromatograms were analyzed and edited (to remove ambiguity, where appropriate) using the Chromas sequence analysis software. Sequences generated from the sequencing reactions were compared with the *SLC37A4* reference sequence. As previously mentioned, any variation detected was compared against mutation databases to confirm their involvement in the pathogenesis of GSD 1b in these patients. In cases where no match was found, the probable involvement of suspected pathogenic mutations is supported by additional *in silico* prediction analyses.

(See section 4.4.1 Predictive analysis on the effect of mutation on protein function). A summary of the mutations found were listed in Table 4.4 and the sequencing result for each fragment is presented below: -

i) Exon 1

Exon 1 was PCR amplified with an expected size of 232 bp. No sequence variation was found within this exon for all patients.

ii) Exon 2

Exon 2 was PCR amplified with an expected size of 337 bp. No sequence variation was found within this exon for all patients.

iii) Exon 3

Exon 3 was PCR amplified with an expected size of 341 bp. No sequence variation was found within this exon for all patients.

iv) Exon 4

Exon 4 was PCR amplified with an expected size of 249 bp. No sequence variation was found within this exon for all patients.

v) Exon 5

Exon 5 was PCR amplified with an expected size of 143 bp. Sequencing analysis revealed one sequence variation within exon 5 of Patient 2. He was found to be homozygous for a c.196insT mutation (Figure 4.16). This mutation has not been previously reported as pathogenic. However, preliminary analysis suggested that this mutation will lead to a frameshift, and subsequently generate premature stop codon 6 amino acids downstream (See section 4.4.1 for detailed explanation).

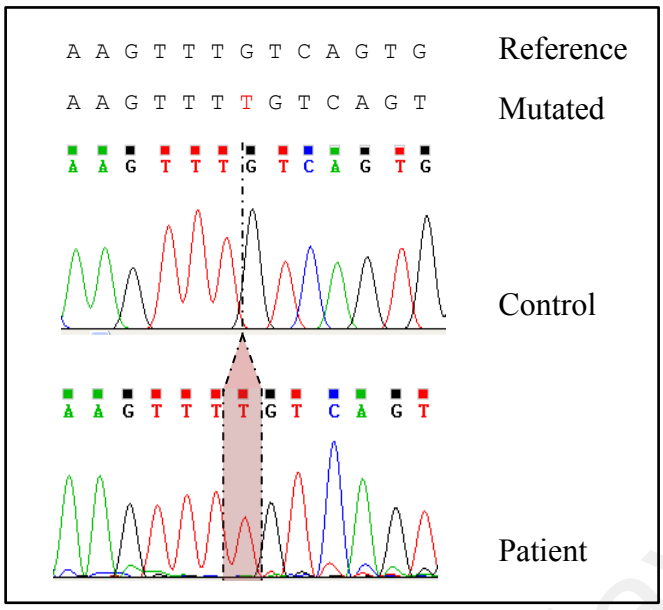


Figure 4.16: Partial sequencing chromatograms of exon 5 showing a T insertion for mutation c.196insT. Red letter in the mutated sequence and the coloured box highlights the inserted nucleotide.

vi) Exon 6

Exon 6 was PCR amplified with an expected size of 191 bp. No sequence variation was found within this exon for all patients.

vii) Exon 7

Exon 7 was PCR amplified with an expected size of 310 bp. No sequence variation was found within this exon for all patients.

viii) Exon 8

Exon 8 was PCR amplified with an expected size of 360 bp. No sequence variation was found within this exon for all patients.

ix) Exon 9

Exon 9 was PCR amplified with an expected size of 251 bp. Sequencing analysis revealed two sequence variations within this exon, one missense mutation and one deletion mutation.

- Mutation c.1265C>T (Figure 4.17a) was detected in exon 9 of Patient 3. He was found to be heterozygous for the mutation. This mutation maintains a serine residue at codon 425 (p.S425S) and has been previously reported to be a polymorphism (rs35010541).
- Mutation c.1286_1290delGAGTG (Figure 4.17b) detected in exon 9 of Patient 1. She was found to be homozygous for the mutation. This mutation has not been previously reported as pathogenic. However, preliminary analysis suggested a translational readthrough may occur and would require further analysis to confirm its role in the pathogenesis of GSD 1b in this patient (see section 4.4.1).

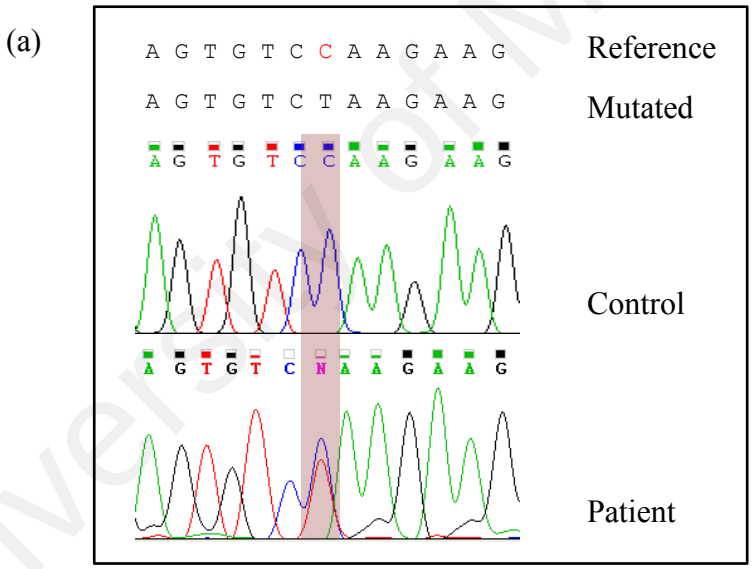


Figure 4.17: Partial sequencing chromatograms for exon 9 of the *SLC37A4* gene. Sequencing chromatograms showing a (a) C>T insertion for mutation c.1265C>T. Red letter in the reference sequence and the coloured box highlights the affected nucleotide.

(b)

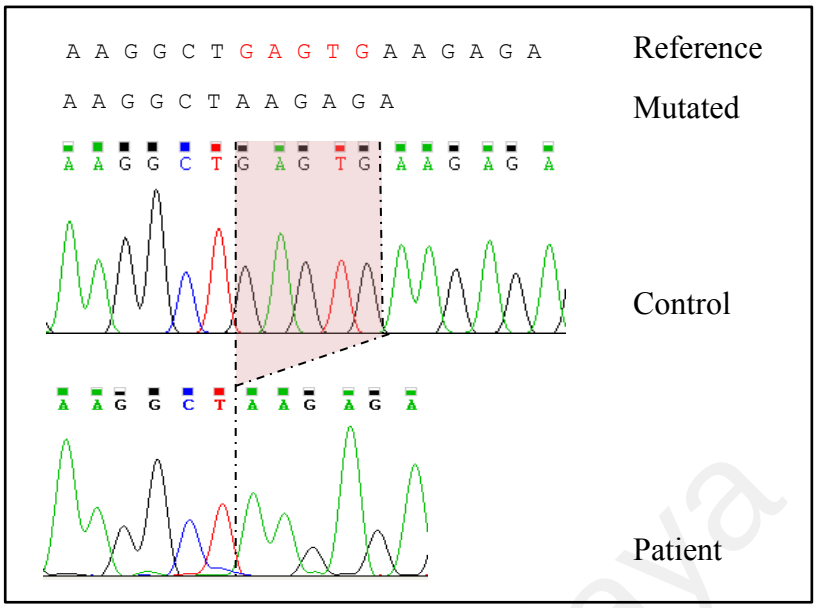


Figure 4.17, continued: Partial sequencing chromatograms for exon 9 of the *SLC37A4* gene. Sequencing chromatograms showing a (b) 5-bp deletion for mutation c.1286_1290delGAGTG. Red letters in the reference sequence and the coloured box highlight the deleted nucleotides.

Table 4.4: Genotypes of five Malaysian GSD 1b patients.

Patient	Race	Source	Genotype	*SLC37A4* Mutation	Effect	Exon	Reference
1	Malay	HKL	Homozygous	c.1286_1290delGAGTG	p.E429Kfs*59	9	Unreported
2	Indian	HKL	Homozygous	c.196insT	p.V66Cfs*6	5	Unreported
3	Indian	HKL	Heterozygous	# c.1275C>T	Non pathogenic	9	rs35010541
4	Malay	HKL		No mutation			
5	Chinese	HKL		No mutation			

Other sequence variation identified, but they have been reported to be polymorphism (not pathogenic).

4.2.3 PCR Optimization for *SLC2A2* Gene Amplification

Nine pairs of primers were used to amplify ten exons of the *SLC2A2* gene. The primers were designed according to gene reference sequence (GenBank accession no. NG_008108.1). *In silico* analysis carried out on the designed primers showed that all primers passed the PCR suitability test for most pre-set criteria (Selected examples are shown in Appendix F).

Annealing temperature and magnesium concentration optimization were carried out for all primer pairs. Different primer pairs were found to have T_A value of either 55 ºC or 60 ºC. All primer pairs were found to have an optimized magnesium concentration of 1.5 mM. A summary of the PCR conditions for *SLC2A2* gene amplification is presented in Table 4.5.

Table 4.5: Primers and optimized PCR condition for *SLC2A2* gene amplification.

Fragment	PCR primers	Exon	Expected size (bp)	Annealing temp. (ºC)	$[Mg^{2+}]$ (mM)
1	SLC2A2-1F/SLC2A2-1R	1	225	55	1.5
2	SLC2A2-2F/SLC2A2-2R	2	232	55	1.5
3	SLC2A2-3F/SLC2A2-3R	3	522	55	1.5
4	SLC2A2-4aF2/SLC2A2-4aR	4a	362	55	1.5
5	SLC2A2-4bF/SLC2A2-4bR	4b	251	55	1.5
6	SLC2A2-5F/SLC2A2-6R2	5,6	1000	60	1.5
7	SLC2A2-7F/SLC2A2-7R	7	348	55	1.5
8	SLC2A2-8F/SLC2A2-8R	8	340	60	1.5
9	SLC2A2-9F/SLC2A2-10R	9,10	629	55	1.5

Results of successful PCR amplification for each fragment are presented as agarose gel electrophoresis images in Figure 4.18.

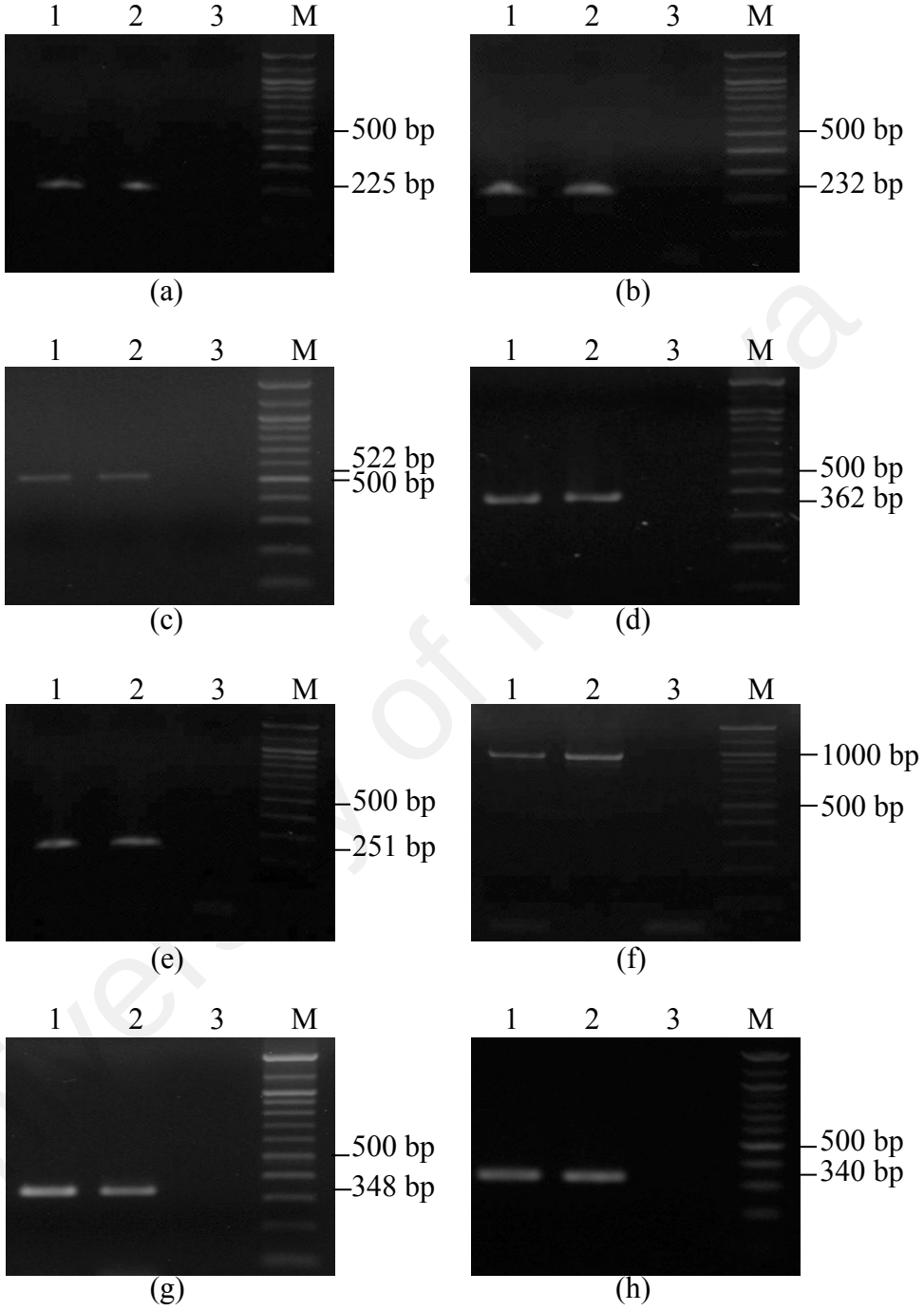
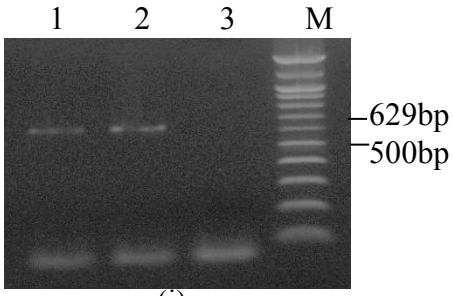


Figure 4.18: Example of agarose gel electrophoresis for *SLC2A2* gene amplification. (a) Fragment 1, 225 bp (b) Fragment 2, 232 bp (c) Fragment 3, 522 bp (d) Fragment 4, 362 bp (e) Fragment 5, 251 bp (f) Fragment 6, 1000 bp (g) Fragment 7, 348 bp (h) Fragment 8, 340 bp. Lanes 1-2: PCR product from patients' sample, Lane 3: Negative control (no DNA template), Lane M: Molecular weight marker (100 bp ladder).



(i)

Figure 4.18, continued: Example of agarose gel electrophoresis for *SLC2A2* gene amplification. (i) Fragment 9, 629 bp. Lanes 1-2: PCR product from patients' sample, Lane 3: Negative control (no DNA template), Lane M: Molecular weight marker (100 bp ladder).

4.2.4 Mutations Found in the *SLC2A2* Gene

Purified PCR products were subjected to sequencing using the same primers used for PCR. Sequencing chromatograms were analyzed and edited (to remove ambiguity, where appropriate) using the Chromas sequence analysis software. Sequences generated from the sequencing reactions were compared with the *SLC2A2* reference sequence. As previously mentioned, any variation detected was compared against mutation databases to confirm their involvement in the pathogenesis of FBS in these patients. In cases where no match was found, the probable involvement of suspected pathogenic mutations is supported by additional *in silico* prediction analyses. (See section 4.4.1). A summary of the mutations found were listed in Table 4.6 and the sequencing result for each fragment is presented below: -

i) Fragment 1

Exon 1 was amplified in fragment 1 with an expected size of 225 bp. No sequence variation was found within this exon for all patients.

ii) Fragment 2

Exon 2 was amplified in fragment 2 with an expected size of 232 bp. No sequence variation was found within this exon for all patients.

iii) Fragment 3

Exon 3 was amplified in fragment 3 with an expected size of 522 bp. No sequence variation was found within this exon for all patients.

iv) Fragment 4

Exon 4a (longer isoform) was amplified in fragment 4 with an expected size of 362 bp. No sequence variation was found within this exon for all patients.

v) Fragment 5

Exon 4b (shorter isoform) was amplified in fragment 5 with an expected size of 251 bp. No sequence variation was found within this exon for all patients.

vi) Fragment 6

Exon 5 and 6 were amplified as a single fragment 6 with an expected size of 1000 bp fragment. Sequencing analysis revealed three different variations which include two missense mutations and one deletion mutation.

- Mutation c.589G>C (Figure 4.19a) was detected in exon 5 of Patient 2. He was found to be homozygous for the mutation. This mutation has not been previously reported as pathogenic. However, the mutation replaces a valine residue with a leucine at codon 197 (p.V197L), and would require further analysis to confirm its role in the pathogenesis of FBS in this patient (See section 4.4.1).
- The c.872_873delTA (Figure 4.19b) mutation was detected in exon 6 of Patient 1. She was found to be heterozygous for the mutation. This mutation has not been previously reported to be pathogenic. However, preliminary analysis suggested that this mutation will lead to a frameshift, and subsequently generate premature stop codon 100 amino acids downstream (See section 4.4.1).

- Mutation c.952G>A (Figure 4.19c) was found in exon 6 of Patient 1. She was found to be heterozygous for the mutation. This mutation that replaces a glycine residue with an arginine at codon 318 (p.G318R) has been previously reported to be pathogenic.

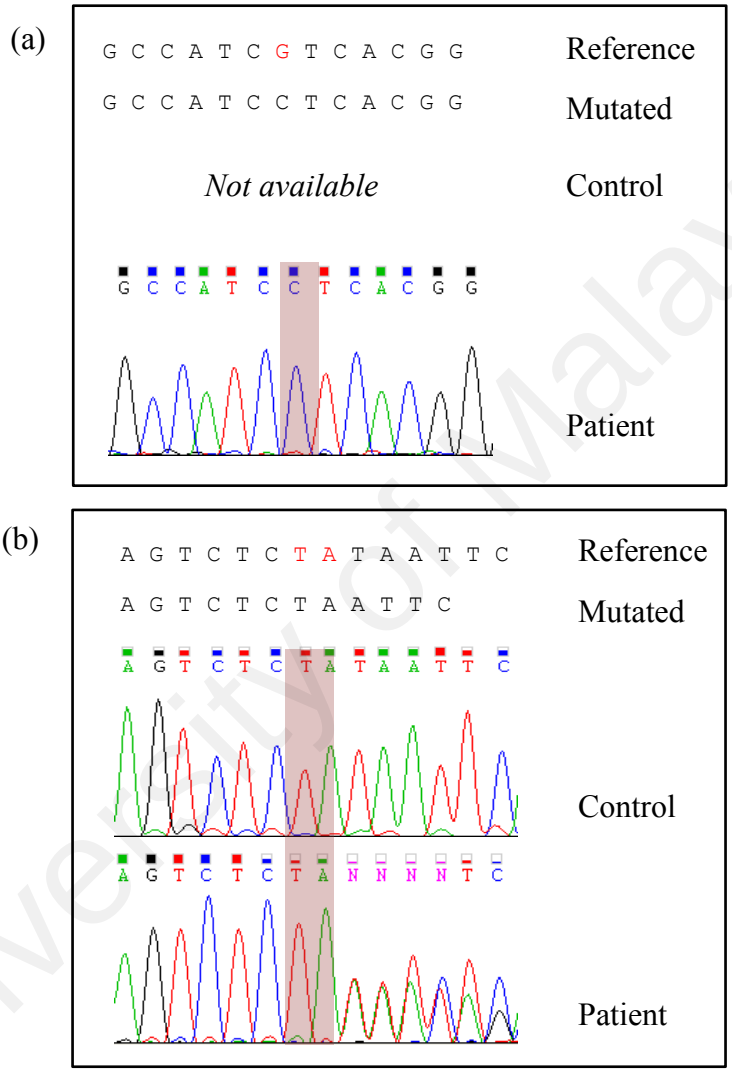


Figure 4.19: Partial sequencing chromatograms showing mutations found within fragment 6 of the *SLC2A2* gene. The chromatograms showing a (a) G>C substitution for mutation c.589G>C (b) 2-bp deletion for mutation c.872_873delTA. Red letters in the reference sequence and the coloured box highlight the affected nucleotides.

(c)

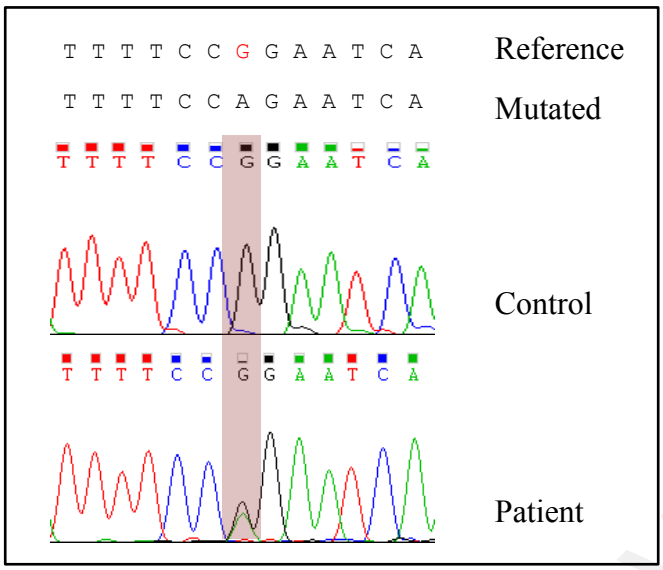


Figure 4.19, continued: Partial sequencing chromatograms showing mutations found within fragment 6 of the *SLC2A2* gene. The chromatogram (c) showing a G>A substitution for mutation c.952G>A. Red letter in the reference sequence and the coloured box highlights the affected nucleotide.

vii) Fragment 7

Exon 7 was amplified in fragment 7 with an expected size of 348 bp. No sequence variation was found within this exon for all patients.

viii) Fragment 8

Exon 8 was amplified as in fragment 8 with an expected size of 340 bp. No sequence variation was found within this exon for all patients.

ix) Fragment 9

Exon 9 and 10 were amplified as a single fragment 9 with an expected size of 629 bp. No sequence variation was found within this exon for all patients.

Table 4.6: Genotypes of four Malaysian FBS patients.

Patient	Race	Source	Genotype	Mutation	Effect	Exon	Reference
1	Chinese	UMMC	Heterozygous	c.952G>A	p.G318R	6	Santer et al., 2002
				c.872_873delTA	p.I291Nfs*100	6	Unreported
2	Indian	HKL	Homozygous	c.589G>C	p.V197L	5	Unreported
3	Malay	UMMC		No mutation			
4	Chinese	UMMC		No mutation			

4.3 Mutation Screening Results of the *MUT* Gene (MMA Mut-Type) and the *MMACHC* Gene (MMA CblC-Type)

A single MMA (mut-type) patient and two MMA cblC-type patients were recruited for mutation screening of the *MUT* gene and the *MMACHC* gene respectively. The results for PCR optimization and mutation screening of the *MUT* gene are presented in section 4.3.1 and section 4.3.2 respectively. The results for PCR optimization and mutation screening of the *MMACHC* gene are presented in section 4.3.3 and section 4.3.4 respectively.

4.3.1 PCR Optimization for *MUT* Gene Amplification

Four pairs of primers were used to amplify the 13 exons covering the coding region of *MUT* gene. The primers were designed according to gene reference sequence (GenBank accession no. M65131.1). *In silico* analysis carried out on the designed primers showed that all primers passed the PCR suitability test for most pre-set criteria (Selected examples are shown in Appendix G).

Optimization of the annealing temperature and magnesium concentration was carried out for all primer pairs. All primer pairs were found to have an optimized T_A value of 55 ºC and optimized magnesium concentration of 1.5 mM. A summary of the PCR conditions for *MUT* gene amplification is presented in Table 4.7.

Table 4.7: Primers and optimized PCR condition for *MUT* gene amplification.

Fragment	PCR primers	Exon	Expected size (bp)	Annealing temp. (ºC)	$[Mg^{2+}]$ (mM)
1	mut_1F/ mut_2R	1, 2	570	55	1.5
2	mut_3F/ mut_4R	3, 4, 5	647	55	1.5
3	mut_5F/ mut_6R	6, 7, 8, 9	724	55	1.5
4	mut_7F/ mut_8R	10, 11, 12, 13	738	55	1.5

Results of successful PCR amplification for each fragment are presented as agarose gel electrophoresis images in Figure 4.20 below.

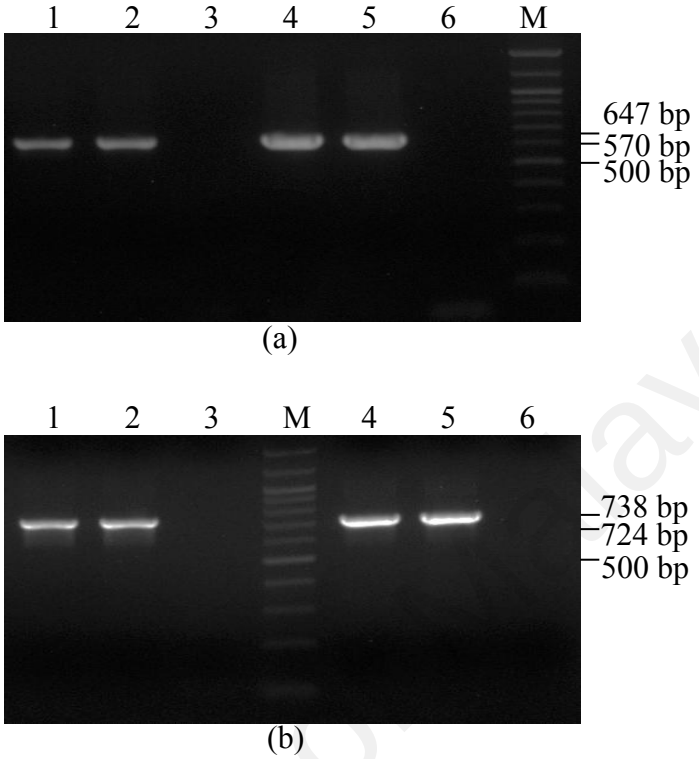


Figure 4.20: Example of agarose gel electrophoresis for *MUT* gene amplification. (a) Fragment 1, 570 bp and fragment 2, 647 bp (b) Fragment 3, 724 bp and fragment 4, 738 bp. Lanes 1-2, 4-5: PCR product from patients' sample, Lane 3, 6: Negative control (no DNA template), Lane M: Molecular weight marker (100 bp ladder).

4.3.2 Mutations Found in the *MUT* Gene

Purified PCR products were subjected to sequencing using the same primers used for PCR. Sequencing chromatograms were analysed and edited (to remove ambiguity, where appropriate) using Chromas sequence analysis software. Sequences generated from the sequencing reactions were compared with *MUT* reference sequence. As previously mentioned, any variation detected was compared against mutation databases to confirm their involvement in the pathogenesis of MMA in the patient. The mutations found were listed in Table 4.8 and the sequencing result for each fragment is presented below: -

i) Fragment 1

Exon 1 and exon 2 were amplified as a single fragment 1 with an expected size of 570 bp. No sequence variations were found in both exons for this patient.

ii) Fragment 2

Exon 3, 4 and 5 were amplified as a single fragment 2 with an expected size of 647 bp. No sequence variations were found in all exons for this patient.

iii) Fragment 3

Exon 6, 7, 8 and 9 were amplified as a single fragment 3 with an expected size of 724 bp. Sequencing analysis revealed two sequence variations within this fragment which include a missense mutation and insertion/deletion mutation.

- Mutation c.1280G>A (Figure 4.21a) was found in exon 6. He was found to be heterozygous for the mutation. Mother and father samples were available for sequencing and analysis on their sample revealed the father as the carrier for mutation. This mutation that replaces a glycine with an aspartic acid at codon 427 (p.G427D) had been previously reported to be pathogenic (Worgan et al., 2006).
- Mutation c.1630_1631delGGinsTA (Figure 4.21b) was found in exon 9. He was found to be heterozygous for the mutation. Similarly, the mother and the father samples were available for sequencing and analysis on their sample revealed mother as the carrier for the mutation. This mutation that causes a frameshift and subsequent premature stop codon (p.G544X) had been previously reported (Champattanachai et al., 2003).

iv) Fragment 4

Exon 10, 11, 12 and 13 were amplified as a single fragment 4 with an expected size of 738 bp. No sequence variations were detected in all exons for the patient.

(a)

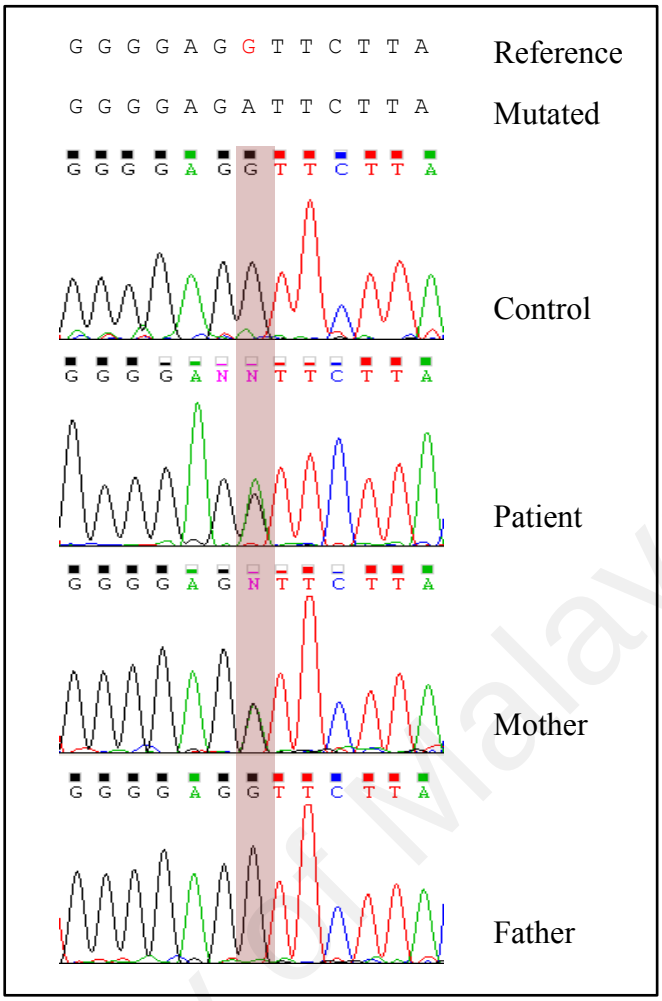


Figure 4.21: Partial sequencing chromatograms of exon 6 showing a (a) G>A substitution for mutation c.1280G>A. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

(b)

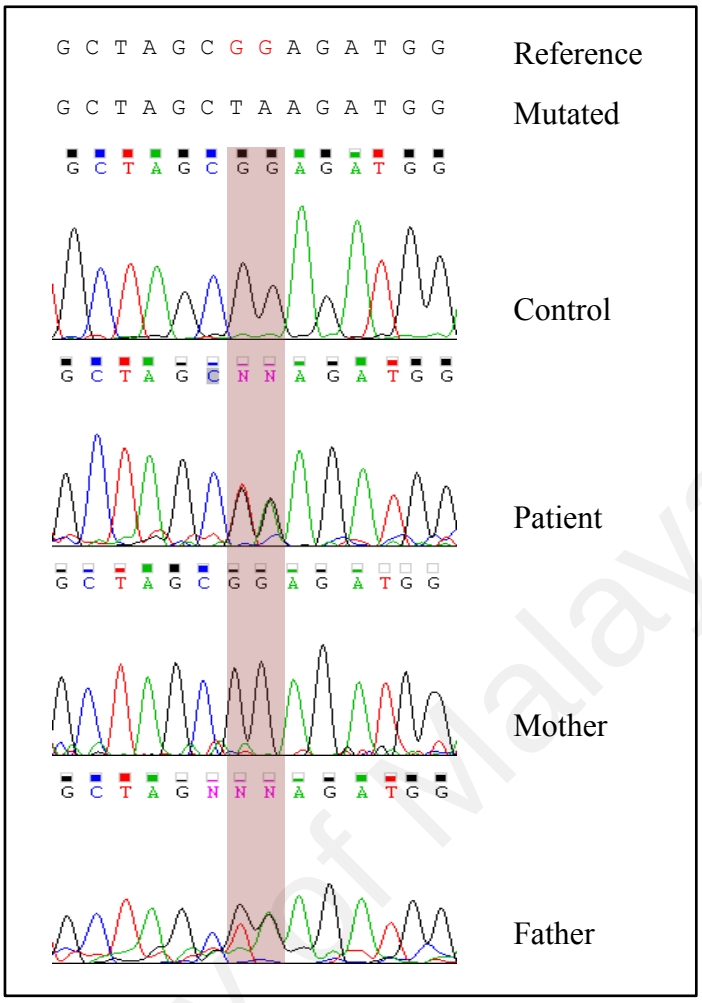


Figure 4.21, continued: Partial sequencing chromatograms of exon 9 showing a (b) 2-bp indels for mutation c.1630_1631delGGinsTA. Red letters in the reference sequence and the purple box highlight the deleted nucleotides.

Table 4.8: Genotypes of a Malaysian MMA patient.

Race	Source	Mutation	Effect	Exon	Remark
Chinese	UMMC	c.1280G>A	p.G427D	6	Worgan et al., 2006
		c.1630_1631delGGinsTA	p.G544X	9	Champattanachai et al., 2003

4.3.3 PCR Optimization for *MMACHC* Gene Amplification

Four pairs of primers were used to amplify four exons covering the coding region of the *MMACHC* gene. The primers were designed according to gene reference sequence (GenBank accession no. NG_013378.1). *In silico* analysis carried out on the designed primers showed that all primers passed the PCR suitability test for most pre-set criteria (Selected examples are shown in Appendix H).

Optimization of the annealing temperature was carried out for all primer pairs. The temperature that yielded fragment with correct expected size was chosen as the optimized annealing temperature (T_A). All primer pairs were found to have a T_A value of 55 ºC. Subsequent magnesium concentration optimizations found that a concentration of 1.5 mM was suitable for all primer pairs to generate fragments with the expected size. A summary of the PCR conditions for the *MMACHC* gene amplification is presented in Table 4.9.

Table 4.9: Primers and optimized PCR condition for *MMACHC* gene amplification.

Primers for PCR	**Exon**	**Expected size (bp)**	**Annealing temp. (ºC)**	**[Mg^{2+}] (mM)**
MMACHC_1F/ MMACHC_1R	1	600	55	1.5
MMACHC_2F/ MMACHC_2R	2	467	55	1.5
MMACHC_3F/ MMACHC_3R	3	395	55	1.5
MMACHC_4F/ MMACHC_4R	4	573	55	1.5

Results of successful PCR amplification for each exon are presented in Figure 4.22 as below.

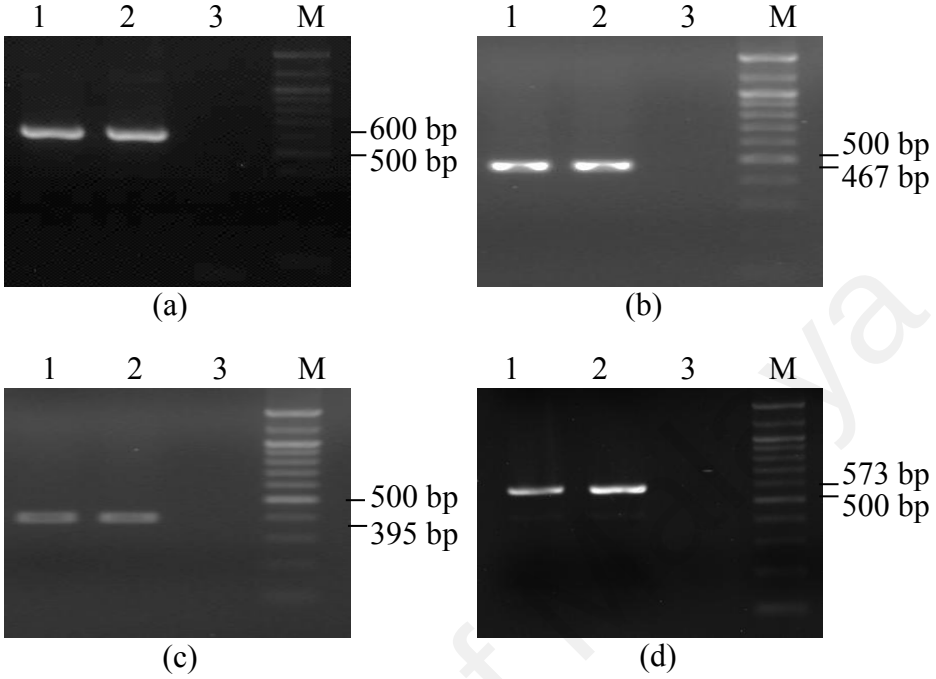


Figure 4.22: Example of agarose gel electrophoresis for *MMACHC* gene amplification. (a) Exon 1, 600 bp (b) Exon 2, 467 bp (c) Exon 3, 395 bp (d) Exon 4, 573 bp. Lanes 1-2: PCR product from patients' sample, Lane 3: Negative control (no DNA template), Lane M: Molecular weight marker (100 bp ladder).

4.3.4 Mutations Found in the *MMACHC* Gene

Purified PCR products were subjected to sequencing analysis using the same primers used for PCR. Sequencing chromatograms were analysed and edited (to remove ambiguity, where appropriate) using Chromas sequence analysis software. Sequences generated from the sequencing reactions were compared with the *MMACHC* reference sequence. As previously mentioned, any variation detected was compared against mutation databases to confirm their involvement in the pathogenesis of MMA cblC-type in these patients. A summary of the mutations found were listed in Table 4.10 and the sequencing result for each fragment is presented below: -

i) Exon 1

Exon 1 was amplified with an expected size of 600 bp. No sequence variation was found within this exon for both patients.

ii) Exon 2

Exon 2 was amplified with an expected size of 467 bp. No sequence variation was found within this exon for both patients.

iii) Exon 3

Exon 3 was amplified with an expected size of 395 bp. No sequence variation was found within this exon for both patients.

iv) Exon 4

Exon 4 was amplified with an expected size of 573 bp. Sequence analysis revealed a single variation, a c.609G>A (Figure 4.23) that was identified in homozygous form in both patients (siblings). Mother and father samples were available for sequencing and analysis on their sample revealed that they are carriers for the mutation. This mutation that causes premature termination at codon 209 (p.W209X) had been previously reported to be causative of MMA cblC-type (Lerner-Ellis et al., 2006).

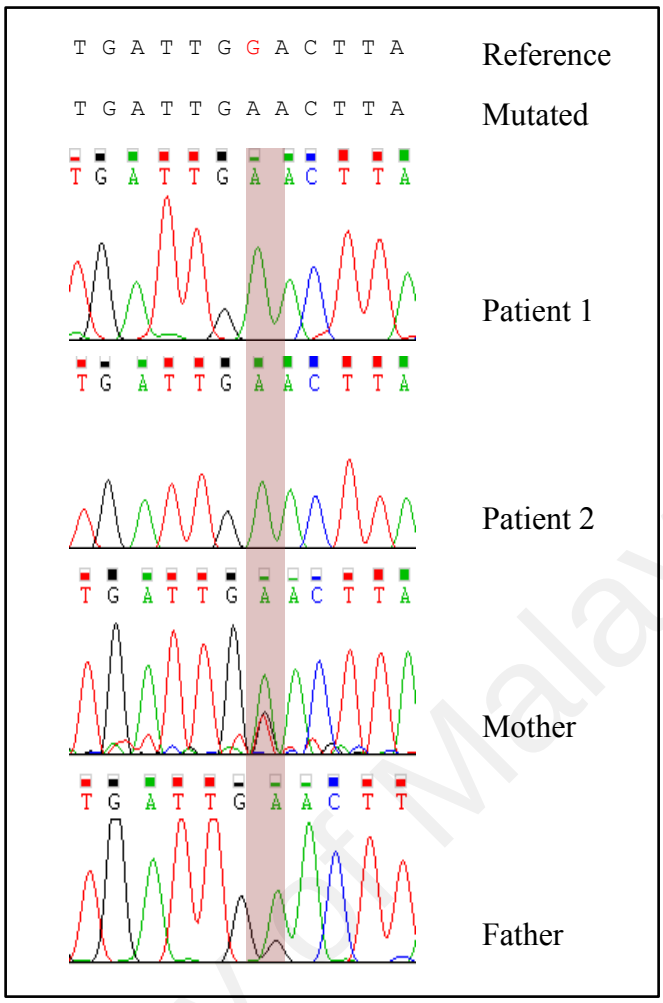


Figure 4.23: Partial sequencing chromatograms of exon 4 showing a G>A substitution for mutation c.609G>A. Red letter in the reference sequence and the coloured box highlights the substituted nucleotide.

Table 4.10: Genotypes of two Malaysian MMA cblC-type patients.

Patient	Race	Source	Mutation	Effect	Exon	Remark
1	Chinese	UMMC	c.609G>A	p.W203X	4	Lerner-Ellis et al., 2006
			c.609G>A	p.W203X	4	Lerner-Ellis et al., 2006
2	Chinese	UMMC	c.609G>A	p.W203X	4	Lerner-Ellis et al., 2006
			c.609G>A	p.W203X	4	Lerner-Ellis et al., 2006

4.4 Inference of Pathogenicity for Unreported Mutations

There are several mutations identified in the present study that are yet unreported. In order to gain insight on whether these mutations are pathogenic and causative of the respective diseases in these patients, predictive analysis and/or experimental assays were carried out to obtain supporting evidence. These include a) predictive functional analysis using some freely available software such as the Sequence Manipulation Suite (www.bioinformatics.org/sms2/), b) screening for the presence (or absence) of the mutation in a panel of healthy individuals, and c) experimental assays to determine changes in normal gene function.

4.4.1 Predictive Analysis on the Effect of Mutation on Protein Function

Prediction on the changes in amino acid coding was carried out using the Sequence Manipulation Suite (SMS) version 2 software (http://www.bioinformatics.org/sms2/reference.html), while prediction on potential protein functionality effect was carried out using PolyPhen-2 software (http://genetics.bwh.harvard.edu/pph2/). Finally, prediction on the effects of intronic mutation on normal splicing was achieved using the HSF version 3 software (http://www.umd.be/HSF3/).

a) Mutations in the *AGL* gene for GSD III

i. c.1423+1G>T

This mutation is predicted to affect mRNA splicing as it occurs at the conserve exon-intron boundary, critical for RNA processing during transcription. *In silico* analysis using Human Splice Finder (HSF) predict two possible outcomes. First, the G>T mutation could result in the creation of a new splicing signal motif (from CCGG***g***tatg to CCGg***t***tatg) 1 base upstream of the normal exon/intron border, and this is predicted with a variation score of 54 % (threshold set at +30 %) (Table 4.11).

Table 4.11: Prediction of the c.1423+1G>T mutation effect on the splicing signal. Prediction was based on two algorithms; Position Weight Matrices and Maximum Entropy. The point of splicing is differentiated with small and capital letters in the motif sequence. Mutated nucleotide is in bold and italicized.

Algorithm type	Prediction algorithm	Splice site type	Motif	New splice site	Wild type	Mutant	Variation (%)
Position Weight Matrices	HSF	Donor	CCGG***g***tatg	CCGg***t***tatg	49.14	75.98	New site +54.62
		Donor	CGG***g***tatgt	CGG***t***tatgt	86.34	59.50	WT site broken -31.09
Maximum Entropy	MaxEnt	-	CGG***g***tatgt	CGG***t***tatgt	8.79	0.28	WT site broken -96.81

Alternatively, the G>T sequence change (from CGG***g***tatgt to CGG***t***tatgt) gave a score variation that is more than -30 % (threshold value = -10 %) indicating that the mutation would result in the disruption of the splicing motif. This latter prediction is further corroborated using the MaxEnt algorithm where the sequence change resulted in a variation score greater than -90 % (threshold value = -30 %). Based on the prediction analysis, intron 12 retention is expected to occur and subsequently generation of a termination signal just two amino acids downstream of exon 12 boundaries as illustrated in Figure 4.24.

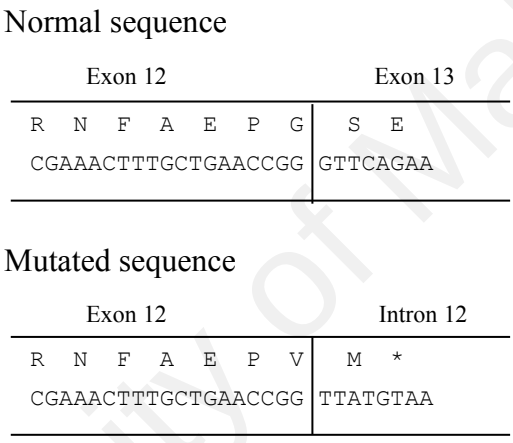


Figure 4.24: Diagram to show the reading frame of correctly spliced exon 12 and exon 13 (top panel) in comparison with the introduction of a termination signal when translation proceeds along intron 12 that has been retained as a result of splicing error (bottom panel).

ii. c.2914_2915delAA

This mutation is predicted to result in translation reading frameshift which is expected to generate premature stop codon 10 amino acids downstream of the deletion site (p.N972Pfs*10) (Figure 4.25). Premature translation termination will result in protein truncation with a loss of 550 amino acids. This loss of almost a third of the C terminal end of the protein includes the glucosidase domain and glycogen binding domain, which renders this mutation to be highly disruptive to normal GDE function and would be expected to be pathogenic.

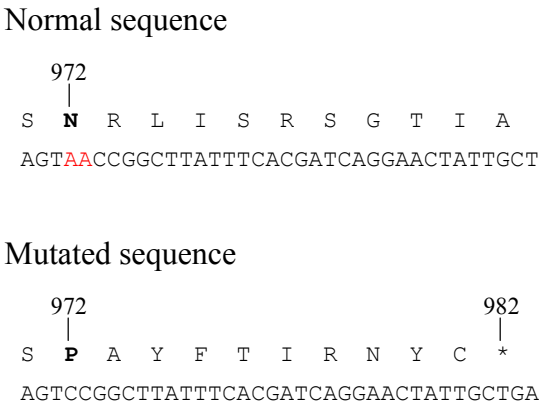


Figure 4.25: Diagram to show the normal reading frame (top panel) in comparison with the introduction of a termination signal (bottom panel) when translation proceeds after deletion of two nucleotides (red letters) at codon 972. The bold letters indicate the affected amino acids.

iii. c.3814_3815delAG

This mutation is predicted to result in translation reading frameshift which is predicted to generate premature stop codon 18 amino acids downstream of the deletion site (p.R1272Rfs*18) (Figure 4.26). Premature translation termination will result in protein truncation with a loss of 242 amino acids. This loss includes the important glycogen binding domain at the N terminal of GDE, which renders this mutation to be highly disruptive to normal GDE function and would be expected to be pathogenic.

Figure 4.26: Diagram to show the normal reading frame (top panel) in comparison with the introduction of a termination signal (bottom panel) when translation proceeds after deletion of two nucleotides (red letters) at codon 1272. The bold letters indicate the affected amino acids.

iv. c.4333T>G

This mutation that occurs due to a T to G substitution at codon 1445 results in replacement of tyrosine residue by an aspartic acid (p.Y1445D). The replacement of the partially hydrophobic tyrosine into charged and polar aspartic acid is predicted to have unfavourable impact on protein function. This prediction is due to tyrosine residue was found to be highly conserved through multiple sequence alignment of *AGL*s from various species (Figure 4.27). *In silico* functional prediction using the PolyPhen-2 analysis software classed the c.4333T>G mutation as "probably damaging" with a score of 0.999 (sensitivity: 0.90, specificity: 0.99), which renders this mutation to be highly disruptive to normal GDE function and would be expected to be pathogenic.

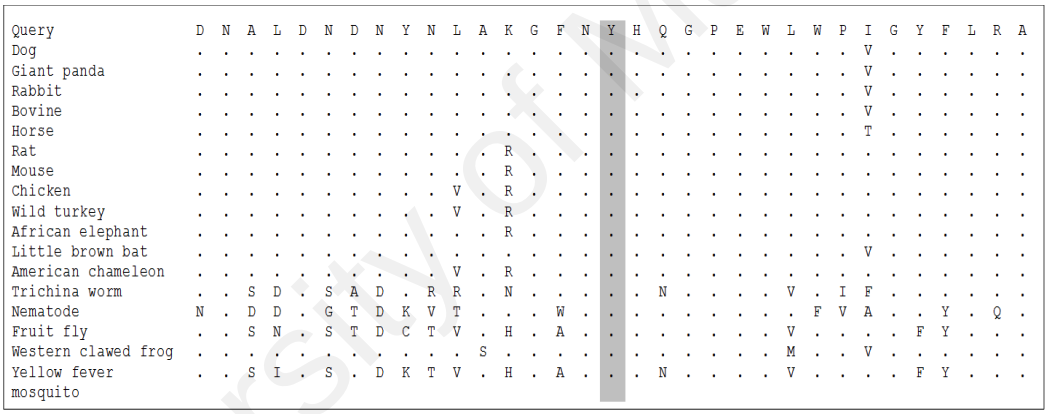
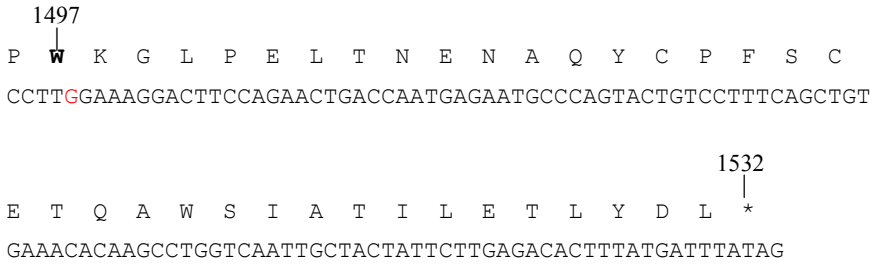


Figure 4.27: Sequence alignment of *AGL*s from various species. The highlighted letters indicate the conserved tyrosine (Y) residue which is substituted to an aspartic acid (D) in mutation c.4333T>G. Dots indicate residues that are identical with the query (human) sequences.

v. c.4490G>A

This mutation generates a premature termination at codon position 1497 (p.W1497X) (Figure 4.28), resulting in protein truncation with a loss of 35 amino acids at the C-terminal region. This missing region spans the glycogen binding domain, important for GDE activity and as such would be expected to disrupt normal protein function.

Normal sequence



Mutated sequence

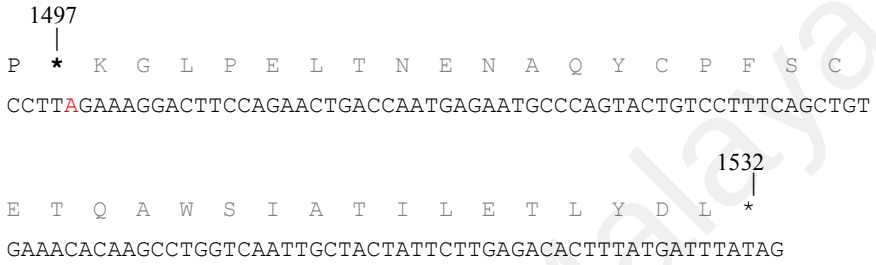


Figure 4.28: Diagram to show the normal reading frame (top panel) in comparison with the introduction of a termination signal (bottom panel) when translation proceeds after the substitution of guanine with adenine (red letters) at codon 1497. The bold letter indicates the affected amino acid and grey letters indicate the loss amino acids.

vi. c.4531_4534delTGTC

The mutation is predicted to result in translation reading frameshift which is expected to disrupt normal stop codon. Subsequent translational readthrough with addition of 14 amino acids forms a new stop codon 35 amino acids downstream (p.C1511Lfs*35) (Figure 4.29). The additional amino acids are predicted to alter the protein structure and as such would be expected to interfere with normal protein function.

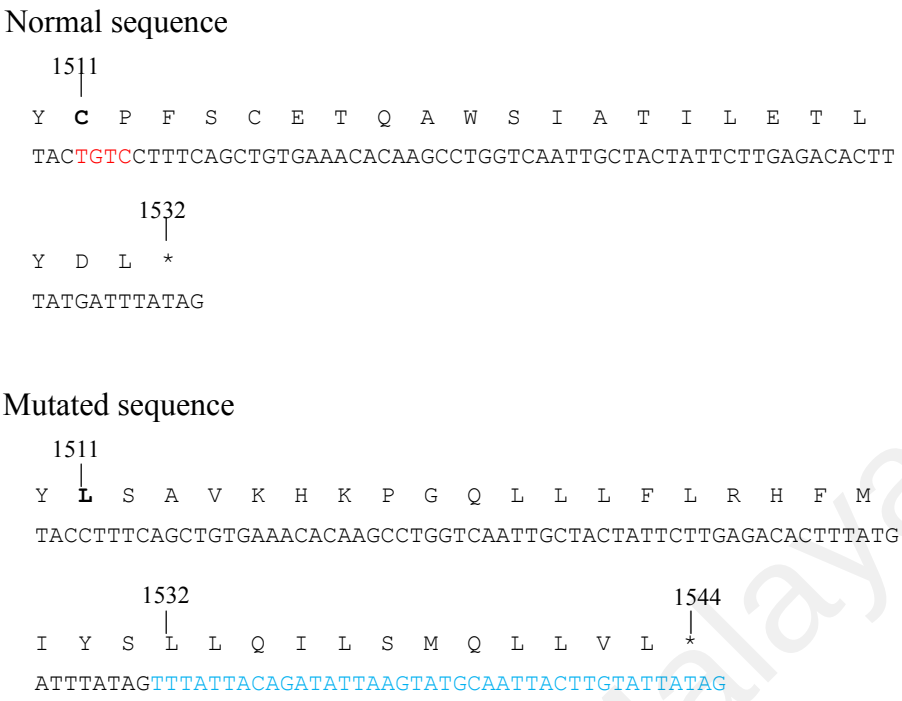


Figure 4.29: Diagram to show the normal reading frame (top panel) in comparison with the introduction of a termination signal (bottom panel) when translation proceeds after the deletion of four nucleotides (red letters) at codon 1511. The bold letters indicate the affected amino acids and blue letters indicate the readthrough sequences.

b) Mutations found in the *SLC37A4* gene for GSD 1b

i. c.196insT

This mutation is predicted to result in translation reading frameshift which is expected to generate premature stop codon six amino acids downstream of the insertion site (p.V66Cfs*6) (Figure 4.30). Premature translation termination will result in protein truncation with loss of 357 amino acids. The missing region spans the signature motif, transmembrane helices as well as the carboxyl terminal ER retention signal of *SLC37A4* gene, and as such would be expected to disrupt normal protein function.

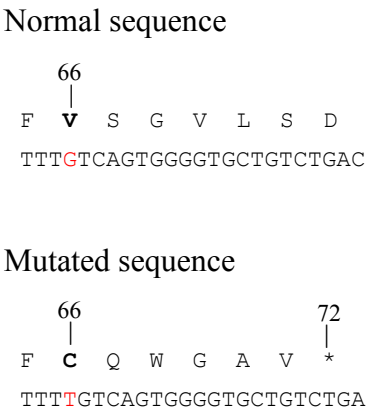


Figure 4.30: Diagram to show the normal reading frame (top panel) in comparison with the introduction of a termination signal (bottom panel) when translation proceeds after the substitution of guanine with tyrosine (red letters) at codon 66. The bold letters indicate the affected amino acids.

ii. c.1286_1290delGAGTG

This mutation is predicted to result in translation reading frameshift which is expected to disrupt normal stop codon. Subsequent translational readthrough would results in formation of a new stop codon 58 amino acids downstream (Figure 4.31). The readthrough is predicted to alter the protein structure and as such would be expected to interfere with normal protein function.

Figure 4.31: Diagram to show the normal reading frame (top panel) in comparison with the introduction of a termination signal (bottom panel) when translation proceeds after the deletion of five nucleotides (red letters) at codon 429. The bold letters indicate the affected amino acids and blue letters indicate the readthrough sequences.

c) Mutations found in the *SLC2A2* gene for FBS

i. c.589G>C

This mutation that occurs due to a G to C substitution at codon 197 results in replacement of valine residue by a leucine (p.V197L). Despite the replacement of hydrophobic valine with another hydrophobic leucine, the mutation is predicted to have unfavourable impact on protein function. This prediction is due to valine residue was found to be highly conserved through multiple sequence alignment of *SLC2A2*s from various species (Figure 4.32). *In silico* functional prediction using the PolyPhen-2 analysis software classed the c.589G>C mutation as "probably damaging" with a score of 0.961 (sensitivity: 0.62, specificity: 0.92), which renders this mutation to be highly disruptive to normal GLUT2 function and would be expected to be pathogenic.

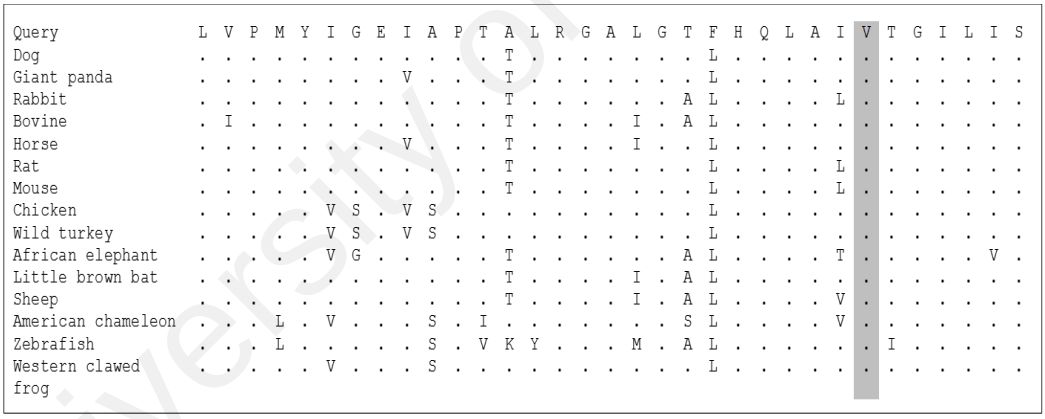


Figure 4.32: Sequence alignment of *SLC2A2*s from various species. The highlighted letters indicate the conserved valine residue which is substituted to a leucine in mutation c.589G>C. Dots indicate residues that are identical with the query (human) sequences.

ii. c.872_873delTA

This mutation is predicted to result in translation reading frameshift which is expected to generate premature stop codon 100 amino acids downstream of the deletion site (p.I291Nfs*100) (Figure 4.33). Premature translation termination will result in protein truncation with loss of 132 amino acids. The missing region includes a sugar signaling domain, sugar tunnel and sugar binding site, and as such would be expected to disrupt normal protein function.

Normal sequence

```
   291
    |
S  I  I  Q  L  F  T  N  S  S  Y  R  Q  P  I  L  V  A  L  M
TCTATAATTCAGCTCTTCACCAATTCCAGCTACCGACAGCCTATTCTAGTGGCACTGATG

L  H  V  A  Q  Q  F  S  G  I  N  G  I  F  Y  Y  S  T  S  I
CTGCATGTGGCTCAGCAATTTTCCGGAATCAATGGCATTTTTTACTACTCAACCAGCATT

F  Q  T  A  G  I  S  K  P  V  Y  A  T  I  G  V  G  A  V  N
TTTCAGACGGCTGGTATCAGCAAACCTGTTTATGCAACCATTGGAGTTGGCGCTGTAAAC

M  V  F  T  A  V  S  V  F  L  V  E  K  A  G  R  R  S  L  F
ATGGTTTTCACTGCTGTCTCTGTATTCCTTGTGGAGAAGGCAGGGCGACGTTCTCTCTTT

L  I  G  M  S  G  M  F  V  C  A  I  F  M  S  V  G  L  V  L
CTAATTGGAATGAGTGGGATGTTTGTTTGTGCCATCTTCATGTCAGTGGGACTTGTGCTG

L  N  K
CTGAATAAG
```

Mutated sequence

```
   291
    |
S  N  S  A  L  H  Q  F  Q  L  P  T  A  Y  S  S  G  T  D  A
TCTAATTCAGCTCTTCACCAATTCCAGCTACCGACAGCCTATTCTAGTGGCACTGATGCT

A  C  G  S  A  I  F  R  N  Q  W  H  F  L  L  L  N  Q  H  F
GCATGTGGCTCAGCAATTTTCCGGAATCAATGGCATTTTTTACTACTCAACCAGCATTTT

S  D  G  W  Y  Q  Q  T  C  L  C  N  H  W  S  W  R  C  K  H
TCAGACGGCTGGTATCAGCAAACCTGTTTATGCAACCATTGGAGTTGGCGCTGTAAACAT

G  F  H  C  C  L  C  I  P  C  G  E  G  R  A  T  F  S  L  S
GGTTTTCACTGCTGTCTCTGTATTCCTTGTGGAGAAGGCAGGGCGACGTTCTCTCTTTCT

N  W  N  E  W  D  V  C  L  C  H  L  H  V  S  G  T  C  A  A
AATTGGAATGAGTGGGATGTTTGTTTGTGCCATCTTCATGTCAGTGGGACTTGTGCTGCT

   392
    |
E  *
GAATAA
```

Figure 4.33: Diagram to show the normal reading frame (top panel) in comparison with the introduction of a termination signal (bottom panel) when translation proceeds after the deletion of two nucleotides (red letters) at codon 291. The bold letters indicate the affected amino acids.

4.4.2 Population Screening of Unreported Mutations Found in the *AGL* Gene

To further lend support for the involvement of these unreported mutations as causative of the respective disorder, panels of unrelated healthy control individuals were screened for their presence. This is done on the premise that if these mutations are neutral polymorphisms, there is a higher likelihood that they will be present in the normal population at a frequency of 1 % or higher, ≥ 1 %. Due to time limitations, this part of the study onwards was focused on mutations in the *AGL* gene, causing GSD III.

Assays to detect mutant alleles were designed either using restriction enzyme digestion or TaqMan® SNP genotyping (when a suitable restriction enzyme assay is not available) (Table 4.12). A total of 150 control individuals were included as representative of the healthy population (Malay *n*=50, Chinese *n*=50, and Indian *n*=50).

Table 4.12: Assays to determine the presence of unreported mutations in healthy control samples.

Mutation	Strategy
c.1423+1G>T	TaqMan® SNP genotyping assay
c.2914_2915delAA	Restriction enzyme: Cfr10I (BsrFI)
c.3814_3815delAG	TaqMan® SNP genotyping assay
c.4333T>G	Restriction enzyme: Sau3AI (Bsp143I/BfuCI)
c.4490G>A	Restriction enzyme: StyI *
	TaqMan® SNP genotyping assay
c.4531_4534delTGTC	Restriction enzyme: HpyCH4III

*Note: Restriction enzyme assay produced ambiguous result. A TaqMan® SNP genotyping assay was then carried out for definitive results.

a) c.1423+1G>T

Presence of the c.1423+1G>T mutation could not be assayed using restriction enzymes as the sequence change does not result in creation or disruption of any restriction site. As such, TaqMan® SNP genotyping assay was designed and analysis was carried out for patient carrying the mutation and 150 healthy control individuals.

An example of the output (allele discrimination plot) is given in Figure 4.34, where individuals genotyped as homozygous for the normal allele is represented by red dots, individuals homozygous for the mutant allele (if present) is represented by blue dots, while heterozygous individuals carrying both alleles are represented by green dots. Non-template control reactions are represented by black dots.

As shown in Figure 4.34, only one green dot can be seen which represent the sample for Patient 7 who is heterozygous for this mutation. All other samples are represented by the cluster of red dots, which represent homozygous individuals from the healthy control group. There was no blue signal (homozygous mutant) detected in this analysis.

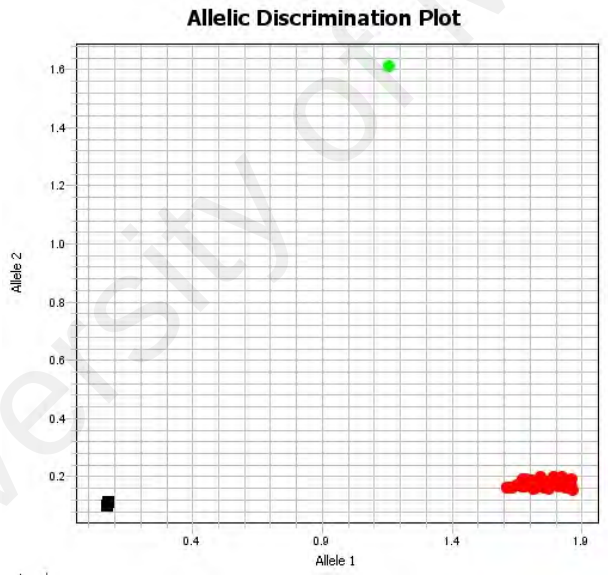


Figure 4.34: Allelic discrimination plot of TaqMan® SNP genotyping assay on unreported c.1423+1G>T mutation. Green dot: heterozygous mutant allele; red dots: homozygous normal allele; black squares: non-template control (water).

b) c.2914_2915delAA

Presence of the c.2914_2915delAA mutation was assayed using Cfr10I (BsrFI) restriction endonuclease. This enzyme cuts amplicons carrying the normal sequence at 5'-R↓CCGGY-3', generating fragments with molecular weights of 180 bp and 493 bp. The mutation abolishes the RE cut site, leaving the fragment intact at 673 bp.

Following enzyme digestion and subsequent electrophoresis, Patient 1 was found to be heterozygous for this mutation when three bands were observed (180 bp, 493 bp and 673 bp) as shown in Figure 4.35. The assay also showed that the patient's father is the mutation carrier as their fragment's patterns are similar. A similar assay carried out on 150 healthy control samples showed that none of them carried this mutation.

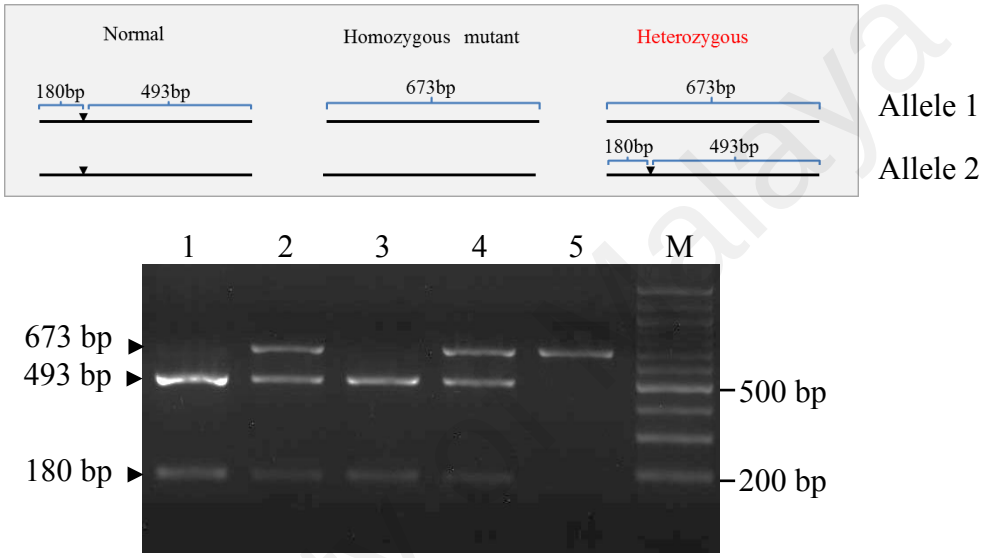


Figure 4.35: Agarose gel electrophoresis of Cfr10I (BsrFI) digestion on PCR amplified genomic DNA from patient carrying the c.2914_2915delAA mutation and control sample. Lane 1: Control sample, Lane 2: Patient's sample, Lane 3: Mother's sample, Lane 4: Father's sample (carrier), Lane 5: Uncut control (No enzyme), Lane M: Marker (100 bp).

c) c.3814_3815delAG

Presence of the c.3814_3815delAG mutation could not be assayed using restriction enzymes as the sequence change does not result in creation or disruption of any restriction site. As such, TaqMan® SNP genotyping assays were designed and analysis was carried out for patient carrying the mutation and 150 healthy control individuals. An example of the output (allele discrimination plot) is given in Figure 4.36, where individuals genotyped as homozygous for the normal allele is represented by red dots, individuals homozygous for the mutant allele is represented by

blue dots, while heterozygous individuals carrying both alleles (if present) are represented by green dots. Non-template control reactions are represented by black dots.

As shown in Figure 4.36, only one blue dot can be seen which represent the sample for Patient 10 who is homozygous for this mutation. All other samples are represented by the cluster of red dots, which represent homozygous individuals from the healthy control group. There was no green signal (heterozygous mutant) detected in this analysis.

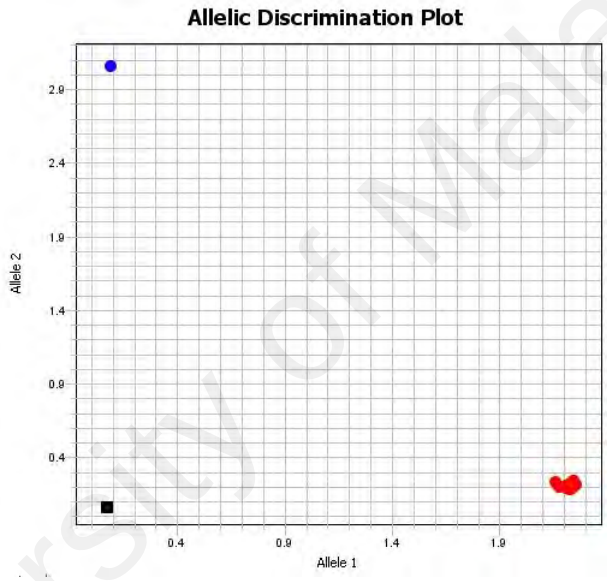


Figure 4.36: Allelic discrimination plot of TaqMan® SNP genotyping assay on unreported c.3814_3815delTGTC mutation. Blue dot: homozygous mutant allele; red dots: homozygous normal allele; black square: non-template control (water).

d) c.4333T>G

Presence of the c.4333T>G mutation was assayed using Sau3AI (Bsp143I/BfuCI) restriction endonuclease. This enzyme cuts amplicons carrying the normal sequence at 5'-↓GATC-3', generating fragments with molecular weight of 101 bp and 487 bp. The mutation introduces an additional cutting site, giving three fragments of 101 bp, 113 bp and 374 bp in length respectively.

Following enzyme digestion and subsequent electrophoresis, Patient 2 and Patient 3 (siblings) were found to be heterozygous for this mutation (Figure 4.37). It was also found that the 101 bp and 113 bp fragments were not distinguishable in the gel image. However, this does not represent a problem because the presence of c.4333T>G mutation is indicated by the occurrence of a 374 bp fragment. A similar assay carried out on 150 healthy control samples showed that none of them carried this mutation.

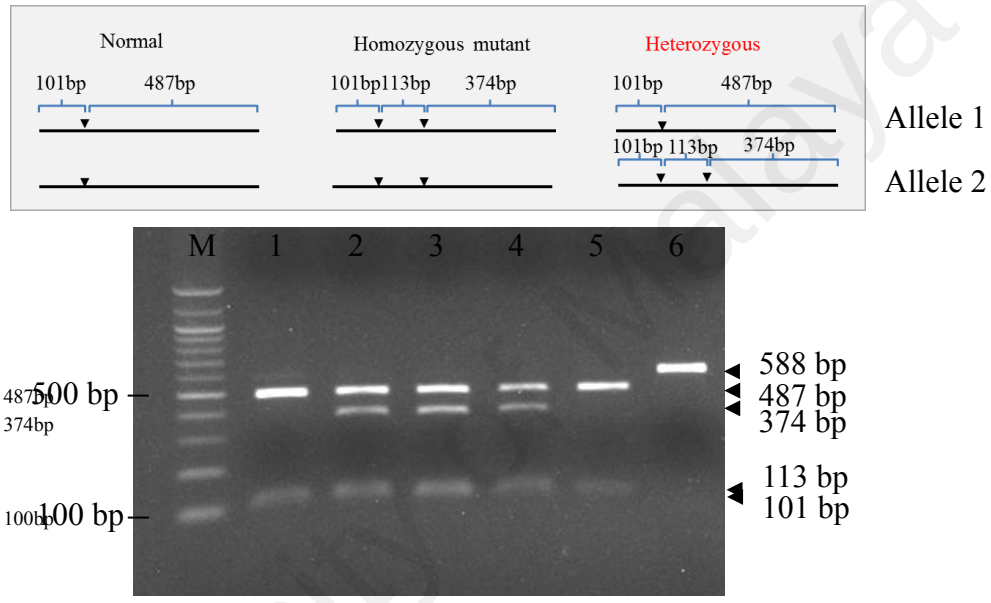


Figure 4.37: Agarose gel electrophoresis of Sau3AI (Bsp143I/BfuCI) digestion on PCR amplified genomic DNA from patient carrying the c.4333T>G mutation and control sample. Lane M: Marker (100 bp), Lane 1: Control sample, Lanes 2-3: Patients' samples (siblings), Lane 4: Mother's sample (carrier), Lane 5: Father's sample, Lane 6: Uncut control (No enzyme).

e) c.4490G>A

Presence of the c.4490G>A mutation was assayed using StyI restriction endonuclease. This enzyme cuts amplicons carrying the normal sequence at 5'-C↓CWWGG-3', generating fragments with molecular weights of 76 bp and 259 bp. The mutation abolishes the RE cut site, leaving the fragment intact at 335 bp.

Following enzyme digestion and subsequent electrophoresis, Patient 4 was found to be homozygous for this mutation when a single band was observed (335 bp) as shown in Figure 4.38. Unexpectedly, restriction digestion of healthy control samples

produced a pattern comprising three bands (76 bp, 259 bp and 335 bp respectively), suggesting that most, if not all of them are heterozygous carriers of the mutation. Another possibility of this result could be due to inefficiency of the enzyme activity, and presence of the intact 335 bp band could be the result of incomplete digestion. Due to this ambiguity, population screening of this mutation was repeated using the TaqMan® SNP genotyping assay method. The allele discrimination plot (Figure 4.39) revealed that there is one homozygous mutant sample (Patient 4), while the rest which are healthy control samples, are in actual fact homozygous normal. Two mock heterozygous samples were included in this assay. Mock samples were prepared when aliquots of samples known to be homozygous for the normal allele were mixed with aliquots of samples known to be homozygous for the mutant allele. These mock samples can be seen as two green dots in the plot.

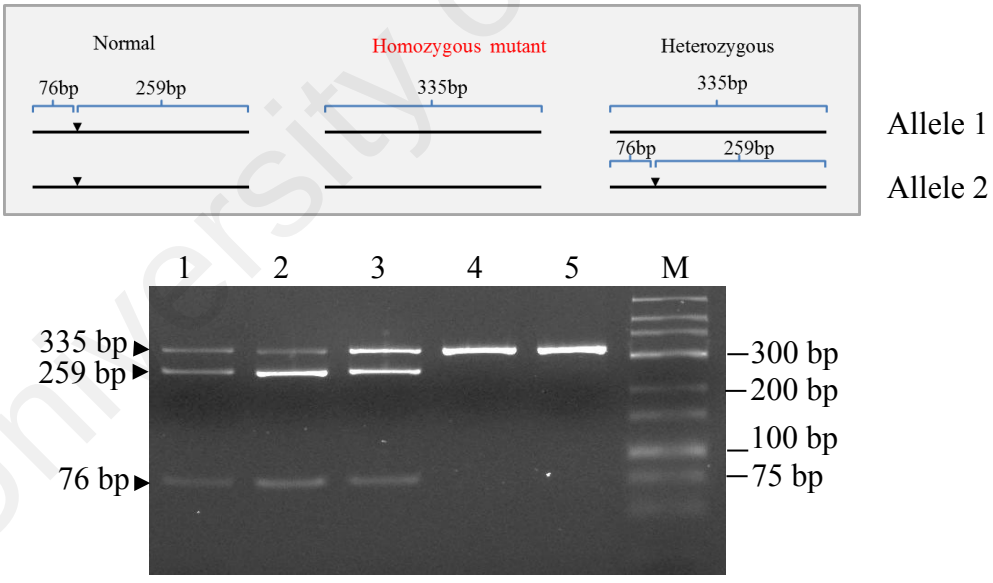


Figure 4.38: Agarose gel electrophoresis of StyI digestion on PCR amplified genomic DNA from patient carrying the c.4490G>A mutation and control sample. Lane 1: Control sample, Lane 2: Control sample, Lane 3: Mock heterozygous sample, Lane 4: Patient's sample, Lane 5: Uncut control (No enzyme), Lane M: Marker (25 bp).

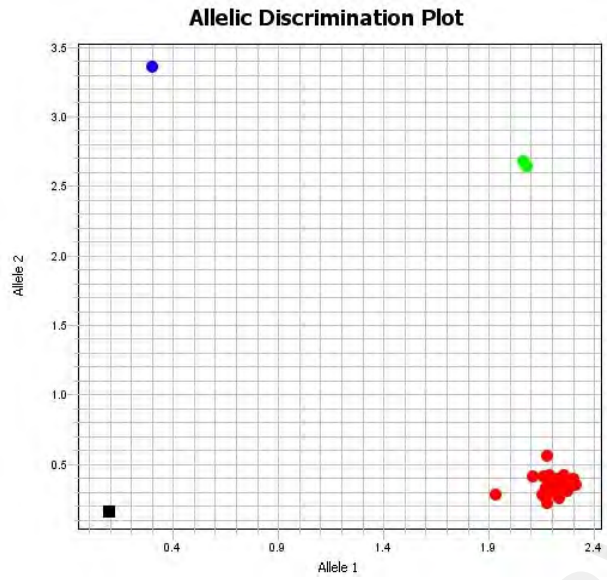


Figure 4.39: Allelic discrimination plot of TaqMan® SNP genotyping assay on unreported c.4490G>A mutation. Blue dot: homozygous mutant allele; green dots: heterozygous mutant allele, red dots: homozygous normal allele; black square: non-template control (water).

f) c.4531_4534delTGTC

Presence of the c.4531_4534delTGTC mutation was assayed using HpyCH4III restriction endonuclease. This enzyme cuts amplicons carrying the normal sequence at 5'-ACN↓GT-3', generating fragments with molecular weights of 121 bp and 214 bp. The mutation abolishes the RE cut site, leaving the fragment intact at 335 bp.

Following enzyme digestion and subsequent electrophoresis, both Patient 2 and Patient 3 (siblings) were found to be heterozygous for this mutation when three bands were observed (121 bp, 214 bp, and 335 bp) as shown in Figure 4.40. The assay also showed that their father as the mutation carrier as similar patterns were observed. A similar assay carried out on 150 healthy control samples showed that none of them carried this mutation.

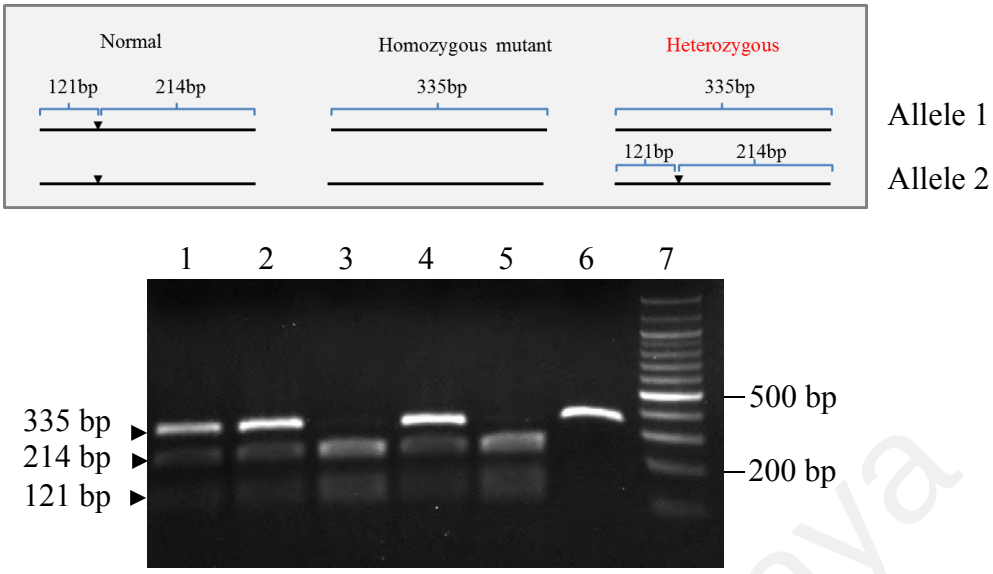


Figure 4.40: Agarose gel electrophoresis of HpyCH4III digestion on PCR amplified genomic DNA from patient carrying the c.4531_4534delTGTC mutation and control sample. Lanes 1-2: Patients' samples (siblings), Lane 3: Mother's sample, Lane 4: Father's sample, Lane 5: Control sample, Lane 6: Uncut control (No enzyme), Lane M: Marker (100 bp).

4.4.3 Assessment of the Effects of Mutations on Normal Protein Function

To further explore the effect of unreported mutations on protein function, two types of experimental work were carried out which include mRNA splicing assay (section 4.4.3.1) and enzyme functional assay on mutant *AGL* (section 4.4.3.2). Splicing assay was carried out on the mutation c.1423+1G>T while constructs carrying the other mutations (c.2914_2915delAA, c.3814_3815delAG, c.4333T>G, c.4490G>A and c.4531_4534delTGTC) were created to be subsequently used in functional enzyme assays.

4.4.3.1 Splicing Assay

One of the mutations found in this study is the c.2681+1G>A, which is a G to A substitution in the first base of intron 21. Usually mutations located within the exon/intron boundary causes abnormal splicing, and indeed this mutation was shown to result in skipping of exon 21 from the mature transcript (Hadjigeorgiou et al., 1999). A

control splicing experiment assaying the c.2681+1G>A was initially carried out to validate the report findings. RT-PCR assay were designed to amplify across several exons and introns so that any change in splicing activity should result in a shift from the expected amplicon length (Figure 4.41).

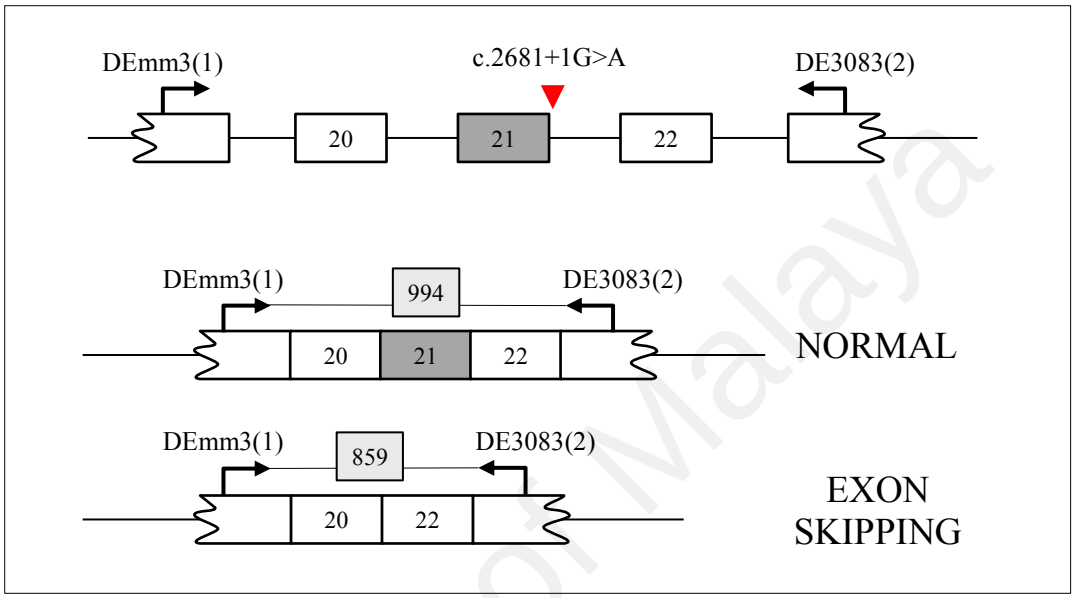


Figure 4.41: The panel shows the position of mutation c.2681+1G>A in reference to exons 20, 21 and 22. Boxes represent the exons, black line the introns, black arrows indicate the RT-PCR primers; forward primer: DEmm3(1), reverse primer: DE3083(2).

RT-PCR for mutation c.2681+1G>A analysis in Patient 7 produced two bands corresponding to the normal mRNA (994 bp) and to the abnormally spliced version (859 bp) respectively, as shown in Figure 4.42. Further analysis also showed that this mutation was inherited from the mother as the band representing the spliced version was present. Both the fragments were sequenced and analysed for verification. Sequencing product from normal sample showed that exon 21 was transcribed (Figure 4.43a), while sequencing product from patient cDNA revealed that exon 21 was completely missing (Figure 4.43b).

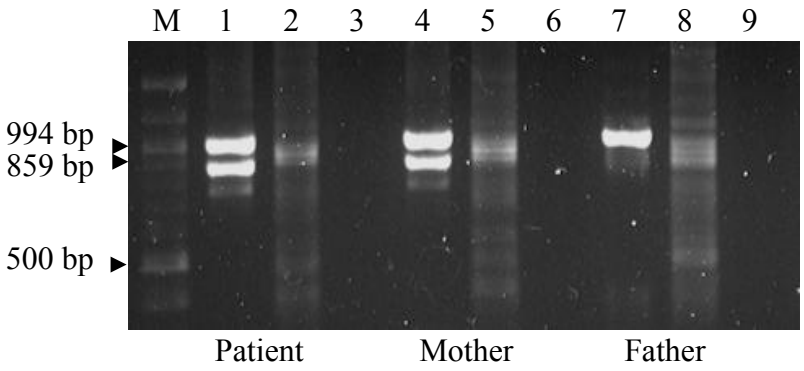


Figure 4.42: Agarose gel electrophoresis of RT-PCR product for c.2681+1G>A splicing assay. Lane M: RNA marker (100 bp), Lane 1: Patient's sample amplified with Superscript polymerase, Lane 4: Mother's sample amplified with Superscript polymerase, Lane 7: Father's sample amplified with Superscript polymerase, Lane 2, 5, 8: Positive control (without superscript), Lane 3, 6, 9: Negative control (without RNA).

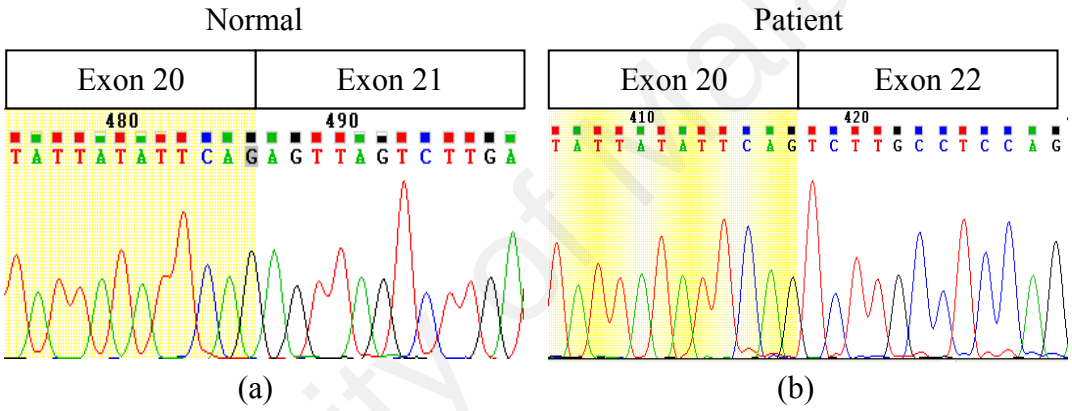


Figure 4.43: Partial sequencing chromatograms of RT-PCR products from normal individual and patient sample. Results showing the sequences of (a) exon 20 and exon 21 (b) exon 20 and exon 22.

In this study, another exon/intron boundary mutation was also identified, which is the c.1423+1G>T involving the first base of intron 12. Prediction analysis using the HSF software suggested that this mutation will lead to the disruption of the splicing motif (see section 4.4.1, Table 4.11). To confirm this prediction, a splicing experiment was designed to assay changes of splicing behaviour resulting from this mutation. Similar to the previous assay, RT-PCR primers were designed to amplify across several exons and introns so that any change in splicing activity should result in a shift from the expected amplicon length (Figures 4.44).

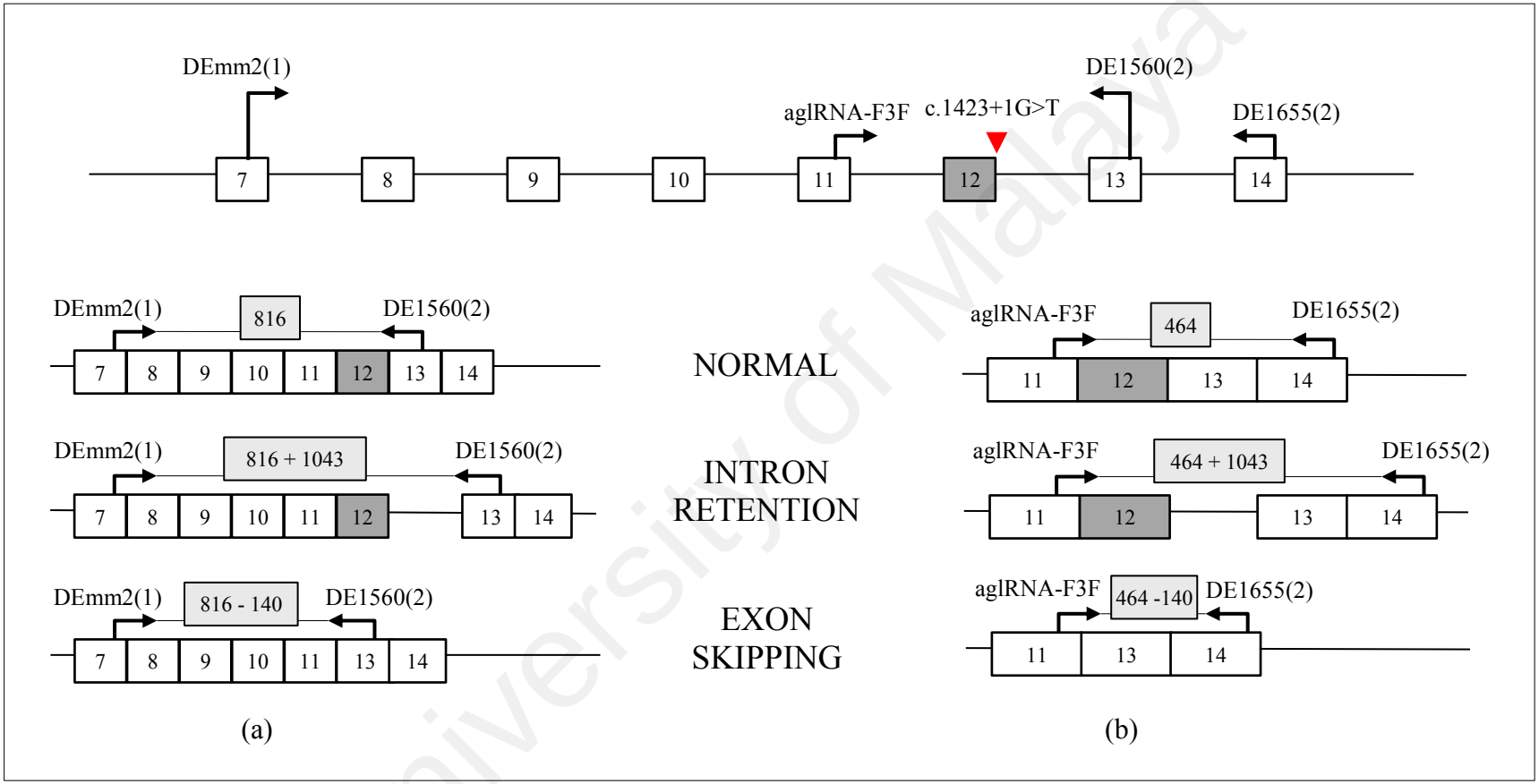


Figure 4.44: The panel shows the position of mutation c.1423+1G>T in reference to exons 7-14. Boxes represents the exons, black line the introns, black arrows indicate the RT-PCR primers; (a) forward primer: DEmm2(1), reverse primer: DE1560(2) and (b) forward primer: aglRNA-F3F, reverse primer: DE1655(2).

RT-PCR results for c.1423+1G>T also showed splicing aberration. However unlike the reported c.2681+1G>A mutation, a longer amplification fragment was observed instead of shorter. Further scrutiny showed that the length corresponded to the retention of intron 12 (sizing 1043 bp) thus yielding a fragment of 1859 bp in length (Figure 4.45a). Unfortunately because this result was not expected and due to the fact that the largest size reference in that particular electrophoresis experiment was 1500 bp, an additional confirmatory analysis was carried out. A second RT-PCR primer pair was designed with an expected product length of 1507 bp if intron 12 is retained (Figure 4.44b). Indeed, following RT-PCR and electrophoresis, a 1507 bp band was observed (Figure 4.45b) thus providing confirmation that intron retention has occurred.

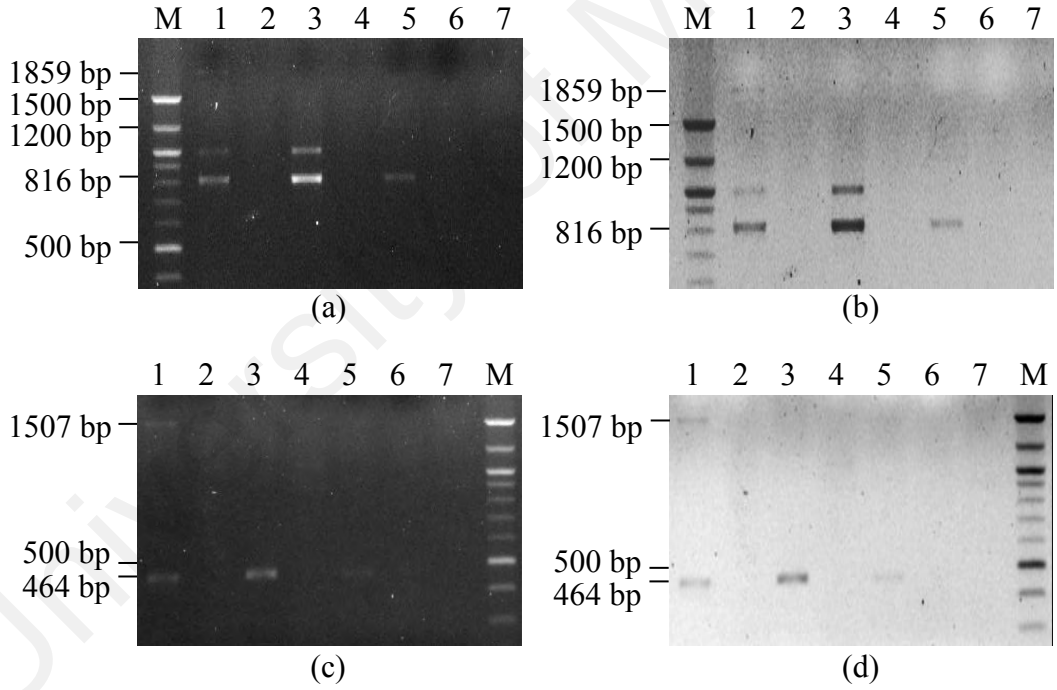


Figure 4.42: Agarose gel electrophoresis of RT-PCR products for mutation c.1423+1G>T splicing assay. (a) RT-PCR product of primer set DEmm2(1) and DE1560(2), (b) inverted and enhanced image of (a) to confirm the presence of band 1859 bp in Lane 1 and lack of the band in other lanes, (c) RT-PCR product of primer set aglRNA-F3F and DE1655, (d) inverted and enhanced image of (c) to confirm the presence of band 1507 bp in Lane 1 and lack of the band in other lanes. Lane 1: Patient's sample amplified with Superscript polymerase, Lane 3: Mother's sample amplified with Superscript polymerase, Lane 5: Father's sample amplified with Superscript polymerase, Lane 2, 4, 6: Positive control (without superscript), Lane 7: Negative control (without RNA), Lane M: RNA marker (100 bp).

4.4.3.2 Functional Study

The other five unreported mutations found in the *AGL* gene (c.2914_2915delAA, c.3814_3815delAG, c.4333T>G, c.4490G>A and c.4531_4534delTGTC) occurred within the exonic region. As such it was postulated that these mutations would directly result in impairment of GDE function. While the fate of mutated GDE may vary depending on the nature of the mutation, it is possible that impairment of function may be variable too. To further understand the effect of each mutation, functional studies were attempted to elucidate the possible changes in protein activity resulting from these mutations.

a) Cloning and Site-Directed Mutagenesis of *AGL*

Expression constructs carrying individual unreported *AGL* mutations were created using commercially available mammalian expression vector *AGL*-pCMV6 (OriGene Technologies, USA) containing wild-type *AGL* sequence. Following sequence verification, *AGL*-pCMV6 was used as template for site-directed mutagenesis (SDM) PCR. Information on the affected nucleotides and substituted codon for each mutation is presented in Table 4.13 and the gel electrophoresis images of the SDM PCR products are presented in Figure 4.46.

Table 4.13: Mutant *AGL* constructs obtained through site-directed mutagenesis.

Construct ID	Mutation	Codon	Amino acid change
SDM_972	c. 2914_2915delAA	AAC>CCG	Asparagine>Proline
SDM_1272	c.3814_3815delAG	AGA>AGG	Arginine>Arginine
SDM_1445	c.4333T>G	TAT>GAT	Tyrosine>Aspartic acid
SDM_1497	c.4490G>A	TGG>TAG	Tryptophan>Stop
SDM_1511	c.4531_4534delTGTC	TGT>CTT	Cysteine>Leucine

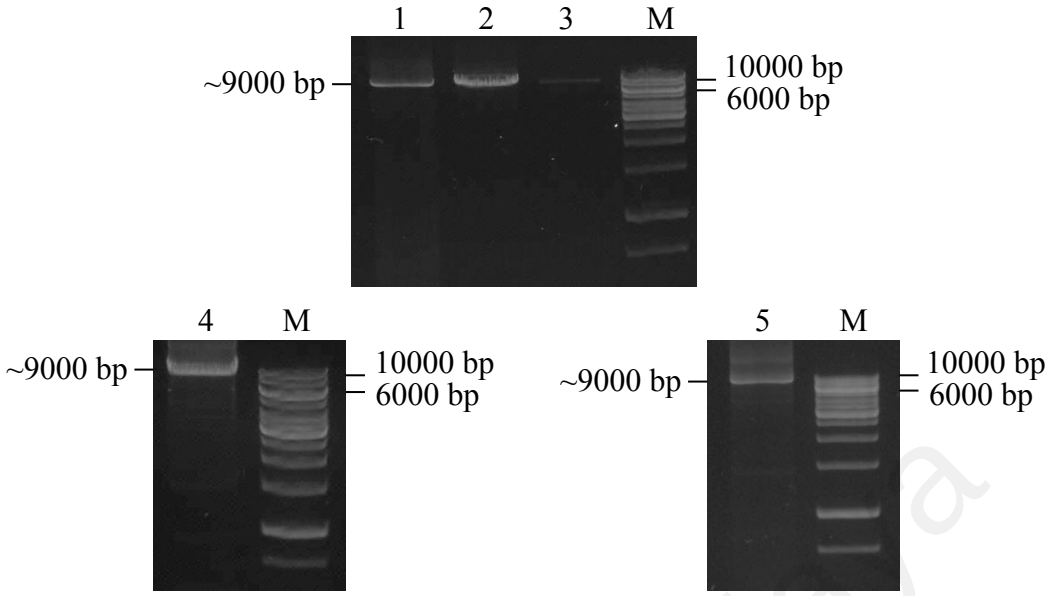


Figure 4.43: Agarose gel electrophoresis of site-directed mutagenesis PCR product. Lane 1: SDM_972, Lane 2: SDM_1272, Lane 3: SDM_1497, Lane 4: SDM_1511a, Lane 5: SDM_1511b, Lane M: Marker (1 kb).

Products of SDM were subsequently ligated and transformed into competent *E. coli* strain TOP10. Colonies generated were selected and size-verified via colony PCR. Several combinations of insert primers (as previously listed in Table 3.6) were used. Candidate clones for each construct were later extracted and sequenced for their full length. The result for each construct is presented below.

i. Mutation c.2914_2915delAA (SDM_972)

Transformation of SDM_972 product into *E.coli* generated multiple colonies. Clones carrying this insert was PCR verified with insert forward primer DE1996(1) and insert reverse primer DEmm3(2). Gel electrophoresis of the PCR products revealed amplicons of variable size. However, only clones which produced PCR bands with expected size of ~1152 bp (Figure 4.47a) were chosen for subsequent plasmid extraction (Figure 4.47b) and used for sequencing. Sequence analysis showed a 2-bp deletion in the sample which correlates with the mutation c.2914_2915delAA (Figure 4.47c).

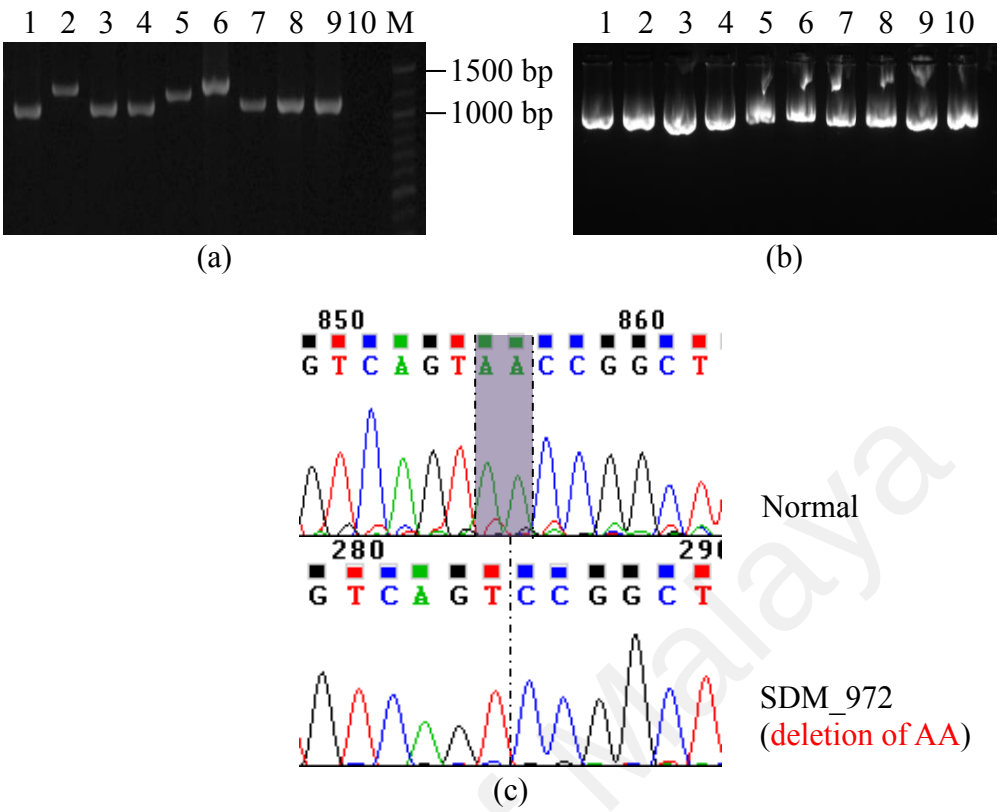


Figure 4.44: Agarose gel electrophoresis of (a) colony PCR for SDM_972 clones bands of varying sizes; Lanes 1-9: Transformed SDM_972 clones, Lane 10: Negative control (no template), Lane M: Marker (100 bp ladder) (b) extracted c.2914_2914delAA mutant plasmids; Lanes 1-10; representative plasmids and (c) partial sequencing chromatogram of the clone carrying c.2914_2914delAA mutation. The highlighted letters indicate deleted nucleotides.

ii. Mutation c.3814_3815delAG (SDM_1272)

Transformation of SDM_1272 product into *E.coli* generated multiple colonies. Clones carrying this insert were PCR verified with insert forward primer aglRNA_28F and insert reverse primer aglRNA_32R. Gel electrophoresis of the PCR products revealed amplicons of variable size. However, only clones which produced PCR bands with the expected size of 622 bp (Figure 4.48a) were chosen for plasmid extraction (Figure 4.48b) and used for sequencing. Sequence analysis revealed a 2-bp deletion in the sample which correlates with the mutation c.3814_3815delAG (Figure 4.48c).

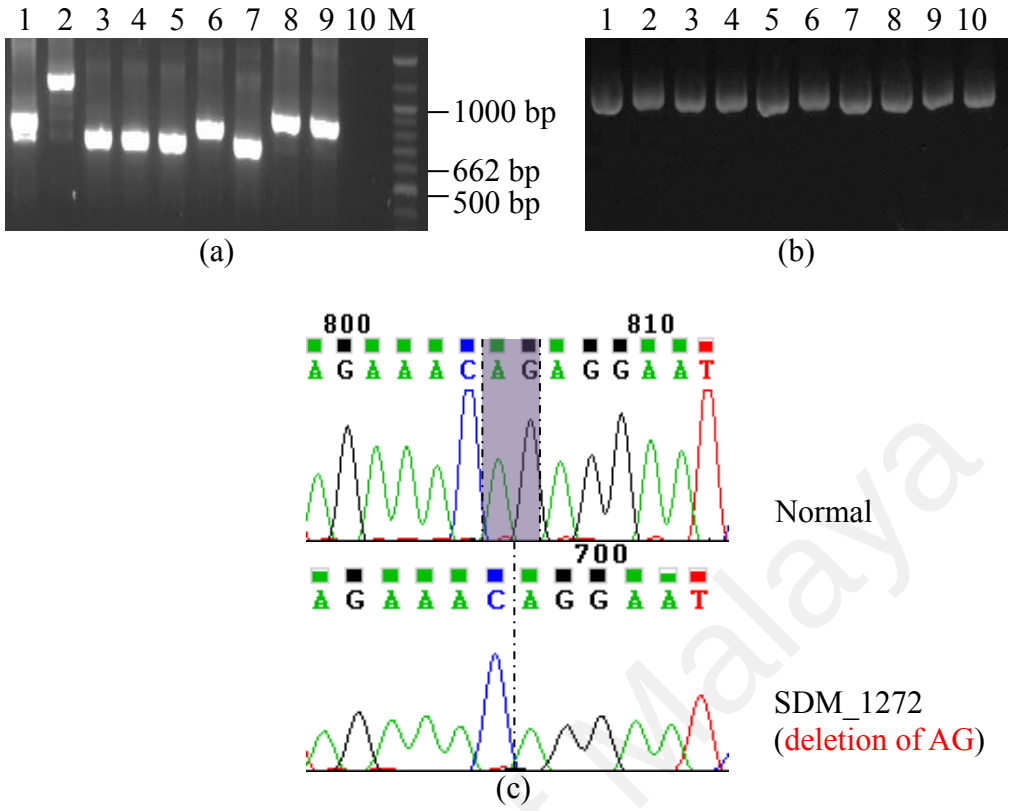


Figure 4.45: Agarose gel electrophoresis of (a) colony PCR for SDM_1272 clones; Lanes 1-9: Transformed SDM_1272 clones, Lane 10: Negative control (no template), Lane M: Marker (100 bp ladder) (b) extracted c.3814_3815delAG mutant plasmids. Lane 1-10: representative plasmids and (c) partial sequencing chromatogram of the clone carrying c.3814_3815delAG mutation. The highlighted letters indicate deleted nucleotides.

iii. Mutation c.4333T>G (SDM_1445)

Transformation of SDM_1445 product into *E.coli* generated multiple colonies. Clones carrying this insert was PCR verified with insert forward primer DE1690(1) and insert reverse primer DEmm2(2). Gel electrophoresis of the PCR products revealed amplicons of variable size. However, only clones which produced PCR bands with the expected size of 662 bp (Figure 4.49a) were chosen for plasmid extraction (Figure 4.49b) and used for sequencing. Sequence analysis showed a T>G change in the sample which correlates with the mutation c.4333T>G (Figure 4.49c).

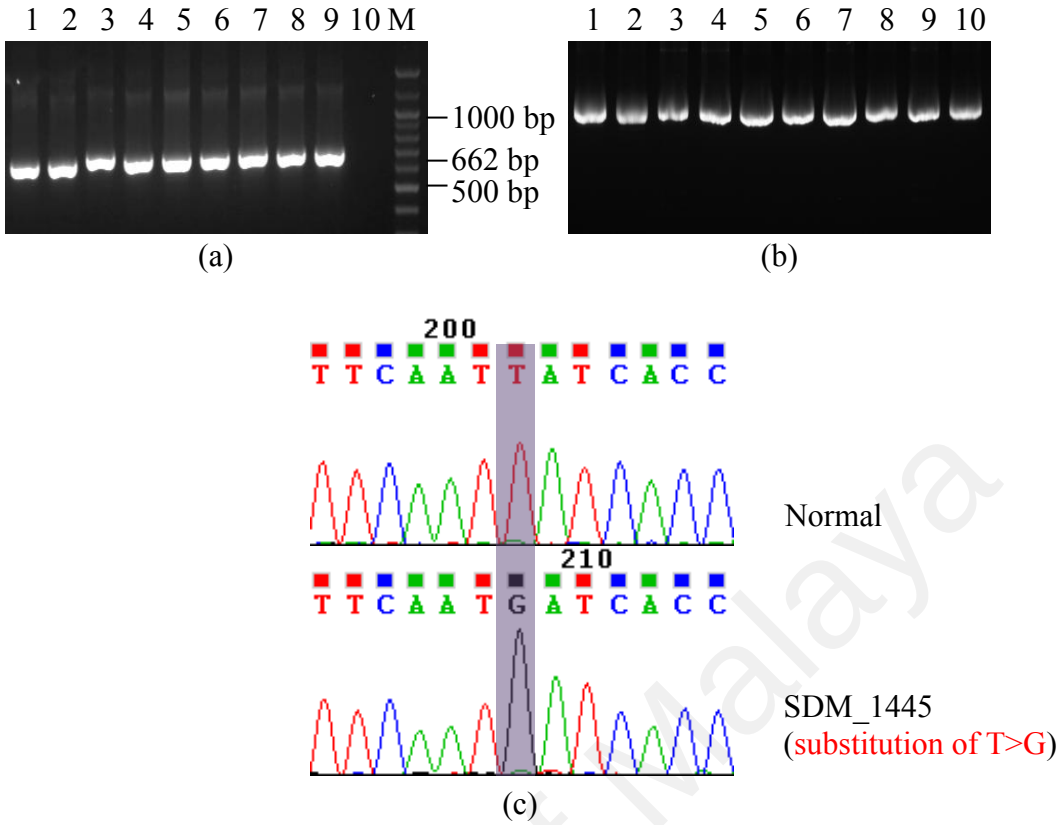


Figure 4.46: Agarose gel electrophoresis of (a) colony PCR for SDM_1445 clones; Lanes 1-9: Transformed SDM_1445 clone, Lane 10: Negative control (no template), Lane M: Marker (100 bp ladder) (b) extracted c.4333T>G mutant plasmids; Lane 1-10: representative plasmids and (c) partial sequencing chromatogram of the clone carrying c.4333T>G mutation. The highlighted letters indicate affected nucleotides.

iv. Mutation c.4490G>A (SDM_1497)

Transformation of SDM_1497 product into *E.coli* generated multiple colonies. Clones carrying this insert was PCR verified with insert forward primer DE1690(1) and insert reverse primer DEmm2(2). Gel electrophoresis of the PCR products revealed amplicons of similar sizes of 662 bp (Figure 4.50a). Selected clones were chosen for subsequent plasmid extraction (Figure 4.50b) and used for sequencing. Sequence analysis showed a G>A change in the sample which correlates with the mutation c.4490G>A (Figure 4.50c).

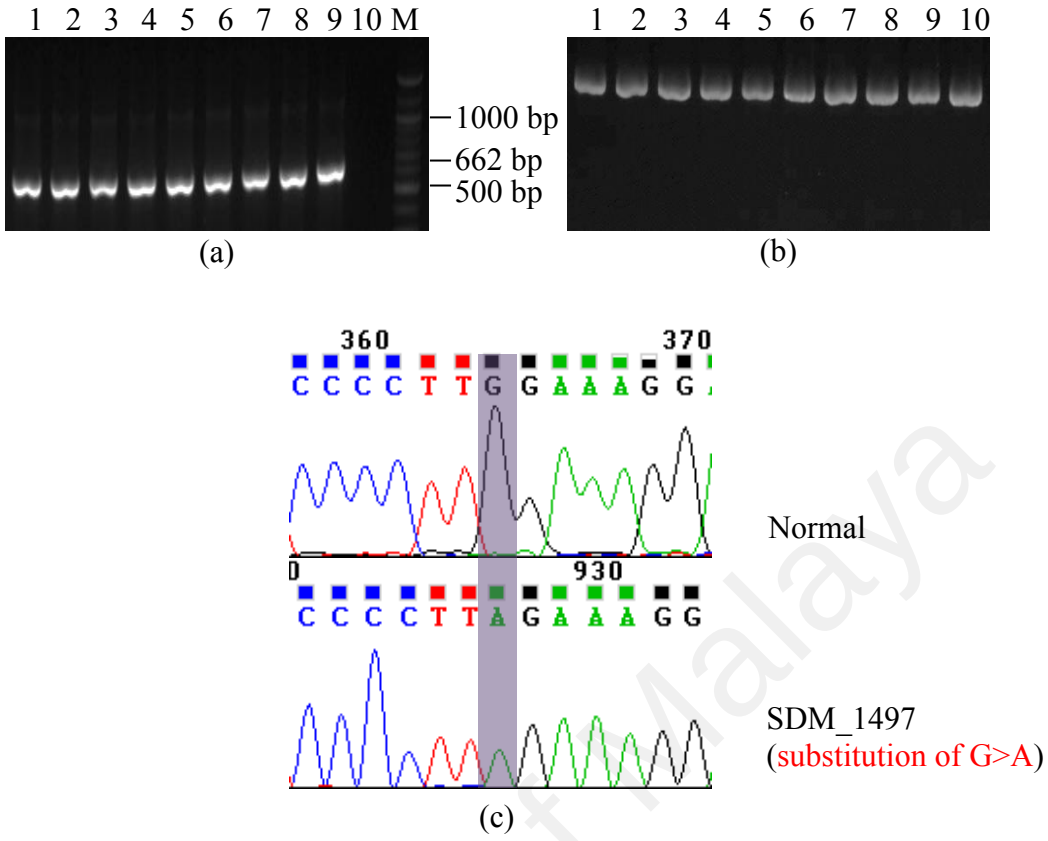


Figure 4.47: Agarose gel electrophoresis of (a) colony PCR for SDM_1497 clones; Lanes 1-9: Transformed SDM_1497 clones, Lane 10: Negative control (no template), Lane M: Marker (100 bp ladder) (b) extracted c.4490G>A mutant plasmids; Lanes 1-10: representative plasmids and (c) partial sequencing chromatograms of the clone carrying c.4490G>A mutation. The highlighted letters indicate affected nucleotides.

v. Mutation c.4531_4534delTGTC (SDM_1511)

The construction of this mutant was achieved in two steps. SDM was first carried out on wild type *AGL* construct to create a 4-bp deletion mutant (SDM_1511a) and another SDM was carried out on the same plasmid to create the readthrough region (SDM_1511b). Results for each step are presented in the following.

- SDM_1511a

Transformation of SDM_1511a product into *E.coli* generated multiple colonies. Clones carrying this insert was PCR verified with insert forward primer DE1996(1) and insert reverse primer DEmm3(2). Gel electrophoresis of the PCR products revealed amplicons of similar sizes of 662 bp (Figure 4.51a). Selected

clones were chosen for extraction (Figure 4.51b) and used for sequencing. Sequence analysis showed a homozygous 4-bp deletion in the sample which correlates with the mutation c. 4531_4534delTGTC (Figure 4.51c).

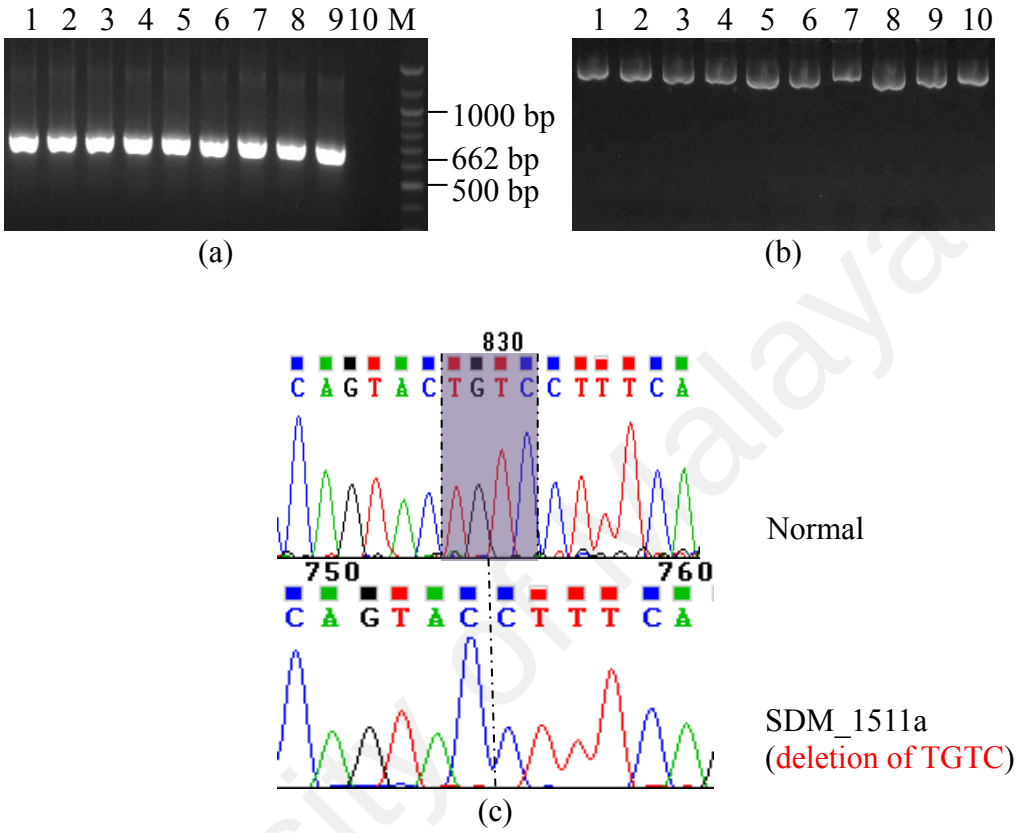


Figure 4.48: Agarose gel electrophoresis of (a) colony PCR for SDM_1511a clones; Lanes 1-9: Transformed SDM_1511a clones, Lane 10: Negative control (no template), Lane M: Marker (100 bp ladder) (b) extracted c.4531_4534delTGTC mutant plasmids; Lanes 1-10: representative plasmids and (c) partial sequencing chromatograms of the clone carrying c.4531_4534delTGTC mutation. The highlighted letters indicate deleted nucleotides.

- SDM_1511b

  Clone that had been confirmed to carry the 4-bp deletion of TGTC then subjected to the second SDM (SDM_1511b). Transformation of this SDM product into *E.coli* generated multiple colonies. Clones carrying this insert was PCR verified with insert forward primer DE1690 (1) and insert reverse primer DEmm2(2). Gel electrophoresis of the PCR products revealed amplicons of variable sizes. However, only clones which produced PCR

bands with the expected size of 622 bp (Figure 4.52a) were chosen for plasmid extraction (Figure 4.52b) and used for sequencing. Sequence analysis showed a 40-bp insertion in sample which correlates with the readthrough region sequence generated by mutation c.4531_4534delTGTC (Figure 4.52c).

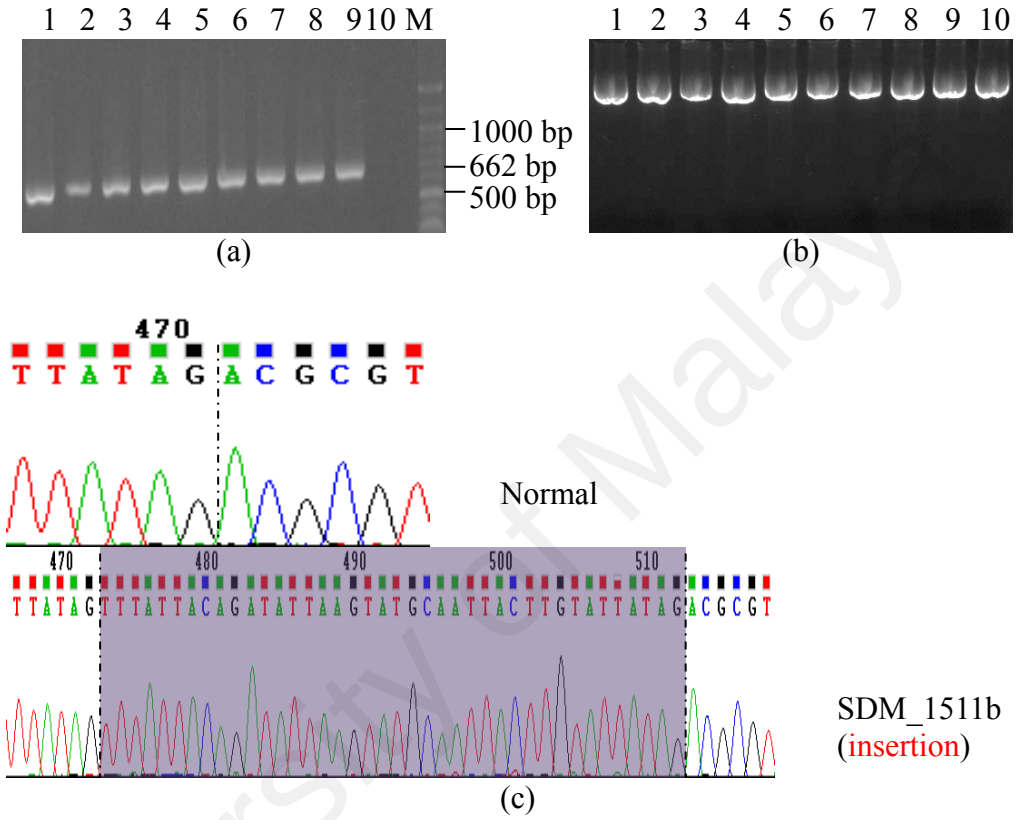


Figure 4.49: Agarose gel electrophoresis of (a) colony PCR for SDM_1511b clones; Lanes 1-9: Transformed SDM_1511b clones, Lane 10: Negative control (no template), Lane M: Marker (100 bp ladder) (b) extracted c.4531_4534delTGTC mutant plasmids; Lanes 1-10: representative plasmids and (c) partial sequencing chromatograms of the clone carrying c.4531_4534delTGTC mutation. The highlighted letters indicate inserted nucleotides.

b) Verification of Successful Transfection and Transcription

Clones that had been identified to carry the desired mutant construct were transfected into Human hepatocellular carcinoma (HepG2) cell line using calcium phosphate transfection method. Following stable selection using Geneticin (G418), formed colonies were isolated and harvested. Total DNA and RNA were extracted from the cells and used for PCR and RT-PCR. PCR was carried out to confirm successful

transfection and RT-PCR to confirm transcription (gene expression) of the mutant constructs. Both the PCR and RT-PCR products were then subjected to sequencing analysis. Results for each construct are presented in the following section.

i. WT

DNA sample obtained from cells transfected with the plasmid carrying WT *AGL* sequence (*AGL*-pCMV6) were PCR amplified using forward primer AGL1497F and reverse primer AGL1497R. Gel electrophoresis of the PCR product revealed a single band of expected size 399 bp (Figure 4.53a). RNA sample from the same cell was also subjected to RT-PCR using the same primer combination used in PCR reaction. Gel electrophoresis of the RT-PCR also revealed the amplification of a single band with expected size 399 bp (Figure 4.53b). Sequencing of both the PCR and RT-PCR products showed the presence of normal *AGL* sequence (Figure 4.53c).

ii. The SDM_972

DNA sample obtained from cells transfected with the plasmid carrying c.2914_2916delAA mutant were PCR amplified using forward primer AGL972F and reverse primer AGL972R. Gel electrophoresis of the PCR product revealed a single band of expected size 451 bp (Figure 4.54a). RNA sample from the same cell was also subjected to RT-PCR using the same primer combination used in PCR reaction. Gel electrophoresis of the RT-PCR also revealed the amplification of a single band with expected size 451 bp (Figure 4.54b). Sequencing of both the PCR and RT-PCR products showed the presence of peaks corresponding to 2-bp deletion of mutation c.2914_2915delAA and peaks corresponding to normal endogenous *AGL* sequence (Figure 4.54c).

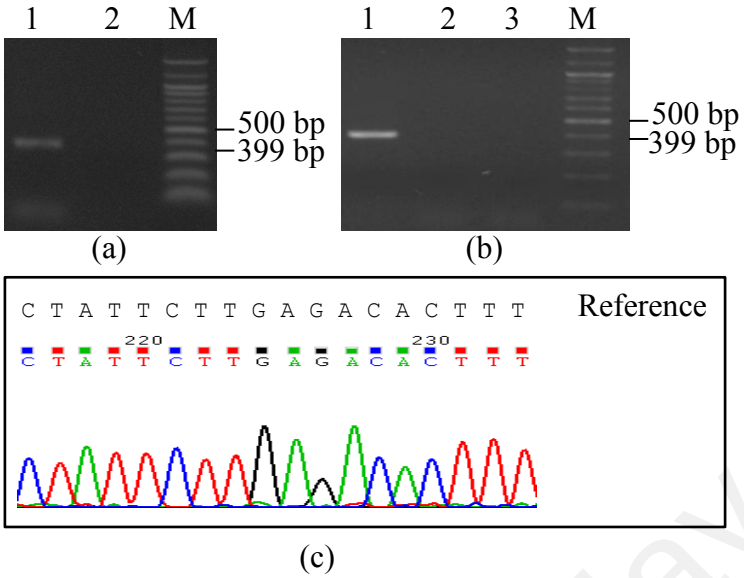


Figure 4.50: Agarose gel electrophoresis of (a) PCR product from cell sample carrying WT construct with expected size of 399 bp; Lane 1: Sample with WT construct, Lane 2: Negative control (No DNA template), Lane M: Molecular weight marker (100 bp ladder) (b) RT-PCR product; Lane 1: Sample amplified with Superscript polymerase, Lane 2: Positive control (without superscript), Lane 3: Negative control (without RNA), Lane M: Molecular weight marker (100 bp ladder). (c) Partial sequencing chromatogram showing normal *AGL* sequence obtained from the cell sample.

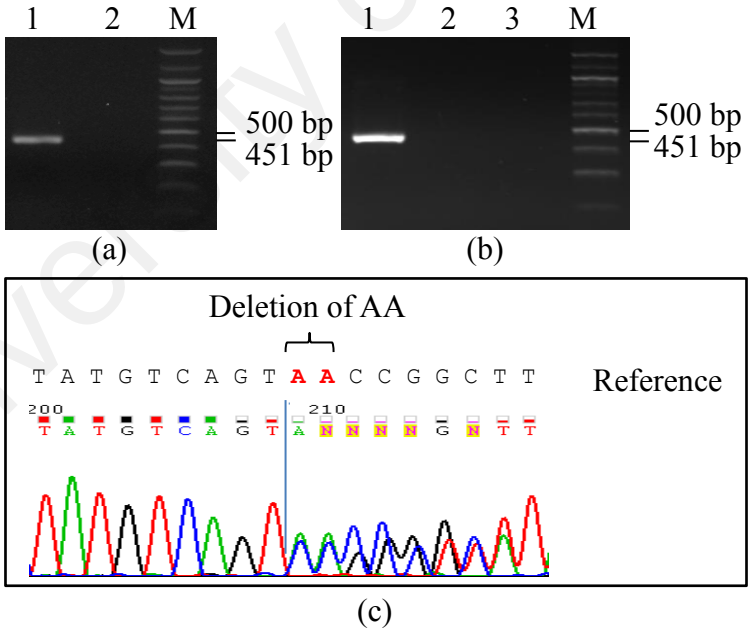


Figure 4.51: Agarose gel electrophoresis of (a) PCR product from cell sample carrying SDM_972 construct with expected size of 451 bp; Lane 1: Sample with SDM_972 construct, Lane 2: Negative control (No DNA template), Lane M: Molecular weight marker (100 bp ladder) (b) RT-PCR product; Lane 1: Sample amplified with Superscript polymerase, Lane 2: Positive control (without superscript), Lane 3: Negative control (without RNA), Lane M: Molecular weight marker (100 bp ladder) (c) Partial sequencing chromatogram showing a 2-bp deletion from the cell sample with SDM_972 construct. Red letters in the reference sequence indicate deleted nucleotides.

iii. The SDM_1272

DNA sample obtained from cells transfected with the plasmid carrying c.3814_3816delAG mutant were PCR amplified using forward primer AGL1272F and reverse primer AGL1272R. Gel electrophoresis of the PCR product revealed a single band of expected size 466 bp (Figure 4.55a). RNA sample from the same cell was also subjected to RT-PCR using the same primer combination used in PCR reaction. Gel electrophoresis of the RT-PCR also revealed the amplification of a single band with expected size 466 bp (Figure 4.55b). Sequencing of both the PCR and RT-PCR products showed the presence of peaks corresponding to 2-bp deletion of mutation c.3814_3816delAG and peaks corresponding to normal endogenous *AGL* sequence (Figure 4.55c).

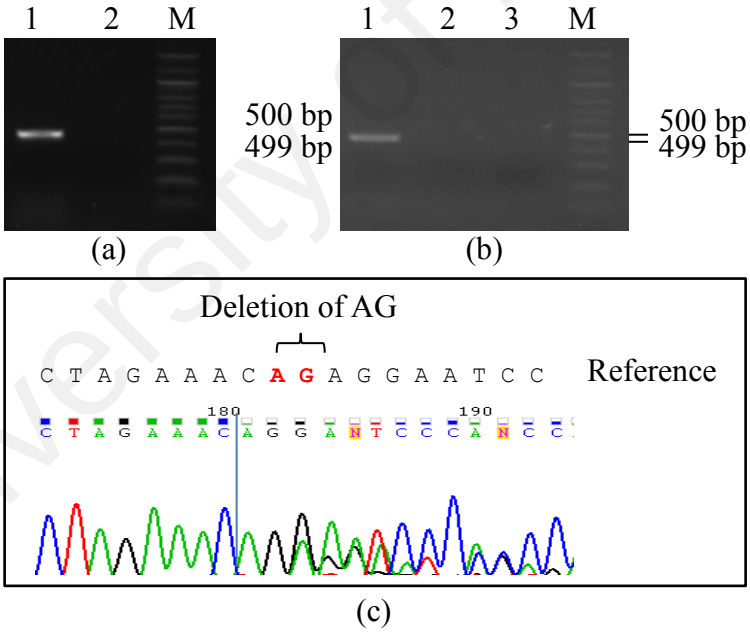


Figure 4.52: Agarose gel electrophoresis of (a) PCR product from cell sample carrying SDM_1272 construct; Lane 1: Sample with SDM_1272 construct, Lane 2: Negative control (No DNA template), Lane M: Molecular weight marker (100 bp ladder) (b) RT-PCR product; Lane 1: Sample amplified with Superscript polymerase, Lane 2: Positive control (without superscript), Lane 3: Negative control (without RNA), Lane M: Molecular weight marker (100 bp ladder) (c) Partial sequencing chromatogram showing a 2-bp deletion from the cell sample with SDM_1272 construct. Red letters in the reference sequence indicate deleted nucleotides.

iv. The SDM_1445

DNA sample obtained from cells transfected with the plasmid carrying c.4333T>G mutant were PCR amplified using forward primer AGL1445F and reverse primer AGL1445R. Gel electrophoresis of the PCR product revealed a single band of expected size 466 bp (Figure 4.56a). RNA sample from the same cell was also subjected to RT-PCR using the same primer combination used in PCR reaction. Gel electrophoresis of the RT-PCR also revealed the amplification of a single band with expected size 466 bp (Figure 4.56b). Sequencing of both the PCR and RT-PCR products showed the presence of two peaks corresponding to a mutant guanine and an endogenous thymine (Figure 4.56c).

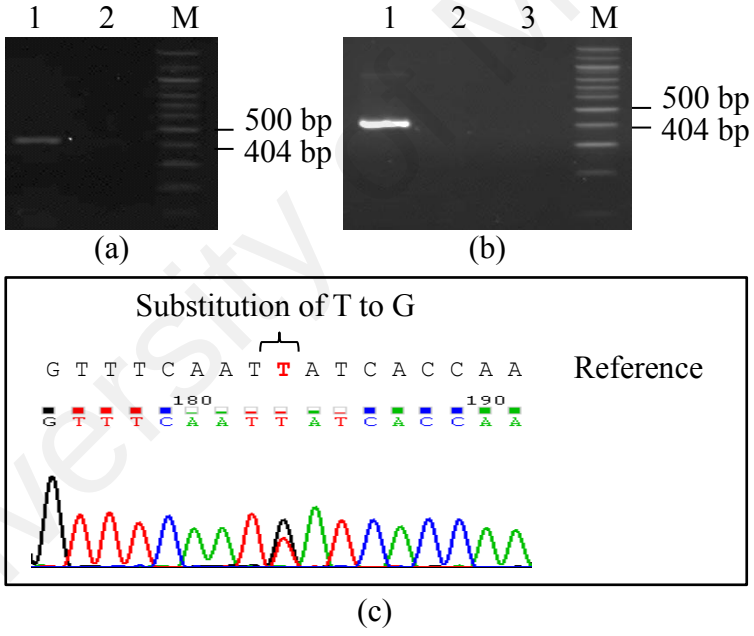


Figure 4.53: Agarose gel electrophoresis of (a) PCR product from cell sample carrying SDM_1445 construct with expected size of 404 bp; Lane 1: Sample with SDM_1445 construct, Lane 2: Negative control (No DNA template), Lane M: Molecular weight marker (100 bp ladder) (b) RT-PCR product; Lane 1: Sample amplified with Superscript polymerase, Lane 2: Positive control (without superscript), Lane 3: Negative control (without RNA), Lane M: Molecular weight marker (100 bp ladder) (c) Partial sequencing chromatogram showing a T to G substitution from the cell sample with SDM_1445 construct. Red letter in the reference sequence indicates substituted nucleotide.

v. The SDM_1497

DNA sample obtained from cells transfected with the plasmid carrying c.4490G>A mutant were PCR amplified using forward primer AGL1497F and reverse primer AGL1497R. Gel electrophoresis of the PCR product revealed presence of the band with expected size 399 bp (Figure 4.57a). RNA sample from the same cell was also subjected to RT-PCR using the same primer combination used in PCR reaction. Gel electrophoresis of the RT-PCR also revealed the amplification of a single band with the expected size of 399 bp (Figure 4.57b). Sequencing of both the PCR and RT-PCR products showed the presence of two peaks corresponding to a mutant adenine and an endogenous guanine (Figure 4.57c).

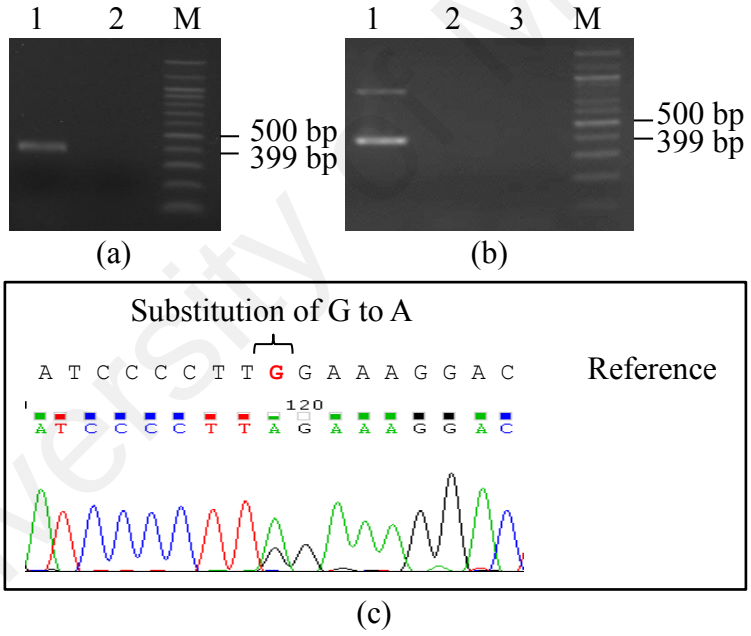


Figure 4.54: Agarose gel electrophoresis of (a) PCR product from cell sample carrying SDM_1497 construct with expected size of 399 bp; Lane 1: Sample with SDM_1497 construct, Lane 2: Negative control (No DNA template), Lane M: Molecular weight marker (100 bp ladder) (b) RT-PCR product, Lane 1: Sample amplified with Superscript polymerase, Lane 2: Positive control (without superscript), Lane 3: Negative control (without RNA), Lane M: Molecular weight marker (100 bp ladder) (c) Partial sequencing chromatogram showing a G to A substitution from the cell sample with SDM_1497 construct. Red letter in the reference sequence indicates substituted nucleotide.

vi. The SDM_1511

DNA sample obtained from cells transfected with the plasmid carrying c.4531_4534delTGTC mutant were PCR amplified using forward primer AGL1497F and reverse primer AGL1497R. Gel electrophoresis of the PCR product however revealed no amplification of the band with expected size 435 bp. RNA sample from the same cell was also subjected to RT-PCR using the same primer combination used in PCR reaction. Gel electrophoresis of the RT-PCR also revealed no amplification of the expected band.

c) Detection of Mutant Protein Using SDS-PAGE and Western Blotting

Proteins were harvested from cell colonies that had been identified to carry the desired mutant genes. Untransfected cells, cells transfected with empty vector and cells carrying the WT construct were also harvested. Cell lysate from these samples were then subjected to SDS-PAGE to detect the presence of GDE protein. The SDS-PAGE results showed multiple bands for all samples (Figure 4.58a). To determine the presence of GDE protein, these samples were then subjected to Western blotting. Initially, detection was carried out using C-terminal anti-*AGL* antibody. Mutant constructs were expected to generate bands at lower molecular weight (SDM_972: 112.2 kDa, SDM_1272: 146.8 kDa, SDM_1445: 174.6 kDa and SDM_1497: 170.5 kDa) compared to the WT. However, only the WT protein was detected with the expected size of 175 kDa (Figure 4.58b). Western blotting was then repeated using N-terminal anti-*AGL* antibody. However, a similar observation was obtained where only the wild type protein was detected (Figure 4.58c).

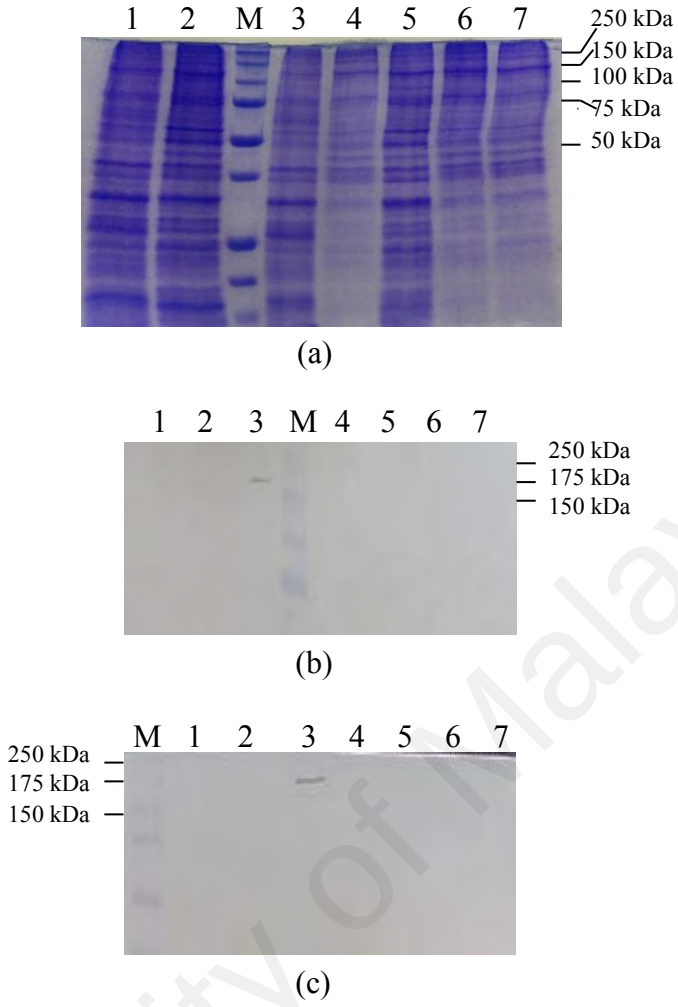


Figure 4.55: Examples of protein verification using (a) SDS-PAGE (b) Western blotting using C-terminal anti-*AGL* antibody and (c) Western blotting using N-terminal anti-*AGL* antibody. Lane 1: HepG2, Lane 2: Empty vector, Lane 3: WT, Lane 4: SDM_972, Lane 5: SDM_1272, Lane 6: SDM_1445, Lane 7: SDM_1497, Lane M: Protein marker.

CHAPTER 5: DISCUSSION

This chapter discusses the results obtained in this study and compares it with existing knowledge available in the literature. This chapter is organized as follows:-

Section 5.1 interpret and discusses the study findings on mutations identified in the *AGL* gene as well as other genes. Section 5.2 discusses the negative mutation identification in some of the patients.

5.1 Mutation Identification Through Molecular Analysis

The present study revolved around using molecular techniques to look for DNA lesions that could help explain the genetic cause of disorders in question. It has been mentioned that there are various types of GSDs and different types sometime show similar clinical symptoms that complicates accurate diagnosis (Burton, 1998; Mak et al., 2013). Table 5.1 for example illustrates that at least six different types of GSDs are presented with hypoglycaemia and almost all types share hepatomegaly as a common feature. Accurate diagnosis on the other hand is crucial as in various cases the course of treatment required by different types of GSDs and their prognosis vary (Adeva-Andany et al., 2016).

Achieving accurate diagnosis has been greatly facilitated by the advent of molecular diagnostics. In addition, molecular diagnostics data is not only useful for the patient, but also in the identification and determination of carrier status which would be valuable for genetic counselling purposes. Understanding the pathogenesis of disorders also contributes to the knowledge of the disease pathway, which is important in the effort to develop new strategies of treatment.

Table 5.1: Phenotype presentation of different types of GSDs.

Type	Subtype	Features
0		Hypoglycaemia
I	a	Hepatomegaly, hypoglycaemia, hyperuricemia, hyperlipidemia, short stature and gout
	b	Hepatomegaly, hypoglycaemia, hyperuricemia, gout, predisposing to infection due to neutropenia
II		Cardiomegaly, cardiorespiratory failure, death, and a milder adult form being a gradual skeletal myopathy that sometimes causes respiratory problems
III	a	Hepatomegaly, hypoglycaemia, short stature while muscle involvement are characterized by progressive muscle wasting and weakness, cardiac dysfunction, congestive heart failure, and rarely sudden death
	b	Hepatomegaly and hypoglycaemia
	c	Hepatomegaly and hypoglycaemia
	d	Hepatomegaly and hypoglycaemia
IV		Hepatosplenomegaly, liver cirrhosis, progressive hepatic failure, and death
V		Muscle cramps and pain, fatigue during exercise (exercise intolerance), weakness
VI		Hepatomegaly, mild to moderate hypoglycaemia, mild ketosis, short stature
VII		Myopathy, hemolytic anemia, multisystem involvement (seizures, cardiopathy)
IX		Hepatomegaly, hypoglycaemia, short stature, myopathy
XI		Exercise intolerance, cramps, skin lesions

5.1.1 Mutation Spectrum of the *AGL* Gene

Molecular analysis on GSD III patients has revealed that their genetic causes are highly heterogeneous (Goldstein et al., 2010; Sentner et al., 2016). More importantly, a common theme observed in these reports was that the level of heterogeneity is different when compared between populations (Aoyama et al., 2009; Crushell et al., 2010; Okubo et al., 2000a). Populations such as the Japanese and Caucasian had been reported to

have higher genetic heterogeneity compared to the North African Jewish or the Faroe Islands populations. Different populations also appear to show variation in its genetic trends especially in the identification of common mutations segregating within the population. In a study conducted among eight Japanese GSD III patients, seven mutations were identified which include six that were novel ([1]c.771T>A, c.587delC, c.4216_4217delAG, c.2072_2073insA, c.4735_4736insTAT and IVS29-1G>C) and one previously reported mutation, IVS14+1G>T (Okubo et al., 2000a). Meanwhile, in another study conducted among Caucasian patients, four different mutations (c.2590C>T, c.3894delT, IVS32-12A>G and c.3682C>T) were found to be most frequent, but these mutations only account for 28 % of all mutant alleles reported (Shen et al., 1996).

In contrast, all 12 North African Jewish GSD III patients were found to be homogenous for mutation c.4455delT (Parvari et al., 1997). Similarly, six GSD III patients from the Faroe Island were homozygous for mutation c.1222C>T and only nine heterozygotes were identified out of 272 newborn screened (Santer et al., 2001). Founder effect was proposed and later confirmed as the factor for high prevalence of the respective mutations in these two populations.

Despite showing a variety of trends, there are also several mutations that can be found in various populations. These recurrent mutations such as the c.853C>T for example, was reported to be common in the Korean population had been previously reported in Japanese and Italian patients (Ko et al., 2013). The mutation IVS32-12A>G that was previously reported in North American patients was also found in a Chinese patient and a Japanese patient (Horinishi et al., 2002; Okubo et al., 1998). Another

[1] Corresponding to position c.371T>G, c.187delC, c.3816_3817delAG, c.1672_1673insA and c.4335_4336insTAT respectively according to newer numbering system

example is the mutation c.3980G>A that was reported in Turkish patients (Aoyama et al., 2009) were also identified in patients from Tunisia (Lucchiari et al., 2002), Egypt (Endo et al., 2005) and Canada (Endo et al., 2006). In one of these studies, haplotype analysis was also carried to assess the relationship of mutations between populations, as identical haplotypes suggest highly shared DNA i.e shared a common ancestor. However, Aoyama et al. (2010) found that different haplotypes were segregating, indicating that the c.3980G>A arose independently.

The present study reports the mutation spectrum within the *AGL* gene for Malaysian GSD III patients. From a total of 15 GSD III patients nine different pathogenic/potentially pathogenic mutations were identified in eleven individuals (see Table 4.2 and Figure 5.1 for details). Of these nine, six mutations are potentially novel as they have not been reported elsewhere, and these mutations will be discussed in greater detail in subsequent sections.

The overall results also showed that the most frequent mutation detected among this study cohort is the c.2681+1G>T (n=7/30, in total independent alleles analysed). It is perhaps noteworthy that this mutant allele was only detected in Malay patients and accounted for 43.8 % (n=7/16 alleles) of mutated alleles found in this group. The only other mutation that was recurrently detected in unrelated patients is the c.3814_3815delAG, which was identified in two unrelated Malay patients (refer to Table 4.1). All other mutations were detected in single individuals (except for c.4333T>G and c.4531_4534delTGTC, which were detected in 2 siblings respectively), but no conclusion in terms of its relative frequency can be established as yet, as the number of samples available for this study is limited.

5.1.1.1 Confirming the Pathogenic Status of Potentially Novel *AGL* Mutations

In mutation screening efforts, confirmation that the mutations play a causative role to the pathogenesis of GSD is first achieved by comparison and cross checking against published reports as well as information available in mutation databases. To further obtain support for pathogenicity, screening for the presence of the mutation in question in a panel of healthy control individuals are also carried out (Okubo et al., 2000). However, in cases where the mutations are yet unreported, several additional analyses can be carried out. Ideally, a protein functional assay that demonstrates the exact change (or loss) of normal function is warranted. However, these kinds of effort are quite laborious, as it often requires the establishment of stable and consistent protein expression systems, followed by the availability of techniques to efficiently extract and purify the desired proteins. Subsequently, efficient protein functional assays would have to be designed to observe and assess any shift from the normal protein function. Only then can the exact effect a mutation has on protein function can be established.

Since about a decade ago, an alternative approach to determine the effect of mutation on protein function using computational prediction has been made available. The prediction software relies on the assumption that functionally important amino acid have survived natural selection, and therefore conserved throughout evolution across multiple species. Changes (substitution, deletion or insertion) to these functionally important amino acids are potentially deleterious to gene function. SIFT (Sorting Intolerant From Tolerant), one of the earliest computational prediction software was based on a combined score derived from information on amino acid residues observed at a given position in the sequence alignment (Ng & Henikoff, 2001). More recently, information on structural features and employment of machine-learning techniques were integrated with the conservation-based method to better improve the predictive value. Examples of these include MAPP (Multivariate Analysis of Protein Polymorphism)

prediction, which is derived from information on the physicochemical constraints (hydropathy, polarity, charge, side-chain volume) of an amino acid (Stone & Sidow, 2005) and PolyPhen-2 (Polymorphism Phenotyping v2) prediction, which is derived from information on the sequence alignment and structural properties of a protein (Ramensky et al., 2002). It is important to note however that all three software (SIFT, MAPP and PolyPhen-2) are prediction tools to be effectively used for missense substitution. PROVEAN (Protein Variation Effect Analyser), a relatively new software, is reported to able not only predict the effect of single amino acid substitution but also multiple substitutions and in frame insertion or deletions (Choi et al., 2012). In present study, PolyPhen-2 was employed to predict the effect of mutations because of its generally high quality predictive value and due to its combined conservation-based, structural based as well as machine-learning technique predictive method (Tang & Thomas, 2016). It is also the tool used in various relevant publications (Cross et al., 2015; Leslie & Murray, 2013).

Throughout the course of this study, the six unreported mutations were two deletions causing translation frameshift and premature termination (c.2914_2915delAA and c.3814_3815delAG), one deletion causing translation frameshift and termination readthrough (c.4531_4534delTGTC), one missense (c.4333T>G), one nonsense (c.4490G>A) and one intronic splicing mutation (c.1423+1G>T). For each mutation, investigation of their presence/absence in a panel of control individuals was conducted. This was followed by the prediction of probable effects on splicing or protein function, depending on the type and position of the mutation. Finally, an attempt to obtain the mutant proteins was made by construction of expression cassettes and subsequent protein expression and analysis were carried out.

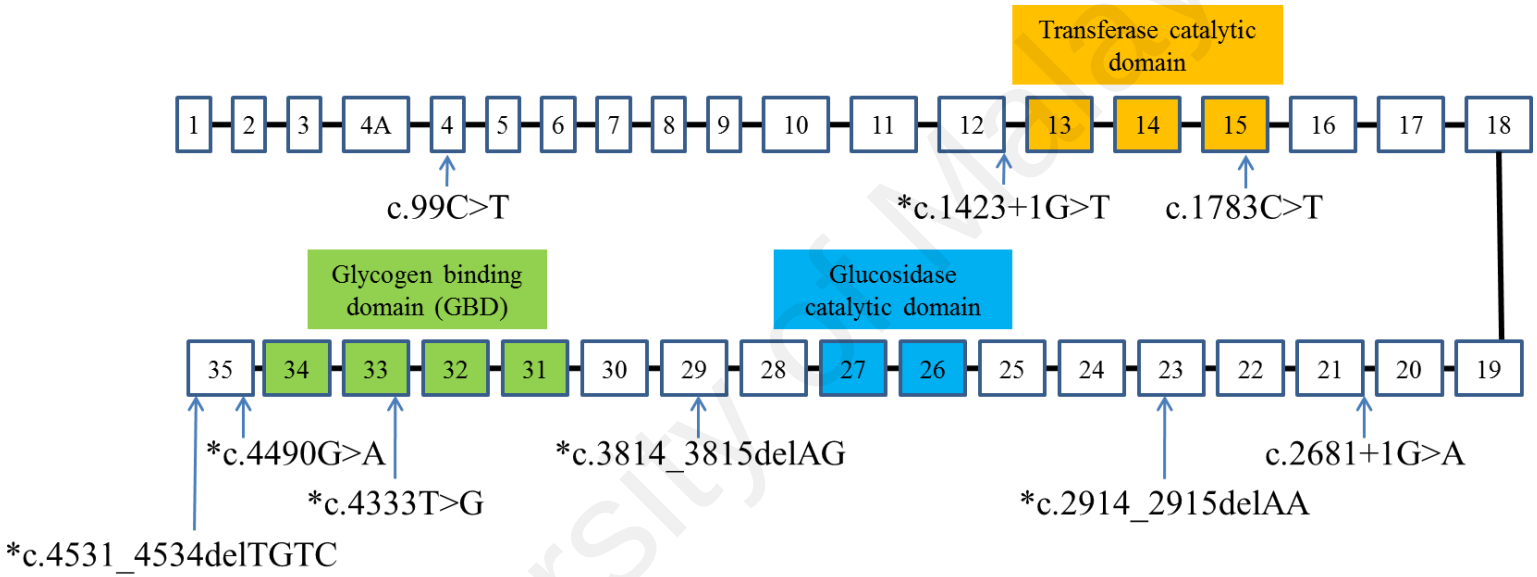


Figure 5.1: Schematic diagram of the *AGL* gene and related mutations found in present study. The numbered boxes represent exons, coloured boxes indicate the transferase catalytic domain (orange), glucosidase catalytic domain (blue) and glycogen binding domain (green). Black lines represent the introns and arrows indicate the position of *AGL* mutations. Unreported mutations are labelled with an asterisk.

i) Mutation c.1423+1G>T

This mutation that occurs at the exon-intron border of exon 12 was predicted to have abnormal RNA splicing. RNA splicing is a process of introns removal from pre-mRNA transcript followed by exon ligation to generate mature mRNA. This complex and tightly regulated process is important for generating functional proteins (Faustino et al., 2003; Hammond & Wood, 2011). The components that are required to direct the splicing process had been established which includes the 5' and 3' splice sites, intron branch site, splice site enhancers and silencers as well as the spliceosome machinery (Hammond & Wood, 2011; Mercer et al., 2015; Scotti & Swanson, 2016). The 5' and 3' splice sites are very critical for splicing process as changes to the invariant GU sequences of the 5' splice site or AG sequence of the 3' splice site will induce incorrect splicing. Changes resulting from the incorrect splicing however varies which can either be exon skipping, activation of a cryptic splice site or intron retention (Liu et al., 2001). A G-to-A substitution in intron 19 of *PFKM* gene for example resulted in exon skipping in a case of GSD Type VII. PCR amplification of patient cDNA showed a 165 bp in-frame truncation which was compatible with exon 19 deletion (Hamaguchi et al., 1994). Another splice site mutation of the *PFKM* gene, a G-to-T substitution at intron 15 activated a cryptic splice site and causes a 75 bp in-frame deletion of *PFKM* mRNA (Nakajima et al., 1990). Examples of mutations affecting the splice site within the *AGL* gene had also been reported. Mutation c.1736+1G>T that occurred at intron 14 was reported to cause exon 14 skipping and formation of truncated protein due to premature termination (Okubo et al., 1996). Another example, the mutation c.293+2T>A that occurred at the second nucleotide at 5' splice site of intron 4 was shown to cause aberrant splicing resulting in exon 4 skipping (Hadjigeorgiou et al., 1999). A similar mutation that occurred at the same location but with different nucleotide changes (c.293+2T>G) was predicted to have its 5' splice site function abolished through

computational splice site analysis (Aoyama et al., 2009). The above mentioned mutations occurred at immediate splice sites, however mutation c.4260-12A>G occurred upstream of 3' splice site of intron 32. The mutation creates a new 3' splice site and insertion of 11 bp intron sequences between exon 32 and exon 33 to the AGL mRNA (Okubo et al., 1998).

The c.1423+1G>T mutation that was found in this study was expected to have impaired splicing as the G-to-T substitution occurred at the conserve 5' splice site of intron 12. Computational splice site analysis using Human Splice Finder software predicted the same outcome as the splicing event is most likely to be affected through disruption of the splicing motif (refer to Table 4.10). The prediction was confirmed when two different reverse transcriptase PCR produced fragments longer than the normal transcript. A normal sample generated either a 816 bp or 464 bp fragments (depending on the primer pair used for amplification) while sample with affected splicing generated either a 1859 bp or 1507 bp fragments. The longer fragments were compatible with intron 12 retention as shown previously in Figure 4.27. Retention of intron 12 consequently generates a termination signal just two codons downstream of exon 12 boundaries. Truncated protein generated from this transcript will subsequently lose all three important GDE domains and ultimately disrupt its function. Disrupted enzyme function may have induced the clinical presentation observed in patient 7 which includes large hepatomegaly, overnight fasting hypoglycaemia and myopathy.

Patient 7 is a compound heterozygote for this c.1423+1G>T mutation and another reported splicing mutation, c.2681+1G>A (described in section 5.1.1.2). Screening of the parents' DNA sample detected the mother as the carrier for mutation c.2681+1G>A, but the mutation c.1423+1G>T was not detected in the father. This finding is unexpected as it is known that for an autosomal recessive mutation to occur,

each parents contributed to one affected allele. There are two possibilities that may explain the absence of mutation c.1423+1G>T in the paternal sample. The first possibility that can be explored is the occurrence of de novo mutation or mosaicism in Patient 7. De novo mutations can occur spontaneously during gamete formation or postzygotically (Lynch, 2010), and can arise both in somatic and germ cells throughout postnatal until adult life (Acuna-hidalgo, Veltman, & Hoischen, 2016). Mutations that occur postzygotically or later in the development/ postnatally can cause cell mosaicism (Acuna-Hidalgo et al., 2015; Rahbari et al., 2016; Youssoufian & Pyeritz, 2002). High-level mosaicism is observed when mutation takes place in the first few cell division cycles after fertilization whereas low-level mosaicism is seen when mutation occurs later in the development. De novo mutations and mosaicism had been described in several diseases particularly severe early-onset diseases (de Ligt et al., 2012; Fischer-Zirnsak et al., 2015; Kong et al., 2012) involving dominant genetic disorders. More recently, a combination of de novo mutation on one allele and inherited mutation on the other allele had also been reported in recessive disorders (Black et al., 2016), and this situation is similar to that seen in Patient 7 in this study. Black and colleagues (2016) reported of two unrelated cases where the first foetus was found to have compound heterozygous loss-of-function variants in *EVC2*, the gene that causes Ellis-van Creveld syndrome. They found that the first mutated allele was a 1-bp deletion in exon 16 inherited from the mother while the second was a de novo nonsense variant in exon 18 confirmed to had arisen on paternal haplotype. In the second case, they found the foetus to have compound heterozygous loss-of-function in *FRAS1*, the gene that causes Fraser syndrome. The first mutated allele was a 20-bp deletion in exon 41 determined to be inherited from the mother while the second mutated allele was a de novo 1-bp deletion in exon 66 confirmed to be arisen on the paternal chromosome (Black et al., 2016).

It is also interesting to note that a common factor shared by patient 7 and the two example cases shown above, is the fact that the de novo mutations were discovered to have paternal origin. This is in parallel with paternal-bias of de novo mutations based on reported studies (Veltman & Brunner, 2012). The other factor that influence de novo mutations is parental-age effects which had been intensely studied in early-onset genetic disorders such as intellectual disabilities and autism spectrum disorder (Goldmann et al., 2016; Kong et al., 2012). Kong et al. (2012) that looks at the association of paternal age in schizophrenic and autistic cases found that rate of de novo mutation increases of about two mutations per year. Francioli et al. (2015) further showed that the proportion of de novo mutations occurring in genic regions increases by 0.26 % per paternal year. That means offspring born to 40-year-old fathers have twice as many genic mutations compared to offspring of 20-year-old fathers. Unfortunately, however there was no information on the father's age for the patient involved hence no postulation can be made regarding paternal age and mutation c.1423+1G>T at the present point.

The second possibility is due to non-paternity event. In non-paternity, someone who is presumed to be an individual's father is in fact not the biological father (Tozzo et al., 2014). Occurrence of misattributed paternity cases have been quoted to be at a frequency of 10 % in the general population but due to lack of published evidence, the exact rate is difficult to be estimated (Bellis et al., 2005; Macintyre & Sooman, 1991). However, the possibility of encountering non-paternity in routine practice and research is acknowledged. Since the true relationship between Patient 7 and the father was not investigated, the possibility of encountering non-paternity in this case remains. To determine the true relationship in such cases, paternity testing can be carried out (Weber-Lehmann et al., 2014) using established methods.

ii) Mutations c.2914_2915delAA (p.N972Pfs*10) and c.3814_3815delAG (p.R1272Rfs*18)

The c.2914_2915delAA mutation is a deletion of two adenine nucleotides in exon 23 of a Malay patient resulting in substitution of an asparagine with a proline residue at codon 972 and also alters the reading frame to create new stop codon 10 amino acids downstream of the mutation site. The new molecular weight of the truncated protein is expected to be 112 kDa compared to 175 kDa of wild type GDE. On the other hand, the c.3814_3815delAG mutation is also a 2-bp deletion, but occurring in exon 29 and involving adenine and guanine nucleotides. This mutation also leads to a reading frameshift, subsequently generating a termination codon 18 amino acids downstream of the lesion. Both these mutations cause protein truncation, resulting to losses of approximately 550 and 240 amino acids at their C-terminal ends respectively. For the former, loss of the C-terminal half corresponds to loss of both the glucosidase active site and the glycogen binding domain (see Figure 5.1), while truncation in the latter involved losing the glycogen binding domain only.

Reports in the literature have established that mutations involving the GBD will severely impair GDE function with an apparent loss of both transferase and glucosidase enzymatic activities (Okubo et al., 2000; Lucchiari et al., 2002; Cheng et al., 2009; Crushell et al., 2010). The authors also reported that other than showing loss of enzymatic function, GDE protein that carry a tyrosine insertion ([2]c.4735_4736insTAT; p.Y1445ins) also resulted in the defective protein being increasingly targeted for degradation through proteosomal-mediated pathway. Similarly, Lucchiari and colleagues (2002) reported that two other GBD-related mutations,

[2] Corresponding to position c.4335_4336 according to the newer numbering system

[3]c.4593G>A; p.W1398X and [4]c.4724insAA; p.K1440Kfs*29 also abolished GBD activity and subsequently both transferase and glucosidase enzymatic functions, and patients carrying these mutations manifested with severe clinical presentations which include hepatosplenomegaly and cardiac hypertrophy (Lucchiari et al., 2002). Finally, a homozygous c.4197delA mutation within the GBD (introduces a termination signal 15 amino acids downstream of lesion, causing loss of approximately 118 amino acids of the C-terminal end) in an Irish patient showed complete absence of GDE activity in muscle biopsies, and manifests in massive hepatomegaly, severe hypoglycaemia, myopathy, cardiomyopathy and chronic elevation of CK level despite intensive dietary treatment (Crushell et al., 2010). Collectively, these research findings highlighted the importance of retaining normal GBD functions, and the extreme effects of impaired GBD on the overall activity of the GDE enzyme.

Based on these reported findings, it is highly likely that the c.2914_2915delAA and the c.3814_3815delAG mutations carried by the respective patients would result in severe, if not total loss of GDE function for this allele. The loss of the glycogen binding domain due to premature stop codon would be expected to prevent carbohydrate binding to GDE and the ensuing transferase and glucosidase activities are not likely to take place. It is also perhaps noteworthy that, in a recent communication with the attending paediatrician (Prof Sufin Yap, personal communication), enzymology testing on Patient 1 who was heterozygous for the c.2914_2915delAA mutation showed near null GDE activity with a reading of 0.1 pmol/min/mg.

[3] Corresponding to position c.4193G>A according to the newer numbering system
[4] Corresponding to position c.4321insAA according to the newer numbering system

iii) Mutation c.4333T>G (p.Y1445D)

A missense mutation, c.4333T>G that was found in exon 33 caused replacement of residue tyrosine to an aspartate at codon 1445 (p.Y1445D). Although this mutation also resides within the GBD like the ones described in the previous subsection, it does not result in total loss of the GBD domain. Only one residue, the tyrosine in amino acid position 1445 is involved, and as such, the exact outcome of the amino acid substitution is less easy to predict. Analysis using the PolyPhen-2 software predicts that this mutation as "probably damaging" perhaps because the substitution involved amino acids with different properties. Tyrosine which is a partially hydrophobic residue prefers to be buried in a hydrophobic core, while charged and polar aspartic acid prefers to be on the surface of a protein (Betts & Russell, 2003). Difference in these residue side chains is expected to influence their structure preference and in consequence, their function. Also, the fact that this tyrosine residue was also found to be highly conserved across various species (Cheng et al., 2009) is indicative of its importance in the protein structure and function.

Perhaps a comparison can be made with a missense mutation situated three amino acids downstream of this mutation. Cheng et al. (2007) who sought to understand the effects of p.G1448R on normal GDE function noted that the mutant protein failed to bind carbohydrate, which is an essential first step for the process of normal glycogen breakdown to progress. Furthermore, their studies suggested that disruption of the GBD also led to increased degradation, which is a similar fate as observed for loss of GBD domain mutations. In a paper published two years later by the same group (Cheng et al., 2009), the author suggested that some missense mutations (p.L620P and p.R1147G) in the GDE protein may lead to incorrect folding and subsequent loss of enzymatic activity, but whether a similar effect is applicable here is not directly clear. In a recent review on the effects of missense mutations, it was generally accepted that they can

have effects on protein stability, protein-protein interactions and characteristics of the active site (Zhang et al. 2012). Definitive description of the effects of mutations would only be available through exhaustive laboratory analyses and experiments, but one could also resort to computer aided predictive analysis which comparatively analyses of protein three dimensional (3D) structures, physicochemical properties of amino acids, energy calculation, or a combination of all of these.

iv) Mutation c.4490G>A (p.W1497X)

The functional consequences of this mutation and for that matter the c.4531_4534delTGTC mutation (discussed in the following section) are less clear as these mutations are situated outside the known important functional domains. The mutation c.4490G>A is located at the last *AGL* exon (exon 35) and was found in homozygous form in a Malay patient. The expected outcome is a truncated protein with a loss of 35 amino acids at the C-terminal region.

Other than this mutation, there are only two other reports of mutations within its vicinity. The first was a p.C1515R amino acid substitution mutation found in a mixed European patient who was speculated to be causative of GSD III through disruption of substrate recruitment (Goldstein et al., 2010; Zhai et al., 2016). However, no definitive experimental evidence was provided in their report. The other mutation is perhaps more relevant, which is an insertion mutation (c.4529insA) that resulted in a protein that is 23 amino acids shorter than the normal GDE protein. Clinically, this mutation is associated with a severe phenotype, causing the affected patient to have recurrent hypoglycaemia, hepatomegaly, severe cardiomyopathy, seizures and finally death at age 4 (Shen et al., 1997). When compared with patient 4 in this study, he/she too was presented with hypoglycaemia, hepatomegaly and myopathy, although no cardiomyopathy was detected (Prof Sufin Yap, personal communication).

The C-terminal region of proteins has been described to have diverse yet important functions. The C-terminal region of USP7 (Ubiquitin-Specific Protease) was shown to be essential for effective catalytic activity and promote sequence specific binding (Faesen et al., 2011) while the C-terminal domain of FGF21 on the other hand is required for ligand binding to β-Klotho (KLB) (Yie et al., 2009). Serial truncation of the C-terminal residues reduces the interaction between FGF21 and KLB (Yie et al., 2009), and removal of five C-terminal amino acids of FGF21 decreased ligand-binding affinity to KLB by up to 10-fold compared to wild type protein (Micanovic et al., 2009). C-terminal sequences are also closely associated with efficient protein folding as shown for Hsp60 protein in yeast, where Hsp60 mutants exhibit defective refolding (Fang & Cheng, 2002). Mutation at the C-terminal of human PCSK9 protein had also been reported to destabilize protein folding and secretion (Ai et al., 2016). Another study involving recombinant Pfg27, a soluble protein essential for sexual development of *Plasmodium falciparum* also found that the C-terminal is important for protein folding and solubility (Sati et al., 2002). All these findings collectively point to the fact that the C-terminal domains have diverse important functions in various types of proteins across different organisms, such that its absence would undoubtedly be expected to be detrimental to the normal function of the protein in question.

v) Mutation c.4531_4534delTGTC (p.C1511Lfs*35)

This mutation is located in exon 35, and it is predicted to produce a mutant GDE with abnormal C-terminal domain due to translation reading frameshift. However, unlike the previous mutation, the reading frameshift in this mutation is expected to change amino acid sequences from amino acid position 1511 (approximately 21 amino acids from the end of C-terminal), with an addition of another 14 new amino acids due to translation termination signal readthrough.

Currently there is no comparable mutation in the *AGL* being reported elsewhere. While stop codon readthrough had been described for viral replicase of bacteriophage and tobacco mosaic virus (Skuzeski et al., 1991), very few readtrough of the prokaryotes and eukaryotes genes had been reported (Beier & Grimm, 2001). However a study involving the metabolic enzyme MDH1, a dehydrogenase that is responsible for the reversible oxidation of malate to oxaloacetate had reported of readthrough isoform (Stiebler et al., 2014). They speculate that normal stop codon readthrough in human cells is used to generate peroxisomal isoforms of cytosolic enzymes, based on earlier study findings involving fungi glycolytic enzymes (Freitag et al., 2012). Ribosomal readthrough in fungi was found to lead to peroxisomal targeting due to the presence of hidden targeting signals in the readthrough extension (Freitag et al., 2012). In fact, another study on MDH1 found that the readthrough extension (termed as MDH1x) contains the peroxisome targeting signal and the MDH1x was in fact localized in the peroxisome (Hofhuis et al., 2016). While these studies report natural translational readthrough, whether the effect of mutational readthrough that occur in the present study is detrimental to normal protein function remains to be proven. However, as have been described in the preceding section, any change in the C-terminal region would be expected to result in aberrant protein function. Therefore, the contribution of mutation c.4531_4534delTGTC towards the pathogenicity of GSD III is thought to be most likely, since it was detected in a pair of siblings (Patients 2 and 3) who both have been diagnosed with this disorder.

5.1.1.2 Expression Study of Unreported *AGL* Gene Mutation

Attempts to further understand the effects of unreported mutations on normal GDE function were made by constructing plasmids carrying the mutated sequence. All five exonic mutations (c.2914_2915delAA, c.3814_3815delAG, c.4333T>G, c.4490G>A and c.4531_4534delTGTC) were successfully constructed through gene

cloning and subsequent mutagenesis steps. (The c.1423+1G>T mutation is intronic and affects the splicing mechanism, and therefore not included in this part of the study.) Upon transfection into HepG2 cells, the expressions of four out of five constructs were verifiable through specific RT-PCR analysis. However, when the expression of RT-PCR-positive clones were further analysed to confirm protein expression through Western blotting, none of them yielded positive results (Table 5.2).

Table 5.2: Summary of expression study on unreported *AGL* mutation.

Mutation	Verified site-directed mutagenesis	Transfection into HepG2 cells	Expression detection using RT-PCR	Western Blot
c.2914_2915delAA	/	/	/	x
c.3814_3815delAG	/	/	/	x
c.4333T>G	/	/	/	x
c.4490G>A	/	/	/	x
c.4531_4534delTGTC	/	/	x	x

The failure to detect protein expression through Western blotting is perhaps attributed to the loss of stability of the mutant protein. As had been explained previously, mutations that disrupts GBD domain affects protein stability and results in increased degradation rate (Cheng et al., 2007). In their study, Cheng and colleagues found that WT protein have a half-life of at least 9 hours while the GBD mutants (p.Y1445ins, p.G1448R and ΔGBD) were shown to have shorter half-life of 3 hours or less. Furthermore, they noted that the GBD mutants have enhanced ubiquitination (Cheng et al., 2007). In eukaryotic cells, most proteins destined for degradation are labelled first by ubiquitin before subsequent proteolytic action of the 26S proteasome (Goldberg, 2003; Pickart, 2001) in a process termed as the ubiquition-proteasome pathway. In a Lafora disease study involving the Laforin protein, E3 ubiquitin ligase Malin was shown to interact with mutant Laforin promoting its ubiquitination and

finally direct its degradation (Gentry 2005). Similar to GDE, Laforin is a dual specificity phosphatase protein consisting of a C-terminal catalytic domain and an N-terminal carbohydrate binding domain (Ganesh et al., 2000). Additional studies by Cheng's group later revealed that upon treatment with the proteasome inhibitor MG-132, the GBD mutant's level was significantly stabilized while there was no effect on the WT or non-GBD mutants (p.L620P or p.R1147G) (Cheng et al., 2009). They later proposed that the shorter half-life of GBD mutants were suggestive of ubiquitin-proteasome system involvement (Cheng et al., 2009). As most mutations found in this study (c.1423+1G>T, c.2914_2915delAA, c.3814_3815delAG and c.4333T>G) have their GBD domain affected, they would most probably have a shorter half-life due to decreased protein stability, and therefore results in failure to be detected by Western blotting. On the other hand, c.4490G>A and c.4531_4534delTGTC are non-GBD mutants and would not be expected to undergo the same fate. However, disruption to the C-terminal has also been reported to affect protein stability, hence the expression product of these two mutations are most likely to undergo degradation also.

It is also generally known that mutations can also give rise to misfolded proteins, which ultimately result in degradation of the affected proteins. As ER is the site for protein synthesis and packaging, the misfolded proteins may alter ER homeostasis and function giving rise to an ER stress state (Bravo et al., 2013). The cellular response to ER stress is known as the unfolded protein response (UPR) (Goldberg, 2003), which is an important process that protects ER from the accumulation of toxic misfolded proteins. The UPR functions to restore the normal ER condition and activity through multiple strategies. If in conditions where ER homeostasis cannot be restored, UPR switches its signaling towards a pro-apoptotic mode to eliminate damaged cells (Carreras-Sureda et al., 2017). In such conditions the misfolded protein aggregates and triggers the cell's proteolytic systems and molecular chaperones to act

(Ni & Lee, 2007). In the present study, cell cultures containing the mutant constructs were observed to have aggregate formation and most cells eventually undergo cell death. This observation has indeed been reported in many cloned, misfolded protein that aggregate in the cytosol of bacteria (Goldberg, 2003). Similar inclusions are also found in various inherited and neurodegenerative diseases, in which, if the inclusions failed to be solubilized, a chaperone-mediated solubilisation pathway will be initiated, after which they are later ubiquitylated and degraded by proteasome (Houck et al., 2012; Sherman & Goldberg, 2001). In the present analysis, only the WT control protein was detected by Western blotting and no results were obtained for all mutant constructs. Thus it would be logical to predict that these mutants have most likely been degraded. This notion can be confirmed by emulating Cheng et al. (2009) assay of proteasome inhibition which will allow detection of change in degradation rate through incorporation of compound MG-132.

5.1.1.3 Reported *AGL* Gene Mutations

In this study, mutation screening of the *AGL* gene also revealed three reported mutations which are the c.99C>T (p.R34X), c.1783C>T (p.R595X) and c.2681+1G>A. These reported mutations account for 50 % (11/ 22) of affected alleles. The detection of many different mutations in a limited number of patients corroborates with other reports that *AGL* gene mutation is highly heterogeneous.

i) Mutation c.2681+1G>A

This mutation that occurred at the donor splice site of intron 21 has been previously mentioned in patients of Italian (Hadjigeorgiou et al., 1999), Japanese (Uotani et al., 1995) and Germany (Endo et al., 2006) origin. This mutation that sometimes reported as IVS21+1G>A has been shown to be the most frequent mutation of GSD III found in each mentioned study. While this mutation is presumably to be a

recurrent mutation and could have occurred due to a founder effect, only one study had reported of haplotyping data (Endo et al., 2006) so no comparison can be made to confirm this assumption. In the present study, seven out of 22 independent alleles analysed showed to have this mutation, suggesting that this mutation also has high frequency in Malaysian patients. Moreover, this mutant allele was only detected in the Malays and not Chinese patients. Although the patient number in this study is small, 43.8 % (7 out of 16 alleles) detected in Malay patients possibly worth mentioning. Two patients (Patient 5 and Patient 9) were homozygotes for this mutation and three patients were heterozygotes (Patient 1, Patient 6 and Patient 7). All the patients carrying this mutation appear to have hypoglycaemia, severe hepatomegaly and either no detectable or very low enzyme activity (Prof Sufin Yap, personal communication). Homozygous form of this mutation was reported by Hadjigeorgiou et al. (1999) to be more severe than the heterozygous form. Lucchiari et al. (2002) also reported the same observation with their patients and further mentioned that there was variation in clinical symptoms of patients with the same homozygous status.

The c.2681+1G>A mutation was expected to affect splicing mechanism of transcription process as the mutation occur at the conserved GT sequence of exon-intron boundary. Hadjigeorgiou et al. (1999) proved that this mutation causes skipping of exon 21 in the mRNA sample of the patient that they had studied. The same finding was obtained when patient's RNA sample from this study was subjected to RT-PCR and sequence analysis of the amplicon confirmed the missing exon. Skipping of exon 21 is expected to affect GDE function as the glucosidase and glycogen binding domains will be disrupted.

ii) Mutation c.99C>T (p.R34X) and c.1783C>T (p.R595X)

The reported c.99C>T mutation that occur at exon 4 replaces an arginine residue with a stop codon at position 34. This mutation was found in homozygous form in a Malay patient that was presented with hepatomegaly and myopathy (Prof Sufin Yap, personal communication), a similar observation as seen in a homozygous Italian patient reported by Lucchiari et al. in 2002 . Another patient of Chinese origin had been reported to be a compound heterozygote for this mutation but no information on the symptoms presented (Horinishi et al., 2002).

Another reported mutation found in this study, the c.1783C>T occur at exon 15 and also causes the substitution of an arginine residue with a stop codon at position 595. This mutation was found in homozygous form in a Bajau patient that was presented with hypoglycaemia and moderate hepatomegaly (Prof Sufin Yap, personal communication), a similar observation as seen in a homozygous Turkish patient reported by Aoyama et al. in 2009.

Despite the similar clinical presentation observed between patients of this study and other patients reported in the literature for each mutation above, the genotype-phenotype correlation however was not determined for present study. While claims on the existence of phenotype-genotype correlation may exist (Shaiu et al., 2000), other studies however suggested the opposite (Aoyama et al., 2009; Mili et al., 2012). In study conducted by Aoyama, they found that even in patients with the same mutation, clinical features varied (Aoyama et al., 2009). A similar observation was reported by Mili and colleagues where a pair of siblings with same the genotype exhibited different clinical features (Mili et al., 2012).

5.1.1.4 SNPs Found in the *AGL* Gene

Other than the unreported mutations discussed in the previous section, thirteen other sequence variations were identified from GSD III samples in this study. Ten out of these variations (c.83-3C>T, c.894C>T, c.959-18G>A, c.2001+8T>C, c.2812+11G>A, c.2950-21T>A, c.3199C>T, c.3259+37G>A, c.3343G>A, and c.3836+45G>A) had been reported to be non-pathogenic SNPs. The other two variations (c.1185+15T>C and c.3836+53T>A) has not been reported elsewhere but are not thought to be involved in the pathogenesis of GSD in the present cohort as the patients that carry these SNPs have already been shown to carry other pathogenic mutations, or mutations that are more likely to be pathogenic based on functional prediction and population screening results.

SNPs are common throughout the genome and usually but not always occur at the intronic sequences (Castle, 2011). SNPs are expected to occur in a population at a frequency of 1 % or higher (≥ 1 %), because they are deemed harmless and have no effect on health and development. However, some SNPs (often categorized as non-synonymous SNPs) may introduce slight protein property changes which include substrate specificity, binding efficiencies and specificity of enzyme activity (Wang & Moult, 2001), and therefore a few have been associated with rare human diseases (Wang & Zhou, 2009). Other examples include SNPs within the *CYP* gene which had been associated with different rates of drug metabolism or susceptibility to certain diseases (Zhou et al., 2009). Polymorphism within this gene reduces the cytochrome P450 enzyme efficiency to metabolize drug (Nebert & Russell, 2002).

5.1.2 Mutations Screening of Selected Inherited Metabolic Disorders

As previously mentioned, mutation analyses of a few other inherited metabolic disorders were also included in the present study. These diseases include the glycogen storage disease type 1b, Fanconi-Bickel syndrome (used to be referred to as glycogen

storage disease type XI), methylmalonic aciduria (mut-type) and methylmalonic aciduria with homocystinuria (cblC-type).

Screening of the *SLC37A4* gene responsible for GSD 1b revealed two different mutations, in which both are unreported. Screening of the *SLC2A2* gene for FBS revealed three different mutations where only one have been previously reported as pathogenic (two mutations are yet unreported). Screening of the MMA of both types identified two mutations for the *MUT* gene and one mutation for the *MMACHC* gene, all three of which have been previously reported as pathogenic. These findings are discussed in greater detail in the following subsections.

5.1.2.1 Mutation Screening of the *SLC37A4* Gene for GSD 1b

Two unreported mutation were identified in GSD 1b patients in this study (see Table 4.4 and Figure 5.2). The c.196insT mutation, which is an insertion mutation that occurs at exon 5 of the *SLC37A4* gene was detected in Patient 2. The other mutation was the c.1286_1290delGAGTG deletion that was found in Patient 1, which occurred at exon 12. The predicted effect of each mutation on glucose-6-phosphate transporter enzyme is discussed below.

i) Mutation c.196insT (p.V66Cfs*6)

This mutation results in an uncommon replacement of a valine residue with a cysteine at position 66 and also result in a premature stop codon six amino acids downstream of the location of the mutation (p.V66Cfs*6). The predicted G6PT protein truncation is expected to destabilize the protein and eventually disrupt G6P uptake activity.

G6PT as previously stated is a ten helices transmembrane protein (Pan et al., 2009) with four cytoplasmic loops, four luminal loops and both the N and C termini

facing the cytoplasm (Pan et al., 1999). The mutation c.196insT is located in the 51-residue-long first luminal loop (Figure 5.2). Luminal Loop-1 is critical for G6P uptake activity as missense mutations within the loop had been reported to result in an enzyme with no detectable activity (Chou & Mansfield, 2014). As such, the protein product of c.196insT mutant will most likely have no enzyme activity.

The resulting premature stop codon will produce a truncated protein of only 71 amino acids long. The loss of the remainder 357 amino acids included several important regions such as the signature motif (QFTGTWWAILSTSMNLAG) located at amino acid 133-149 (Pan et al., 2003) and the carboxyl terminal ER retention signal (KKAE) located at amino acid 426-429 (Jackson et al., 1990). The signature motif of *SLC37A4* gene had also been described to be essential for the enzyme's normal function. In their report, Chen et al. (2002) had identified two mutations that alter the first (p.Q133P) and the last (p.G149E) residues of the signature motif. Both these mutations were later reported to show complete abolishment of microsomal G6P uptake activity (Pan et al., 2003). Another study carried out by Pan et al. (2003) found that mutations at the invariant residues 135, 146 and 149 also resulted in abolished G6P uptake activity. They further reported that mutations of the near invariant (p.L147A and p.A148D) also resulted in abolishment of uptake activity (Pan et al., 2003). Taken together, these findings underline the fact that both the invariant and near invariant residues of the signature motif are important for G6P uptake activity.

While absence of both the Luminal Loop-1 and the signature motif contributes to the predicted loss of G6P uptake activity in the c.196insT mutant protein product, absence of the transmembrane helices and the C-terminal domain may be more damaging. Mutations within the transmembrane helices had been reported to be damaging towards the stability of G6PT protein. Nonsense mutations (eg. p.W393X,

p.E401X and p.T408X) that resulted in proteins having no helix-10 structure caused the protein to not fold properly and undergo degradation (Chen et al., 2000; Hiraiwa et al., 1999). Additional argument supporting the possible protein destabilization due to the c.196insT mutation is the loss of C-terminal domain (amino acid 415-429). Located in this C-terminal domain is the ER retention signal (amino acid 426-429), and loss of the C-terminal domain, which in turn leads to the disruption of the ER retention signal would be disruptive to the overall protein function. However, it has also been reported that a naturally occurring mutation, the p.R415X that lacked the entire C-terminal domain (and thus the ER retention signal) however retain about 47% of wild-type G6PT activity (Chen et al., 2000), even though protein stability analysis interestingly showed that this mutant form was found to degrade more rapidly than the wild-type G6PT (Chen et al., 2000). Further study on G6PT C-terminal mutants (p.N416X, p.I417X, R418X, p.T419X and p.K420X) determined that the first three residues of cytoplasmic tail (amino acid 415-417) are associated with stability (Chen et al., 2000).

ii) Mutation c.1286_1290delGAGTG (p. E429Kfs*58)

The effect of mutation c.1286_1290delGAGTG (p.E429Kfs*58) is less easy to predict. The deletion mutation that occurs at the last amino acid residue of G6PT protein (Figure 5.2) disrupts the immediate stop codon and is expected to produce a protein with longer C-terminal, as it causes the incorporation of 58 additional amino acids. There is no comparable mutation in the *SLC37A4* gene being reported elsewhere. However, explanation on the effect of this mutation is attempted based on existing knowledge of the G6PT protein structural requirement.

Theoretically, the mutation is predicted to have less severe impact on G6P uptake into the lumen. As previously mentioned both the Luminal Loop-1 and the signature motif of *SLC37A4* gene are crucial for optimum G6P uptake activity (Chou &

Mansfield, 2014; Pan et al., 2003). The fact that mutation c.1286_1290GAGTG occurs at codon 429, the Luminal Loop-1 and signature motif are expected to be undisrupted which would then correspond to normal or active G6P uptake. However, because the C-terminal now is longer than usual, its potentially harmful effects towards normal protein folding cannot be ruled out. This in turn would be expected to affect the stability of the mutant protein, as any change from the normal sequence or structure of the C-terminal end has been shown to lead to increased degradation (see explanation for *AGL*, for example).

At present, the effects of additional 58 amino acids to the cytoplasmic tail of mutant protein cannot be determined, and at best, the arguments presented herein are speculative in nature. Additional studies directed towards the understanding of the effect of mutation c.1286_1290delGAGTG on normal protein function are therefore warranted.

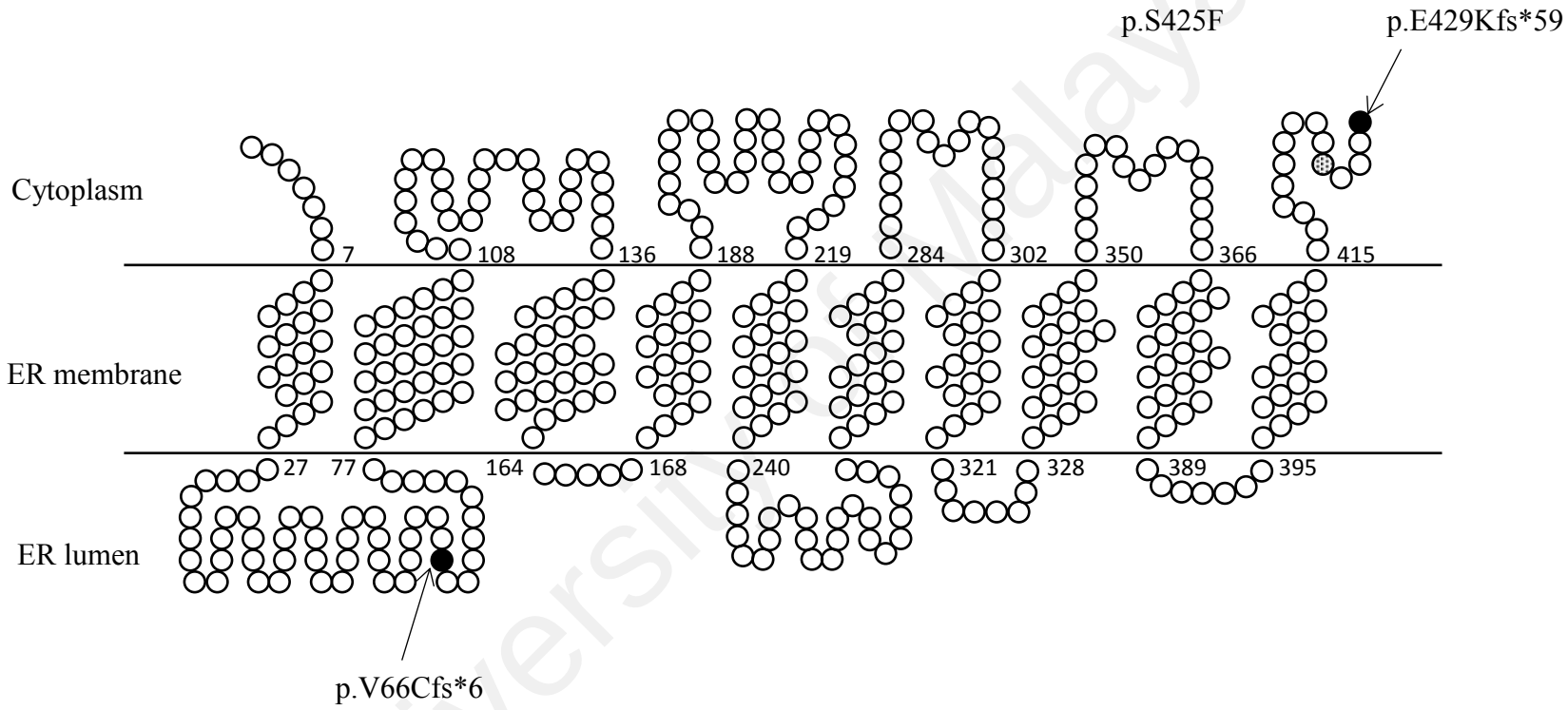


Figure 5.2: Schematic diagram of the G6PT protein and related mutations found in present study. Circles represent amino acids, black-coloured circles indicate unreported mutations and patterned circle indicates reported SNP (Adapted from Pan et al., 2009).

5.1.2.2 Mutation Screening of the *SLC2A2* Gene for FBS

Screening of the *SLC2A2* gene for FBS revealed three different mutations for which only one (c.952G>A) have been previously reported as pathogenic and two mutations (c.589G>C and c.872_873delTA) are still unreported.

i) Mutation c.952G>A (p.G318R)

The reported c.952G>A mutation occurs at exon 6 of the *SLC2A2* gene replaces a conserved glycine at position 318 with an arginine (p.G318R) (Figure 5.3). This mutation was found in heterozygous form in a Chinese patient. Santer et al. (2002) had previously reported the presence of this mutation in an Italian and a Caucasian patient. Occurrence of this mutation in various populations may be useful for routine *SLC2A2* gene mutation screening, but data from more populations need to be obtained to render its suitability.

ii) Mutation c.589G>C (p.V197L)

This unreported missense mutation c.589G>C substitutes a valine residue with a leucine at position 197 (p.V197L) (Figure 5.3). This mutation occurs at exon 5 of the *SLC2A2* gene, and was found in homozygous state in a patient of Indian ethnicity. Both valine and leucine are hydrophobic and aliphatic residues. The difference between the two however lies in their size. Valine is a small residue meanwhile leucine is relatively large. Despite this size difference, substitutions between these two amino acids are frequently encountered and have been suggested to have minor effect on normal protein function (Betts & Russell, 2003). Upon analysis using the PolyPhen-2 prediction software however, the predicted effect on protein function was categorized as "probably damaging". The valine residue was noted to be highly conserved through the alignment of multiple *SLC2A2* sequences from various species. It is therefore speculated that

different residues may have different affinities for substrate recognition, and affect the overall efficiency of protein function.

Other studies had also reported the deleterious effect of similar (valine to leucine) substitution. Substitution of valine to leucine at codon 865 of the human androgen receptor (*AR*) gene had been associated with partial form of androgen resistance (Kazemi-Esfarjani et al., 1993). The researchers concluded that such finding was attributed to selective androgen-binding affinity of the mutant. Another study by Rost et al. (2004) found that mutation p.V29L in *VKORC1* gene reduced the activity of vitamin K epoxide reductase multiprotein complex (VKOR) by 96 % compared to the wild-type. VKOR is an essential cofactor for several blood coagulation factors. Reduced VKOR activity had been associated with increased vitamin K demand and subsequent death from spontaneous bleeding (Rost et al., 2004).

iii) Mutation c.872_873delTA (p.I291Nfs*100)

The other unreported mutation is c.872_873delTA, which is located in exon 6 of the *SLC2A2* gene. Deletion of two nucleotides causes an isoleucine substitution with an asparagine at position 291. This mutation also results in reading frameshift and consequently formation of a premature stop codon 100 amino acids downstream of the mutation, and a loss of 132 amino acids of the C-terminal end.

Isoleucine at position 291 is located within the large central intracellular loop of 65 amino acids (Figure 5.3) (Mueckler et al., 1985) that contain the sugar transporter family signature motif PESPR (amino acids 240-244) and sugar signaling domain (amino acid 295) (Michau et al., 2013). A missense mutation that occur within the signature motif (p.S242R) expression was barely detected and not targeted at the plasma membrane (Michau et al., 2013). Conversely however, the missense mutation p.F295Y that occur at the highly conserved amino acid within the sugar signaling domain of

central intracellular loop reported to retain its expression and transport activity (Michau et al., 2013). This is perhaps due to the fact that phenylalanine replacement with tyrosine can be classified as a conservative mutation, and therefore not expected to be disruptive (Betts & Russell, 2003). On the other hand, a substitution with a non-conservative residue is expected to have a more unfavourable outcome.

More importantly, the mutation c.872_873delTA will also result in a truncated protein due to loss of 132 amino acids. Despite having an intact signature motif which is important for sugar transport activity, the truncating mutation is expected to cause loss of all GLUT2 function. Not only the sugar signaling domain will be disrupted, sugar tunnels and sugar binding-site located at transmembrane 9, 10 and 11 (Michau et al., 2013) respectively, are expected to be missing. The loss of these domains at *SLC2A2* C-terminal consequently prevents the conformational change requirement of GLUT2 to function (Muraoka et al., 1995). An earlier study by Oka et al. (1990) had described that the C-terminal of GLUT1 is required for formation of outward-facing confirmation and therefore mutation within this region locked the formation in an inward-facing form (Oka et al., 1990). Since facilitated glucose transported share similar structure, GLUT2 is expected to adopt this 'model of alternating conformation'.

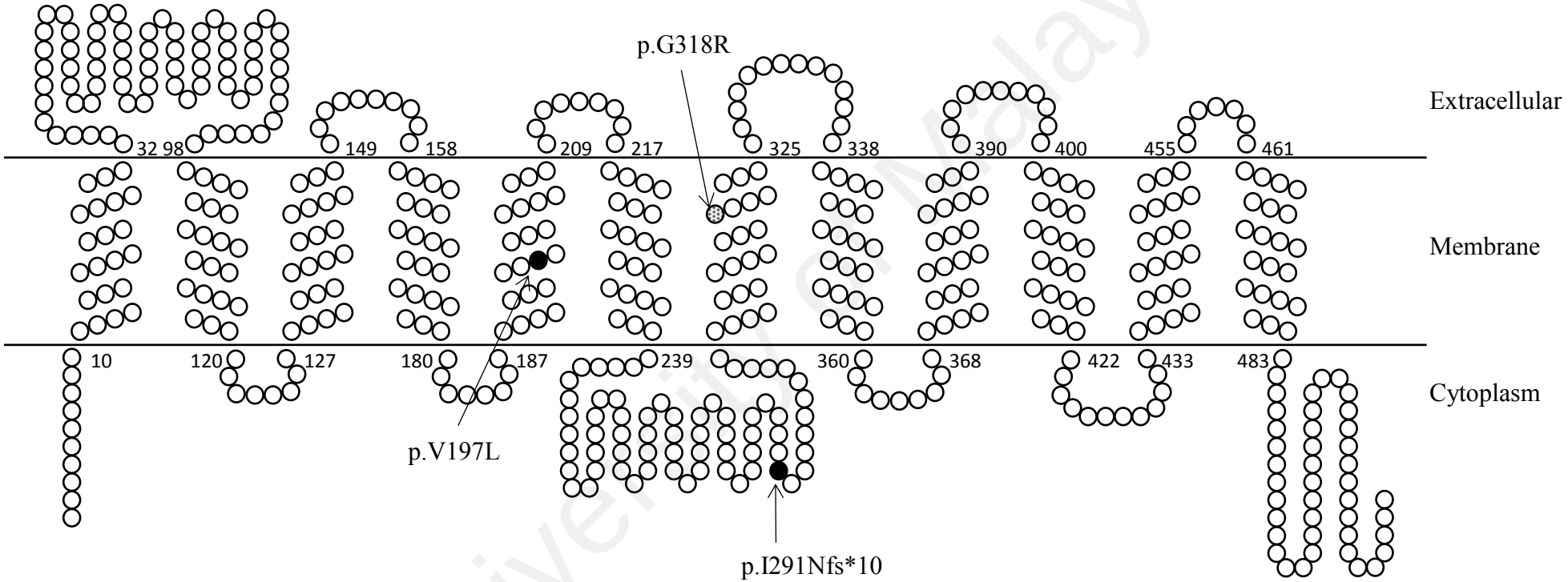


Figure 5.3: Schematic diagram of GLUT2 topology and mutations found in this study. Circles represent amino acids, black-coloured circles indicate unreported mutations and patterned-circle indicates reported mutation (Adapted from Michau et al., 2013).

5.1.2.3 Mutation Screening of the *MUT* Gene for MMA Mut-Type

Screening of the *MUT* gene in a single MMA mut-type patient revealed two different mutations (c.1280G>A and c.1630_1631delGGinsTA). Both mutations have been previously reported elsewhere as pathogenic.

i) Mutation c.1280G>A (p.G427D)

The reported c.1280G>A mutation that occur at exon 6 of the *MUT* gene replaces a glycine at position 427 with an aspartic acid. This mutation was found in heterozygous form in the single MMA mut-type patient of this study. There are a number of *MUT* genes mutations that had been associated with specific ethnic populations (Forny et al., 2014) such as the c.1843C>A in Turkish Asian (Dündar et al., 2012) or c.2150G>T in African Americans (Worgan et al., 2006). It is noteworthy that the Chinese patient in this study have similar mutations as observed in Chinese patients from another report by Liu et al. (2012). Moreover Liu's group found that c.1280G>A is the most frequent *MUT* variation found among Southern Chinese. The first group that reported this c.1280G>A mutation also found it to have occurred in three Asian patient (Worgan et al., 2006). Further study may be required to determine if this mutation can be used for *MUT* gene screening among the Chinese.

ii) Mutation c.1630_1631delGGinsTA (p.G544X)

The reported c.1630_1631delGGinsTA mutation that occurs at exon 9 of the *MUT* gene replaces a glycine residue with a stop codon at position 544 (p.G544X). Mutation c.1630_1631delGGinsTA affects the linker region of *MUT* gene and involved in a highly conserved amino acid (Chu et al., 2016; Liu et al., 2012; Worgan et al., 2006).

Expression study of the mutant protein in *E.coli* showed no detectable enzyme activity (Champattanachai et al., 2003). This finding was expected due to loss of the

MCM domain and cobalamin binding domain (Drennan et al., 1996; Mancia et al., 1996). Furthermore, it was suggested that the mRNA carrying the mutation was less stable than the wild-type mRNA due to the inability to detect the mutation in parent's mRNA (Champattanachai et al., 2003). Mutation c.1630_1631delGGinsTA is now described as one of three common mutations among the Chinese (Han et al., 2015; Liu et al., 2012; Worgan et al., 2006). While the patient in this study is also of Chinese ethnicity, this mutation had also been detected in a patient of Thai origin (Champattanachai et al., 2003) as well as in a Palestinian patient (Chu et al., 2016).

5.1.2.4 Mutation Screening of the *MMACHC* Gene for MMA CblC-Type

Screening of the *MMACHC* gene in two siblings of Chinese origin revealed a single mutation in exon 2. The c.609G>A mutation causes the replacement of a tryptophan residue with a stop codon at position 203 (p.W203X). Previously, this mutation had been reported in five East Asian individuals (Lerner-Ellis et al., 2006). Possible association between the Malaysian Chinese and the East Asian population may exist as Mainland China was the ancestral origin of Malaysian Chinese. This is especially beneficial for diagnosis as possible phenotype-genotype correlations and ethnicity-related trends within the *MMACHC* gene had been proposed (Carrillo-Carrasco et al., 2012; Chang et al., 2011).

5.2 Cases with No Mutation Detected

Mutation screening failed to detect any possible pathogenic mutations in eight patients, namely Patient 12, 13, 14 and 15 for GSD III, Patient 4 and 5 for GSD 1b as well as Patient 3 and 4 for FBS. In most cases, this highlights the challenges and difficulty in interpreting the clinical phenotypes and eventually coming up with an accurate diagnosis. While there are some distinct clinical features differentiating various types of GSDs, overlapping features are also present (Hendriksz & Gissen, 2014). This

is apparent in two GSD III patients (Patient 12 and 13) when they were initially diagnosed with GSD 1a, while ensuing enzymatic assay ruled out GSD 1 as the diagnosis. Another case in point was when Patient 14, despite having positive histological findings suggestive of GSD III diagnosis, enzymatic assay eventually confirmed that Patient 14 did not have the disease.

As there were no enzymatic assay results to help explain the negative mutation detection in the other patients, other possible justifications were explored. Mutation screening in this study was carried out across the exons and immediate intronic regions of about 100 bp. Mutations could remain undetected if they lie outside the boundaries of the current search region. It has indeed been shown that some sequences deep within the intronic region or within the 5'-UTR sometimes bear several regulatory elements such as the TATA-like motif, the CAAT box and the CCAAT/ enhancer binding site (C/EBP) (Takeda et al., 1993) and many others (Cartegni et al., 2002). Several diseases have been shown to be caused by mutation that lies within the regulatory sequences. For example point mutations in the polarizing activity regulatory sequence (ZRS), a *Cis*-regulatory enhancer is known to cause human limb malformations (VanderMeer et al., 2014). In another example, Shatunov and colleagues (2004) described a patient with rapid progressive cognitive decline and pronounced ataxia which was found to have 55 CAG/CAA repeats in the TATA-binding protein compared to normal 37-38 repeats.

Failure to detect mutations could also be due to the fact that these patients actually carry heterozygous long-range deletions that are not detectable using standard PCR analysis (Santer et al., 2002). Linkage analysis of an MMA patient carrying a homozygous c.1280G>A mutation found that the mutation is inherited only from the father, while the mutation was absent in the mother. It was eventually discovered that

both patient and mother harboured a deletion across their chromosome 6 genomic sequence (Liu et al., 2012).

CHAPTER 6: CONCLUSION

This chapter concludes the research work with relevance to research objectives, as well as highlighting some of the research outputs in terms of its contributions and advocating future work. The complete organization of this chapter is as follows.

Section 6.1 describes the fulfilment of research objectives and section 6.2 highlights the contribution of this research. Section 6.3 describes the limitations of the research work and suggestions for future work. Lastly, section 6.4 provides the summary of this study.

6.1 Revisiting the Research Objective

The general purpose of this study is to identify the causative mutations that can be implicated in several IMDs cases in Malaysia and to characterize the mutation spectrum. In the subsequent sections, the five specific objectives of this study were revisited and discussed.

The first objective is to determine the molecular genetic cause of GSD III in Malaysian population. Nine different mutations of the *AGL* gene were successfully identified, six of which are yet unreported while the other three had been previously reported. The findings showed that the mutations were scattered throughout the entire coding region of the *AGL* gene and most of the mutations found are deletion mutations that produce truncated protein. Within the small study group included in this study, the splicing mutation c.2681+1G>A was found to be the most common with 7/22 alleles carrying this mutation.

The second objective of this study is to determine the molecular genetic causes of less common GSD types which are the GSD 1b and FBS. Two unreported mutations were identified upon screening of the *SLC37A4* gene (GSD 1b). Meanwhile, screening

of the *SLC2A2* gene (FBS) found one previously reported mutation and two unreported mutations.

The third objective of this study is to determine the molecular genetic causes of selected IMDs which are the MMA mut-type and MMA with homocystinuria cblC-type. Mutation screening found two reported mutations in the *MUT* gene (MMA mut-type) and one reported mutation in the *MMACHC* gene (MMA with homocystinuria cblC-type).

The fourth objective is to verify the pathogenicity of unreported mutations. In general, all unreported mutations found in this study were predicted to be potentially pathogenic through predictive analysis. The mutations are expected to be causative of their respective disease due to either frameshift that leads to protein truncation, disruption of normal splicing or disruption to protein structure and function. Further population study on unreported *AGL* mutations found that all mutant alleles were absent in the control population. This exclude the possibility of unreported mutations found to be mere polymorphism.

The last objective is to determine the effects of unreported *AGL* mutations on protein function. Splicing analysis on mutation c.1423+1G>T that occur at the exon-intron border found that it causes intron 12 retention leading to a premature termination. Other *AGL* mutations are speculated to be misfolded and degraded due no detection in Western blotting during functional study. This speculation was based on loss or disruption of the GBD and/or C-terminal end. Published research also suggested that sequence changes in mutants protein usually promotes its own degradation, perhaps through pathways that aim to protect and preserve normal cellular functions.

6.2 Contributions

To the author knowledge, prior to this work there is no known study focusing on characterizing and establishing the mutation spectrum for Malaysian GSD and MMA patients. As such, information obtained in this study will be useful in the management of these disorders, including improved diagnosis and treatment, as well as genetic counselling of patients and their family members.

i. Diagnosis

As mentioned previously, diagnosis of inherited metabolic disorders are laborious. Knowledge on the prevalent mutations among Malaysian population will allow for narrowing down of mutations to be tested. This will translate into reduction in cost and time taken for a diagnosis to be made. Where TaqMan® analysis is concerned, the primers and probes that have been designed in this study can be used for fast and reliable screening of selected mutations. The restriction enzyme assay used in this study also can be used as an alternative diagnostic method to detect known mutations.

ii. Treatment

The course of treatment for particular disease may vary to cater for specific symptoms. For diseases like GSD, where symptoms of distinct types may be similar, information obtained from this study may contribute towards the proper planning of treatment needed by affected individuals. Treatment with targeted drugs can be designed for patients carrying certain mutation and the outcome assessed. An established approved therapy for the mutation subset can then be developed. Furthermore, this approach will be in line with the current direction towards personalized therapy that aims to match certain genotypes with effective targeted therapies.

iii. Genetic counselling

Genetic counselling in the country is finally taking shape, with more clinics being established nationwide. Genetic information of diseases, such as those amounted in this study, is invaluable in providing information and support to patients and family members affected by or at risk of inheriting a genetically-influenced conditions. These information may also be used to help families understand the significance of genetic disorders in personal as well as familial situations and to promote informed-choice making.

6.3 Limitations and Future Improvement

The final part of this research project, which was to characterize the effects of the unreported mutation on normal GDE function, was unfortunately only partially achieved. Assays for splicing mutations (including positive control assays involving c.2681+1G>A) were successfully carried out with very unambiguous results, allowing definitive conclusion to be met. However, for the remaining unreported mutations, support for pathogenicity and their causative involvement in GSD III patients were only obtained through population screening, which showed that the mutations were not present in a panel of healthy individuals. Additional support provided through software prediction studies also indicated that all unreported mutations were most likely to be pathogenic.

Perhaps support for pathogenicity is not so crucial for nonsense and truncating mutations (due to their deleterious loss of the GBD and other active sites), but support for missense and readthrough mutations are needed. As had been discussed, such mutations (missense and readthough) may cause misfolding which would then lead to aggregation and ultimately degradation. To prove this conviction would require experimental work that can monitor protein misfolding/aggregation, which could be

achieved by utilizing fluorescent protein reporter such as green fluorescent protein (GFP) (Gregoire et al., 2013). GFP was first used to examine the aggregation and folding status of target protein in *E. coli* (Waldo et al., 1999) and its used had since been extended to mammalian cells. Gregoire and colleagues used enhanced GFP (EGFP) to analyse folding of human copper/zinc superoxide dismutase (SOD1) expressed in HEK293T and NSC-34 cells. They used flow cytometry to provide data that correlates the average cellular fluorescence intensity to the extent of misfolding/aggregation of a target protein in mammalian cells (Gregoire et al., 2013).

6.4 Summary of Study

The present study has enabled the identification of mutations responsible for GSD III, GSD 1b, FBS and MMA (mut-type and cblC-type) patients in Malaysia (Table 6.1). Several potentially novel mutations have been identified, and due to time limitation supporting evidence were only obtained for some of the unreported GSD III mutations. The present study also established that mutations for GSD III are heterogeneous in nature, with c.2681+1G>A being the most commonly found in the current cohort.

Table 6.1: Summary of mutations identified in Malaysian GSD 1b, GSD III, FBS, MMA and MMA cblC-type patients.

Disease	Gene	Unreported mutation	Reported mutation
GSD 1b	*SLC37A4*	c.1280_1290delGAGTG c.196insT	
GSD III	*AGL*	c.1423+1G>T c.2914_2915delAA c.3814_3815delAG c.4333T>G c.4531_4534delTGTC c.4490G>A	c.99C>T c.1783C>T c.2681+G>A
FBS	*SLC2A2*	c.589G>C c.872_873delTA	c.952G>A
MMA	*MUT*		c.1280G>A c.1630_1631delGGinsTA
MMA cblC-type	*MMACHC*		c.609G>A

REFERENCES

Acquaviva, C., Benoist, J. F., Callebaut, I., Guffon, N., Ogier de Baulny, H., Touati, G., … Elion, J. (2001). N219Y, a new frequent mutation among mut0 forms of methylmalonic acidemia in Caucasian patients. *European Journal of Human Genetics: EJHG*, *9*(8), 577–582.

Acuna-Hidalgo, R., Bo, T., Kwint, M. P., Van De Vorst, M., Pinelli, M., Veltman, J. A., … Gilissen, C. (2015). Post-zygotic point mutations are an underrecognized source of de novo genomic variation. *The American Journal of Human Genetics*, *97*, 67–74.

Acuna-hidalgo, R., Veltman, J. A., & Hoischen, A. (2016). New insights into the generation and role of de novo mutations in health and disease. *Genome Biology*, *17*, 1–19.

Adeva-Andany, M. M., González-Lucán, M., Donapetry-García, C., Fernández-Fernández, C., & Ameneiros-Rodríguez, E. (2016). Glycogen metabolism in humans. *BBA Clinical*, *5*, 85–100.

Adjalla, C. E., Hosack, A. R., Matiaszuk, N. V, & Rosenblatt, D. S. (1998). A Common Mutation among Blacks with mut- methylmalonic aciduria. *Human Mutation*, *250*, 248–250.

Ai, X., Palyha, O. C., Ha, S., Quan, S., Chu, D., Zhang, A., … Strack, A. M. (2016). C-terminal loop mutations determine folding and secretion properties of PCSK9. *Biochemistry & Molecular Biology Journal*, *2*(3), 1–12.

Al-Haggar, M., Sakamoto, O., Shaltout, A., El-Hawary, A., Wahba, Y., & Abdel-Hadi, D. (2011). Fanconi-Bickel syndrome: Novel mutations in *GLUT2* gene causing a distinguished form of renal tubular acidosis in two unrelated Egyptian families. *Case Reports in Nephrology*, *2011*, 754369.

Annabi, B., Hiraiwa, H., Mansfield, B. C., Lei, K. J., Ubagai, T., Polymeropoulos, M. H., … Chou, J. Y. (1998). The gene for glycogen-storage disease type 1b maps to chromosome 11q23. *American Journal of Human Genetics*, *62*(2), 400–405.

Aoyama, Y., Ozer, I., Demirkol, M., Ebara, T., Murase, T., Podskarbi, T., … Okubo, M. (2009). Molecular features of 23 patients with glycogen storage disease type III in Turkey: A novel mutation p.R1147G associated with isolated glucosidase deficiency, along with 9 *AGL* mutations. *Journal of Human Genetics*, *54*(11), 681–686.

Arion, J. W., Wallin, B. K., Carlson, P. W., & Lange, A. J. (1972). The specificity of glucose-6-phosphatase of intact liver microsomes. *The Journal of Biological Chemistry, 247*, 2558-2565.

Arion, J. W., Wallin, B. K., Lange, A. J., & Ballas, L. (1975). On the involvement of a glucose-6-phosphate transport system in the function of microsomal glucose-6-phosphatase. *Molecular and Cellular Biochemistry, 6*, 75-83.

Bali, D. S., Chen, Y.-T., Austin, S., & Goldstein, J. L. (2016). *Glycogen storage disease type I. GeneReviews*. Seattle, WA: University of Washington.

Bao, Y., Dawson, T. L., & Chen, Y. T. (1996). Human glycogen debranching enzyme gene (*AGL*): Complete structural organization and characterization of the 5' flanking region. *Genomics*, *38*(2), 155–165.

Basit, S., Malibari, O., Mahmood, A., Balwi, A., Abdusamad, F., & Ismail, F. A. (2014). A founder splice site. *Annals of Saudi Medicine*, *34*(5), 390–395.

Baumgartner, M. R., Hörster, F., Dionisi-Vici, C., Haliloglu, G., Karall, D., Chapman, K. A., … Chakrapani, A. (2014). Proposed guidelines for the diagnosis and management of methylmalonic and propionic acidemia. *Orphanet Journal of Rare Diseases*, *9*(1), 130.

Beier, H., & Grimm, M. (2001). Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. *Nucleic Acids Research*, *29*(23), 4767–4782.

Bellis, M. A., Hughes, K., & Ashton, J. R. (2005). Measuring paternal discrepancy and its public health consequences. *Journal of Epidemiology & Community Health*, *59*(9), 749–754.

Ben-Omran, T., Wong, H., Blaser, S., & Feigenbaum, A. (2007). Late-onset cobalamin-c disorder: A challenging diagnosis. *American Journal of Medical Genetics. Part A*, *143A*(18), 2106–2112.

Berger, I., Shaag, a, Anikster, Y., Baumgartner, E. R., Bar-Meir, M., Joseph, a, & Elpeleg, O. N. (2001). Mutation analysis of the *MCM* gene in Israeli patients with mut(0) disease. *Molecular Genetics and Metabolism*, *73*(1), 107–110.

Berry, G. T., Baker, L., Kaplan, F. S., & Witzleben, C. L. (1995). Diabetes-like renal glomerular disease in Fanconi-Bickel syndrome. *Pediatric Nephrology*, *9*(3), 287–291.

Betts, M. J., & Russell, R. B. (2003). Amino acid properties and consequences of substitutions. In M. R.. Barnes & I. C. Gray (Eds), *Bioinformatics for geneticists* (pp. 289–316). Chichester, UK: John Wiley & Sons, Ltd.

Black, H. A., Parry, D., Atanur, S. S., Ross, D., Rose, E., Russlell, H., … Evans, M. J. (2016). De novo mutations in autosomal recessive congenital malformations. *Genetics in Medicine*, *18*(12), 1325–1326.

Blakemore, S. J., Aledo, J. C., James, J., Campbell, F. C., Lucocq, J. M., & Hundal, H. S. (1995). The GLUT5 hexose transporter is also localized to the basolateral membrane of the human jejunum. *The Biochemical Journal, 309*(1), 7-12.

Bollen, M., Keppens, S., & Stalmans, W. (1998). Specific features of glycogen metabolism in the liver. *The Biochemical Journal*, *336 (1)*, 19–31.

Bravo, R., Parra, V., Gatica, D., Rodriguez, A. E., Torrealba, N., Paredes, F., … Lavandero, S. (2013). Endoplasmic reticulum and the unfolded protein response. Dynamics and metabolic integration. *International Review of Cell and Molecular Biology*, *301*, 215–290.

Brookes, A. J. (1999). The essence of SNPs. *Gene*, *234*, 177–186.

Brown, G. K. (2000). Glucose transporters: Structure, function and consequences of deficiency. *Journal of Inherited Metabolic Disease*, *23,* 237–246.

Brown, T. A. (2016). *Gene cloning and DNA analysis: An introduction* (6th ed.). Chichester: Wiley-Blackwell Publishing, Inc.

Burcelin, R., & Thorens, B. (2001). Evidence that extrapancreatic GLUT2-dependent glucose sensors control glucagon secretion. *Diabetes*, *50*(6), 1282–1289.

Burton, B. K. (1998). Inborn errors of metabolism in infancy: A guide to diagnosis. *Pediatrics*, *102*(6), 1-9.

Burwinkel, B., Sanjad, S. A., Al-Sabban, E., Al-Abbad, A., & Kilimann, M. W. (1999). A mutation in *GLUT2*, not in phosphorylase kinase subunits, in hepato-renal glycogenosis with Fanconi syndrome and low phosphorylase kinase activity. *Human Genetics*, *105*(3), 240–243.

Calderwood, S., Kilpatrick, L., Douglas, S. D., Freedman, M., Smith-Whitley, K., Rolland, M., & Kurtzberg, J. (2001). Recombinant human granulocyte colony-stimulating factor therapy for patients with neutropenia and/or neutrophil dysfunction secondary to glycogen storage disease type 1b. *Blood*, *97*, 376–382.

Carlin, M. P., Scherrer, D. Z., De Tommaso, A. M. A., Bertuzzo, C. S., & Steiner, C. E. (2013). Determining mutations in *G6PC* and *SLC37A4* genes in a sample of Brazilian patients with glycogen storage disease types 1a and 1b. *Genetics and Molecular Biology*, *36*(4), 502–506.

Carreras-Sureda, A., Pihán, P., & Hetz, C. (2017). The unfolded protein response: At the intersection between endoplasmic reticulum function and mitochondrial bioenergetics. *Frontiers in Oncology*, *7*, 55.

Carrillo-Carrasco, N., Chandler, R. J., & Venditti, C. P. (2012). Combined methylmalonic acidemia and homocystinuria, cblC type I. Clinical presentations, diagnosis and management. *Journal of Inherited Metabolic Disease*, *35*(1), 91–102.

Cartegni, L., Chew, S. L., & Krainer, A. R. (2002). Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nature Reviews Genetics*, *3*(4), 285–298.

Castle, J. C. (2011). SNPs occur in regions with less genomic sequence conservation. *PLoS ONE, 6*(6), e 20660.

Cerone, R., Schiaffino, M. C., Caruso, U., Lupino, S., & Gatti, R. (1997). Minor facial animalies in combined methylmalonic aciduria and homocystinuria due to a defect in cobalamin metabolism. *Journal of Inherited Metabolic Disease*, *22*, 247–250.

Champattanachai, V., Cairns, J. R. K., Shotelersuk, V., Keeratichamroen, S., Sawangareetrakul, P., Srisomsap, C., … Svasti, J. (2003). Novel mutations in a Thai patient with methylmalonic acidemia. *Molecular Genetics and Metabolism*, *79*(4), 300–302.

Chang, J. T., Chen, Y. Y., Liu, T. T., Liu, M. Y., & Chiu, P. C. (2011). Combined methylmalonic aciduria and homocystinuria cblC type of a Taiwanese infant with c.609G>A and c.567dupT mutations in the *MMACHC* gene. *Pediatrics and Neonatology*, *52*(4), 223–226.

Chen, L.-Y., Lin, B., Pan, C.-J., Hiraiwa, H., & Chou, J. Y. (2000). Structural requirements for the stability and microsomal transport activity of the human glucose 6-phosphate transporter. *Journal of Biological Chemistry*, *275*(44), 34280–34286.

Chen, L.-Y., Pan, C.-J., Shieh, J.-J., & Chou, J. Y. (2002). Structure–function analysis of the glucose-6-phosphate transporter deficient in glycogen storage disease type Ib. *Human Molecular Genetics*, *11*(25), 3199–3207.

Chen, M. (2011). Glycogen storage diseases. In S. Monga (Eds) *Molecular pathology of liver diseases* (pp. 677–681). Boston, MA: Springer.

Cheng, A., Zhang, M., Gentry, M. S., Worby, C. A., Dixon, J. E., & Saltiel, A. R. (2007). A role for *AGL* ubiquitination in the glycogen storage disorders of Lafora and Cori's disease. *Genes Development*, *21*(19), 2399–2409.

Cheng, A., Zhang, M., Okubo, M., Omichi, K., & Saltiel, A. R. (2009). Distinct mutations in the glycogen debranching enzyme found in glycogen storage disease type III lead to impairment in diverse cellular functions. *Human Molecular Genetics*, *18*(11), 2045–2052.

Choi, R., Park, H.-D., Ko, J. M., Lee, J., Lee, D. H., Hong, S. J., … Choe, Y. H. (2017). Novel *SLC37A4* mutations in Korean patients with glycogen storage disease 1b. *Annals of Laboratory Medicine*, *37*, 261–266.

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., & Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE*, *7*(10), e46688.

Chou, J. Y., Jun, H. S., & Mansfield, B. C. (2010). Neutropenia in type Ib glycogen storage disease. *Current Opinion in Hematology*, *17*(1), 36–42.

Chou, J. Y., & Mansfield, B. C. (2014). The SLC37 family of sugar-phosphate/phosphate exchangers. *Current Topics in Membranes*, *73*, 357–382.

Chu, J., Pupavac, M., Watkins, D., Tian, X., Feng, Y., Chen, S., … Rosenblatt, D. S. (2016). Next generation sequencing of patients with mut methylmalonic aciduria: Validation of somatic cell studies and identification of 16 novel mutations. *Molecular Genetics and Metabolism*, *118*(4), 264–271.

Coleman, R. A., Winter, H. S., Wolf, B., Gilchrist, J. M., & Chen, Y. T. (1992). Glycogen storage disease type III (glycogen debranching enzyme deficiency): Correlation of biochemical defects with myopathy and cardiomyopathy. *Annals of Internal Medicine*, *116*(11), 896–900.

Cross, J. L., Iben, J., Simpson, C. L., Thurm, A., Swedo, S., Tierney, E., … Wassif, C. A. (2015). Determination of the allelic frequency in Smith-Lemli-Opitz syndrome by analysis of massively parallel sequencing data sets. *Clinical Genetics*, *87*(6), 570–575.

Crushell, E., Treacy, E. P., Dawe, J., Durkie, M., & Beauchamp, N. J. (2010). Glycogen storage disease type III in the Irish population. *Journal of Inherited Metabolic Disease*, *33*(3), 215-218.

Das, A. M., Steuerwald, U., & Illsinger, S. (2010). Inborn errors of energy metabolism associated with myopathies. *Journal of Biomedicine and Biotechnology*, *2010*, 340849.

de Ligt, J., Willemsen, M. H., van Bon, B., Kleefstra, T., Yntema, H. G., Kroes, T., … Vissers, L. E. L. M. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. *The New England Journal of Medicine, 367*(20), 1921-1929.

Dieckgraefe, B. K., Korzenik, J. R., Husain, A., & Dieruf, L. (2002). Association of glycogen storage disease 1b and Crohn disease: Results of a North American survey. *European Journal of Pediatrics*, *161*, 88–92.

Ding, J. H., de Barsy, T., Brown, B. I., Coleman, R. A., & Chen, Y. T. (1990). Immunoblot analyses of glycogen debranching enzyme in different subtypes of glycogen storage disease type III. *The Journal of Pediatrics*, *116*(1), 95–100.

Drennan, C. L., Matthews, R. G., Rosenblatt, D. S., Ledley, F. D., Fenton, W. A., & Ludwig, M. L. (1996). Molecular basis for dysfunction of some mutant forms of methylmalonyl-CoA mutase : Deductions from the structure of methionine synthase source. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(11), 5550–5555.

Dronamraju, K. (1992). Profiles in genetics: Archibald E. Garrod (1857-1936). *American Journal of Human Genetics, 51*(1), 216–219.

Dündar, H., Özgül, K., Dursun, A., Sivri, S., Aliefendio, D., & Co, T. (2012). Microarray based mutational analysis of patients with methylmalonic acidemia : Identification of 10 novel mutations. *Molecular Genetics and Metabolism, 106*(4), 419–423.

Dweikat, I. M., Alawneh, I. S., Bahar, S. F., & Sultan, M. I. (2016). Fanconi-Bickel syndrome in two Palestinian children: Marked phenotypic variability with identical mutation. *BMC Research Notes*, *9*, 387.

Ekbote, A. V., Mandal, K., Agarwal, I., Sinha, R., & Danda, S. (2012). Fanconi- Bickel syndrome: Mutation in an Indian patient. *Indian Journal of Pediatrics*, *79*(6), 810–812.

Endo, Y., Fateen, E., Aoyama, Y., Horinishi, A., Ebara, T., Murase, T., … Okubo, M. (2005). Molecular characterization of Egyptian patients with glycogen storage disease type IIIa. *Journal of Human Genetics*, *50*(10), 538–542.

Endo, Y., Horinishi, A., Vorgerd, M., Aoyama, Y., Ebara, T., Murase, T., … Okubo, M. (2006). Molecular analysis of the *AGL* gene: Heterogeneity of mutations in patients with glycogen storage disease type III from Germany, Canada, Afghanistan, Iran and Turkey. *Journal of Human Genetics*, *51*(11), 958–963.

Faesen, A. C., Dirac, A. M. G., Shanmugham, A., Ovaa, H., Perrakis, A., & Sixma, T. K. (2011). Mechanism of USP7/HAUSP activation by its C-terminal ubiquitin-like domain and allosteric regulation by GMP-synthetase. *Molecular Cell*, *44*(1), 147-159.

Falkowska, A., Gutowska, I., Goschorska, M., Nowacki, P., Chlubek, D., & Baranowska-Bosiacka, I. (2015). Energy metabolism of the brain, including the cooperation between astrocytes and neurons, especially in the context of glycogen metabolism. *International Journal of Molecular Sciences*, *16*(11), 25959–25981.

Fang, Y. C., & Cheng, M. (2002). The effect of C-terminal mutations of HSP60 on protein folding. *Journal of Biomedical Science*, *9*(3), 223–233.

Faustino, N. A., Cooper, T. A, & Andre, N. (2003). Pre-mRNA splicing and human disease. *Genes & Development*, *17*(4), 419–437.

Fenton, W. A, Gravel, R. A, & Rosenblatt, D. S. (2001). Disorders of propionate and methylmalonate metabolism. In *The online metabolic and molecular bases of inherited diseases* (pp. 2165–2194). New York, NY: McGraw-Hill Education.

Ferrari, M., Cremonesi, L., & Stenirri, S. (2008). Post-natal molecular diagnosis of inherited diseases. *Journal of the International Federation of Clinical Chemistry and Laboratory Medicine*, *19*(1), 7–12.

Fischer-Zirnsak, B., Escande-beillard, N., Ganesh, J., Tan, Y. X., Bughaili, M. Al, Lin, A. E., … Callewaert, B. (2015). Recurrent de novo mutations affecting residue arg138 of pyrroline-5-carboxylate synthase cause a progeroid form of autosomal-dominant cutis laxa. *American Journal of Human Genetics, 97*(3), 483–492.

Forny, P., Froese, D. S., Suormala, T., Yue, W. W., & Baumgartner, M. R. (2014). Functional characterization and categorization of missense mutations that cause methylmalonyl-CoA mutase (MUT) deficiency. *Human Mutation*, *35*(12), 1449–1458.

Fowler, B., Leonard, J. V., & Baumgartner, M. R. (2008). Causes of and diagnostic approach to methylmalonic acidurias. *Journal of Inherited Metabolic Disease*, *31*(3), 350–360.

Freitag, J., Ast, J., & Bölker, M. (2012). Cryptic peroxisomal targeting via alternative splicing and stop codon read-through in fungi. *Nature*, *485*(7399), 522–525.

Froese, D. S., Zhang, J., Healy, S., & Gravel, R. A. (2009). Mechanism of vitamin B12-responsiveness in cblC methylmalonic aciduria with homocystinuria. *Molecular Genetics and Metabolism*, *98*(4), 338–343.

Froissart, R., Piraud, M., Boudjemline, A. M., Vianey-Saban, C., Petit, F., Hubert-Buron, A., … Labrune, P. (2011). Glucose-6-phosphatase deficiency. *Orphanet Journal of Rare Diseases*, *6*, 27.

Frousios, K., Iliopoulos, C. S., Schlitt, T., & Simpson, M. A. (2013). Predicting the functional consequences of non-synonymous DNA sequence variants - evaluation of bioinformatics tools and development of a consensus strategy. *Genomics*, *102*(4), 223–228.

Fuchshuber, A., Mucha, B., Baumgartner, E. R., Vollmer, M., & Hildebrandt, F. (2000). Mut(0) methylmalonic acidemia: Eleven novel mutations of the methylmalonyl-CoA mutase including a deletion-insertion mutation. *Human Mutation*, *16*(2), 179.

Fukumoto, H., Seino, S., Imura, H., Seino, Y., Eddy, R. L., Fukushima, Y., … Bell, G. I. (1988). Sequence, tissue distribution, and chromosomal localization of mRNA encoding a human glucose transporter-like protein. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(15), 5434–5438.

Ganesh, S., Agarwala, K. L., Ueda, K., Akagi, T., Shoda, K., Usui, T., … Yamakawa, K. (2000). Laforin, defective in the progressive myoclonus epilepsy of Lafora type, is a dual-specificity phosphatase associated with polyribosomes. *Human Molecular Genetics*, *9*(15), 2251–2261.

Gerin, I., Viega-da-Cunha, M., Achouri, Y., Collet, J. F., & VanSchaftingen, E. (1997). Sequence of a putative glucose-6-phosphate translocase, mutated in glycogen storage disease type 1b. *FEBS Letters*, *419*, 235–238.

Gerth, C., Morel, C. F., Feigenbaum, A., & Levin, A. V. (2008). Ocular phenotype in patients with methylmalonic aciduria and homocystinuria, cobalamin C type. *Journal of AAPOS*, *12*(6), 591–596.

Ghazarian, H., Idoni, B., & Oppenheimer, S. B. (2011). A glycobiology review: Carbohydrates, lectins, and implications in cancer therapeutics. *Acta Histochemica*, *113*(3), 236–247.

Ghoraba, D. A., Mohammed, M. M., & Zaki, O. K. (2015). Mutation analysis of methylmalonyl-CoA mutase gene exon 2 in Egyptian families: Identification of 25 novel allelic variants. *Meta Gene*, *3*, 71–88.

Gizicki, R., Robert, M. C., Gómez-López, L., Orquin, J., Decarie, J. C., Mitchell, G. A., … Ospina, L. H. (2014). Long-term visual outcome of methylmalonic aciduria and homocystinuria, cobalamin C type. *Ophthalmology*, *121*(1), 381–386.

Goldberg, A. L. (2003). Protein degration and protection against misfolded proteins. *Nature*, *426*(6968), 895–899.

Goldmann, J. M., Wong, W. S. W., Pinelli, M., Farrah, T., Bodian, D., Stittrich, A. B., … Niederhuber, J. E. (2016). Parent-of-origin-specific signatures of de novo mutations. *Nature Genetics*, *48*(8), 935-939.

Goldstein, J. L., Austin, S. L., Boyette, K., & Kanaly, A. (2010). Molecular analysis of the *AGL* gene: Identification of 25 novel mutations and evidence of genetic heterogeneity in patients with glycogen storage disease type III. *Genetics in Medicine*, *12*(7), 424-430.

Gould, G. W., Thomas, H. M., Jess, T. J., & Bell, G. I. (1991). Expression of human glucose transporters in xenopus oocytes: Kinetic characterization and substrate specificities of the erythrocyte, liver, and brain isoforms. *Biochemistry*, *30*(21), 5139–5145.

Gregoire, S., Irwin, J., & Kwon, I. (2013). Technique for monitoring protein misfolding and aggregation in vitro and in living cells. *The Korean Journal of Chemical Engineering*, *29*(6), 693–702.

Gupta, N., Nambam, B., Weinstein, D. A., & Shoemaker, L. R. (2016). Late diagnosis of Fanconi-Bickel syndrome: Challenges with the diagnosis and literature review. *Journal of Inborn Errors of Metabolism & Screening*, *4*, 1–6.

Hadjigeorgiou, G. M., Comi, G. P., Bardoni, N., Shen, Y. T., Salani, S., Toscano, A., … Lucchiari, S. (1999). Novel donor splice site mutations of *AGL* gene in glycogen storage disease type IIIa. *Journal of Inherited Metabolic Disease*, *22*, 762–763.

Hamaguchi, T., Nakajima, H., Noguchi, T., Ono, A., Kono, N., Tarui, S., … Matsuzawa, Y. (1994). A new variant of muscle phosphofructokinase deficiency in a Japanese case with abnormal RNA splicing. *Biochemical and Biophysical Research Communications*, *202*, 444–449.

Hammond, S. M., & Wood, M. J. A. (2011). Genetic therapies for RNA mis-splicing diseases. *Trends in Genetics*, *27*, 196–205.

Han, L. S., Huang, Z., Han, F., Ye, J., Qiu, W. J., Zhang, H. W., … Gu, X. F. (2015). Clinical features and *MUT* gene mutation spectrum in Chinese patients with isolated methylmalonic acidemia: Identification of ten novel allelic variants. *World Journal of Pediatrics*, *11*(4), 358–365.

Hannibal, L., Kim, J., Brasch, N. E., Wang, S., Rosenblatt, D. S., Banerjee, R., & Jacobsen, D. W. (2009). Processing of alkylcobalamins in mammalian cells: A role for the *MMACHC* (cblC) gene product. *Molecular Genetics and Metabolism*, *97*(4), 260–266.

Hendriksz, C. J., & Gissen, P. (2014). Glycogen storage disease. *Paediatrics and Child Health*, *25*(3), 139–144.

Hiraiwa, H., Pan, C., Moses, S. W., Chou, J. Y., Lin, B., & Yang, J. (1999). Inactivation of the glucose 6-phosphate transporter causes glycogen storage disease type 1b. *The Journal of Biological Chemistry*, *274*(9), 5532–5536.

Hofhuis, J., Schueren, F., Notzel, C., Lingner, T., Gartner, J., Jahn, O., & Thoms, S. (2016). The functional readthrough extension of malate dehydrogenase reveals a modification of the genetic code. *Open Biology*, *6*(11), 160246.

Horinishi, A., Okubo, M., Tang, N. L. S., Hui, J., & Tomohito, K. T. (2002). Mutational and haplotype analysis of *AGL* in patients with glycogen storage disease type III. *Journal of Human Genetics, 47*(2), 55-59.

Hörster, F., & Hoffmann, G. F. (2004). Pathophysiology, diagnosis, and treatment of methylmalonic aciduria-recent advances and new challenges. *Pediatric Nephrology*, *19*(10), 1071–1074.

Houck, S. A., Singh, S., & Cyr, D. M. (2012). Cellular responses to misfolded proteins and protein aggregates. *Methods in Molecular Biology*, *832*(3), 455–461.

Huemer, M., Scholl-Bürgi, S., Hadaya, K., Kern, I., Beer, R., Seppi, K., … Karall, D. (2014). Three new cases of late-onset cblC defect and review of the literature illustrating when to consider inborn errors of metabolism beyond infancy. *Orphanet Journal of Rare Diseases*, *9*(1), 161.

Ihara, K., Takabayashi, A., Terasaki, K., & Hara, T. (1998). Assignment of glucose 6-phosphate translocase (G6PT1) to human chromosome band 11q23.3 by in situ hybridization. *Cytogenetics and Cell Genetics*, *83*, 50–51.

Imtiaz, F., Al-Mubarak, B. M., Al-Mostafa, A., Al-Hamed, M., Allam, R., Al-Hassnan, Z., … Al-Sayed, M. (2015). Spectrum of mutations in 60 Saudi patients with mut methylmalonic acidemia. *JIMD Reports, 29*, 39-46.

Jackson, M. R., Nilsson, T., & Peterson, P. A. (1990). Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *The EMBO Journal*, *9*(10), 3153–3162.

Jansen, R., & Ledley, F. D. (1990). Heterozygous mutations at the mut locus in fibroblast with mut0 methylmalonic acidemia identified by polymerase-chain-reaction cDNA cloning. *American Journal of Human Genetics, 47(5), 808*–814.

Jenkins, G. J., Suzen, H. S., Sueiro, R. A., & Parry, J. M. (1999). The restriction site mutation assay: A review of the methodology development and the current status of the technique. *Mutagenesis*, *14*(5), 439–48.

Jensen, J., Rustad, P. I., Kolnes, A. J., & Lai, Y. (2011). The role of skeletal muscle glycogen breakdown for regulation of insulin sensitivity by exercise. *Frontiers in Physiology*, *2*, 112.

Jequier, E. (1994). Carbohydrates as a source of energy. *The American Journal of Clinical Nutrition*, *59*(3), 682S–685S.

Jung, J. W., Hwang, I. T., Park, J. E., Lee, E. H., Ryu, K. H., Kim, S. H., & Hwang, J. S. (2005). Mutation analysis of the *MCM* gene in Korean patients with MMA. *Molecular Genetics and Metabolism*, *84*(4), 367–370.

Kazemi-Esfarjani, P., Beitel, L. K., Trifiro, M., Kaufman, M., Rennie, P., Sheppard, P., … Pinsky, L. (1993). Substitution of valine-865 by methionine or leucine in the human androgen receptor causes complete or partial androgen insensitivity, respectively with distinct androgen receptor phenotypes. *Molecular Endocrinology*, *7*(1), 37–46.

Keyfi, F., Abbaszadegan, M. R., Rolfs, A., & Orolicki, S. (2016). Identification of a novel deletion in the *MMAA* gene in two Iranian siblings with vitamin B12-responsive methylmalonic acidemia. *Cellular & Molecular Biology Letters*, *21*, 4.

Kim, J., Gherasim, C., & Banerjee, R. (2008). Decyanation of vitamin B12 by a trafficking chaperone. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38), 14551–14554.

Kim, J., Hannibal, L., Gherasim, C., Jacobsen, D. W., & Banerjee, R. (2009). A human vitamin B12 trafficking protein uses glutathione transferase activity for processing alkylcobalamins. *Journal of Biological Chemistry*, *284*(48), 33418–33424.

Kishnani, P. S., Austin, S. L., Abdenur, J. E., Arn, P., Bali, D. S., Boney, A., … Watson, M. S. (2014). Diagnosis and management of glycogen storage disease type I: A practice guideline of the American College of Medical Genetics and Genomics. *Genetics in Medicine, 16*(11), e1.

Kishnani, P. S., Austin, S. L., Arn, P., & Bali, D. S. (2010). Glycogen storage disease type III diagnosis and management guidelines. *Genetics in Medicine*, *12*(7), 446-463.

Ko, J. S., Moon, J. S., Seo, J. K., Yang, H. R., Chang, J. Y., & Park, S. S. (2013). A mutation analysis of the *AGL* gene in Korean patients with glycogen storage disease type III. *Journal of Human Genetics*, *59*(1), 42–45.

Kong, A., Frigge, M. L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., … Wong, W. S. W. (2012). Rate of de novo mutations and the importance of father's age to disease risk. *Nature*, *488*(7412), 471–475.

Kure, S., Suzuki, Y., Matsubara, Y., Sakamoto, O., Shintaku, H., Isshiki, G., … Narisawa, K. (1998). Molecular analysis of glycogen storage disease type Ib: Identification of a prevalent mutation among Japanese patients and assignment of a putative glucose-6-phosphate translocase gene to chromosome 11. *Biochemical and Biophysical Research Communications*, *248*(2), 426–431.

Lachaal, M., Rampal, A. L., Ryu, J., Lee, W., Hah, J. S., & Jung, C. Y. (2000). Characterization and partial purification of liver glucose transporter GLUT2. *Biochimica et Biophysica Acta - Biomembranes*, *1466*(1–2), 379–389.

Lam, C., Lee, A. T., Lam, Y., Wong, T., Mak, T. W., Fung, W., … Tong, S. (2004). DNA-based subtyping of glycogen storage disease type III: Mutation and haplotype analysis of the *AGL* gene in Chinese. *Molecular Genetics and Metabolism*, *83*, 271–275.

Ledley, F. D. (1990). Perspectives on methymalonic acidemia resulting from molecular cloning of methylmalonyl-CoA mutase. *BioEssays*, *12*(7), 335–340.

Ledley, F. D., Crane, A. M., & Lumetta, M. (1990). Heterogeneous alleles and expression of methylmalonyl-CoA mutase in mut methylmalonic acidemia. *American Journal of Human Genetics*, *46*(3), 539–547.

Ledley, F. D., Jansen, R., Nham, S. U., Fenton, W. A., & Rosenberg, L. E. (1990). Mutation eliminating mitochondrial leader sequence of methylmalonyl-CoA mutase causes mut0 methylmalonic acidemia. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(8), 3147–3150.

Ledley, F. D., Lumetta, M. R., Zoghbi, H. Y., VanTuinen, P., Ledbetter, S. A., & Ledbetter, D. H. (1988). Mapping of human methylmalonyl-CoA mutase (*MUT*) locus on chromosome 6. *American Journal of Human Genetics*, *42*(6), 839–846.

Ledley, F. D., & Rosenblatt, D. S. (1997). Mutations in mut methylmalonic acidemia: Clinical and enzymatic correlations. *Human Mutation*, *9*(1), 1–6.

Lempp, T. J., Suormala, T., Siegenthaler, R., Baumgartner, E. R., Fowler, B., Steinmann, B., & Baumgartner, M. R. (2007). Mutation and biochemical analysis of 19 probands with mut0 and 13 with mut- methylmalonic aciduria: Identification of seven novel mutations. *Molecular Genetics and Metabolism*, *90*(3), 284–290.

Leong, Y. H., Gan, C. Y., Tan, M. A., & Majid, M. I. (2014). Present status and future concerns of expanded newborn screening in Malaysia : Sustainability, challenges and perspectives. *The Malaysian Journal of Medical Sciences*, *21*(2), 63–67.

Lerner-Ellis, J. P., Tirone, J. C., Pawelek, P. D., Doré, C., Atkinson, J. L., Watkins, D., … Rosenblatt, D. S. (2006). Identification of the gene responsible for methylmalonic aciduria and homocystinuria, cblC type. *Nature Genetics*, *38*(1), 93–100.

Leslie, E., & Murray, J. (2013). Evaluating rare coding variants as contributing causes to non-syndromic cleft lip and palate. *Clinical Genetics*, *84*(5), 496–500.

Lin, B., Pan, C. J., & Chou, J. Y. (2000). Human variant glucose-6-phosphate transporter is active in microsomal transport. *Human Genetics*, *107*(5), 526–529.

Liu, H.-X., Cartegni, L., Zhang, M. Q., & Krainer, A. R. (2001). A mechanism for exon skipping caused by nonsense or missense mutations in *BRCA1* and other genes. *Nature Genetics*, *27*(1), 55–58.

Liu, M.-Y., Liu, T.-T., Yang, Y.-L., Chang, Y.-C., Fan, Y.-L., Lee, S.-F., & Teng, Y.-T. (2012). Mutation profile of the *MUT* gene in Chinese methylmalonic aciduria patients. *JIMD Reports*, *4*, 113–116.

Lomako, J., Lomako, W. M., & Whelan, W. J. (2004). Glycogenin: The primer for mammalian and yeast glycogen synthesis. *Biochimica et Biophysica Acta - General Subjects*, *1673*(1–2), 45–55.

Lucchiari, S., Donati, M. A., Parini, R., Melis, D., Gatti, R., & Bresolin, N. (2002). Molecular characterisation of GSD III subjects and identification of six novel mutations in *AGL*. *Human Mutation*, *20*(6), 480.

Lucchiari, S., Fogh, I., Prelle, A., Parini, R., Bresolin, N., Melis, D., … Scarlato, G. (2002). Clinical and genetic variability of glycogen storage disease type IIIa: Seven novel *AGL* gene mutations in the Mediterranean area. *American Journal of Medical Genetics*, *190*(3), 183–190.

Lucchiari, S., Santoro, D., Pagliarani, S., & Comi, G. P. (2007). Clinical , biochemical and genetic features of glycogen debranching enzyme deficiency. *Acta Myologica*, *26*(1), 72–74.

Lynch, M. (2010). Rate, molecular spectrum, and consequences of human mutation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(3), 961–968.

Macintyre, S., & Sooman, A. (1991). Non-paternity and prenatal genetic screening. *Lancet*, *338*(8771), 869–871.

Mahdieh, N., & Rabbani, B. (2013). An overview of mutation detection methods in genetic disorders. *Iranian Journal of Pediatrics*, *23*(4), 375–388.

Mak, C. M., Lee, H. H., Chan, A. Y., & Lam, C. (2013). Inborn errors of metabolism and expanded newborn screening: Review and update. *Critical Reviews in Clinical Laboratory Sciences, 50*(6), 142-162.

Malkki, M., & Petersdorf, E. W. (2012). Genotyping of single nucleotide polymorphisms by 5' nuclease allelic discrimination. *Methods in Molecular Biology*, *882*, 173-182.

Mancia, F., Keep, N. H., Nakagawa, a, Leadlay, P. F., McSweeney, S., Rasmussen, B., … Evans, P. R. (1996). How coenzyme B12 radicals are generated: The crystal structure of methylmalonyl-coenzyme A mutase at 2 A resolution. *Structure*, *4*(3), 339–350.

Marcolongo, P., Barone, V., Priori, G., Pirola, B., Giglio, S., Biasucci, G., … Sorrentino, V. (1998). Structure and mutation analysis of the glycogen storage disease type Ib gene. *FEBS Letters*, *436*, 5–8.

Marsh, E. N., McKie, N., Davis, N. K., & Leadlay, P. F. (1989). Cloning and structural characterization of the genes coding for adenosylcobalamin-dependent methylmalonyl-CoA mutase from *Propionibacterium shermanii*. *The Biochemical Journal*, *260*(2), 345–352.

Melendez-Hevia, E., Waddell, T. G., & Shelton, E. D. (1993). Optimization of molecular design in the evolution of metabolism: The glycogen molecule. *The Biochemical Journal, 295*(2), 477–483.

Melis, D., Pivonello, R., Parenti, G., Casa, R. Della, Salerno, M., Lombardi, G., … Andria, G. (2007). Increased prevalence of thyroid autoimmunity and hypothyroidism in patients with glycogen storage disease type I. *The Journal of Pediatrics*, *150*(3), 300–305.

Mercer, T. R., Clark, M. B., Andersen, S. B., Brunck, M. E., Haerty, W., Crawford, J., … Mattick, J. S. (2015). Genome-wide discovery of human splicing branchpoints. *Genome Resource*, *25*(2), 290–303.

Micanovic, R., Raches, D. W., Dunbar, J. D., Driver, D. A., Bina, H. A., Dickinson, C. D., & Kharitonenkov, A. (2009). Different roles of N- and C-termini in the functional activity of FGF21. *Journal of Cellular Physiology*, *219*(2), 227–234.

Michau, A., Guillemain, G., Grosfeld, A., Vuillaumier-Barrot, S., Grand, T., Keck, M., … Le Gall, M. (2013). Mutations in *SLC2A2* gene reveal hGLUT2 function in pancreatic βcell development. *Journal of Biological Chemistry*, *288*(43), 31080–31092.

Mikami, H., Ogasawara, M., Matsubara, Y., Kikuchi, M., Miyabayashi, S., Kure, S., & Narisawa, K. (1999). Molecular analysis of methylmalonyl-CoA mutase deficiency: Identification of three missense mutations in mut0 patients. *Journal of Human Genetics*, *44*(1), 35–39.

Mili, A., Charfeddine, I. Ben, Mamaı, O., Abdelhak, S., Adala, L., Amara, A., … Gribaa, M. (2011). Molecular and biochemical characterization of Tunisian patients with glycogen storage disease type III. *Journal of Human Genetics*, *57*(3), 170-175.

Mili, A., I, B. C., Amara, A., Mamaï, O., Adala, L., T, B. L., … Gribaa, M. A. (2012). A c.3216 _ 3217delGA mutation in *AGL* gene in Tunisian patients with a glycogen storage disease type III: Evidence of a founder effect. *Clinical Genetics*, (4), 534–539.

Morel, C. F., Lerner-Ellis, J. P., & Rosenblatt, D. S. (2006). Combined methylmalonic aciduria and homocystinuria (cblC): Phenotype-genotype correlations and ethnic-specific observations. *Molecular Genetics and Metabolism*, *88*(4), 315–321.

Mueckler, M., Caruso, C., Baldwin, S. A., Allard, W. J., Lienhard, G. E., & Lodish, H. F. (1985). Sequence and structure of a human glucose transporter. *Science, 229*(4717), 941-945.

Mueckler, M., Kruse, M., Strube, M., Riggs, A. C., Chiu, K. C., & Permutt, M. A. (1994). A mutation in the *GLUT2* glucose transporter gene of a diabetic patient abolishes transport activity. *Journal of Biological Chemistry*, *269*(27), 17765–17767.

Muraoka, a, Hashiramoto, M., Clark, A. E., Edwards, L. C., Sakura, H., Kadowaki, T., … Kasuga, M. (1995). Analysis of the structural features of the C-terminus of GLUT1 that are required for transport catalytic activity. *The Biochemical Journal*, *311*(2), 699-704.

Nakajima, H., Kono, N., Yamasaki, T., Hotta, K., Kawachi, M., Kuwajima, M., … Tarui, S. (1990). Genetic defect in muscle phosphofructokinase deficiency. *The Journal of Biological Chemistry*, *265*, 9392–9395.

Narisawa, K., Igarashi, Y., Arai, N., Otake, M., Tada, K., & Kuzuya, T. (1983). Glycogen storage disease type lb: Microsomal glucose-6-phosphatase system in two patients with different clinical findings. *Pediatric Research*, *17*(7), 545–549.

Narisawa, K., Igarashi, Y., Otomo, H., & Tada, K. (1978). A new variant of glycogen storage disease type I probably due to a defect in the glucose-6-phosphate transport system. *Biochemical and Biophysical Research Communications*, *83*(4), 1360–1364.

Nebert, D. W., & Russell, D. W. (2002). Clinical importance of the cytochromes P450. *Lancet*, *360*, 1155–1162.

Ng, P. C., & Henikoff, S. (2001). Predicting deleterious amino acid substitutions. *Genome Research*, *11*, 863–874.

Ni, M., & Lee, A. S. (2007). ER chaperones in mammalian development and human diseases. *FEBS Letters*, *581*(19), 3641–3651.

Nogueira, C., Aiello, C., Cerone, R., Martins, E., Caruso, U., Moroni, I., … Dionisi-Vici, C. (2008). Spectrum of *MMACHC* mutations in Italian and Portuguese patients with combined methylmalonic aciduria and homocystinuria, cblC type. *Molecular Genetics and Metabolism*, *93*(4), 475–480.

Oka, Y., Asano, T., Shibasaki, Y., Lin, J.-L., Tsukuda, K., Katagiri, H., … Takaku, F. (1990). C-terminal truncated glucose transporter is locked into an inward-facing form without transport activity. *Nature*, *345*(6275), 550–553.

Okubo, M., Aoyama, Y., & Murase, T. (1996). A novel donor splice site mutation in the glycogen debranching enzyme gene is associated with glycogen storage disease type III. *Biochemical and Biophysical Research Communications*, *224*, 493–499.

Okubo, M., Horinishi, A., Nakamura, N., Aoyama, Y., Hashimoto, M., Endo, Y., & Murase, T. (1998). A novel point mutation in an acceptor splice site of intron 32 (IVS32 A-G) but no exon 3 mutations in the glycogen debranching enzyme gene in a homozygous patient with glycogen storage disease type IIIb. *Human Genetics*, *102*(1), 1–5.

Okubo, M., Horinishi, A., Suzuki, Y., Murase, T., & Hayasaka, K. (2000). Compound heterozygous patient with glycogen storage disease type III: Identification of two novel *AGL* mutations, a donor splice site mutation of Chinese origin and a 1-bp deletion of Japanese origin. *American Journal of Medical Genetics*, *93*(3), 211–214.

Okubo, M., Horinishi, A., Takeuchi, M., Suzuki, Y., Sakura, N., Hasegawa, Y., … Murase, T. (2000). Heterogeneous mutations in the glycogen debranching enzyme gene are responsible for glycogen storage disease type IIIa in Japan. *Human Genetics*, *106*(1), 108–115.

Orci, L., Thorens, B., Ravazzola, M., & Lodish, H. F. (1989). Localization of the pancreatic beta cell glucose transporter to specific plasma membrane domains. *Science*, *245*(4915), 295–297.

Özen, H. (2007). Glycogen storage diseases: New perspectives. *World Journal of Gastroenterology*, *13*(18), 2541–2553.

Pan, C.-J., Chen, L.-Y., Mansfield, B. C., Salani, B., Varesio, L., & Chou, J. Y. (2003). The signature motif in human glucose-6-phosphate transporter is essential for microsomal transport of glucose-6-phosphate. *Human Genetics*, *112*, 430–433.

Pan, C.-J., Chen, S.-Y., Lee, S., & Chou, J. Y. (2009). Structure–function study of the glucose-6-phosphate transporter, an eukaryotic antiporter deficient in glycogen storage disease type Ib. *Molecular Genetics and Metabolism*, *96*, 32–37.

Pan, C.-J., Lin, B., & Chou, J. Y. (1999). Transmembrane topology of human glucose 6-phosphate transporter. *The Journal of Biological Chemistry*, *274*, 13865–13869.

Parvari, R., Moses, S., Shen, J., Hershkovitz, E., Lerner, A., & Chen, Y. T. (1997). A single-base deletion in the 3'-coding region of glycogen debranching enzyme is prevalent in glycogen storage disease type IIIa in a population of North African Jewish patients. *European Journal of Human Genetics: EJHG*, *5*(5), 266–70.

Permutt, M. A., Koranyi, L., Kellert, K., Lacy, P. E., & Scharp, D. W. (1989). Cloning and functional expression of a human pancreatic islet glucose-transporter cDNA. *Proceedings of the National Academy of Sciences of the Unites States of America*, *86*(22), 8688–8692.

Pickart, C. M. (2001). Mechanisms underlying ubiquitination. *Annual Review of Biochemistry*, *70*(1), 503–33.

Rahbari, R., Wuster, A., Lindsay, S. J., Hardwick, R. J., Ludmil, B., Stratton, M. R., … Hurles, M. E. (2016). Timing, rates and spectra of human germline mutation. *Nature Genetics*, *48*(2), 126–133.

Ramensky, V., Bork, P., & Sunyaev, S. (2002). Human non-synonymous SNPs: Server and survey. *Nucleic Acids Research*, *30*(17), 3894–3900.

Rampal, S., Mahadeva, S., Guallar, E., Bulgiba, A., Mohamed, R., Rahmat, R., … Rampal, L. (2012). Ethnic differences in the prevalence of metabolic syndrome: Results from a multi-ethnic population-based survey in Malaysia. *PLoS ONE*, *7*(9), 8–10.

Roach, P. J., Depaoli‐Roach, A. A., Hurley, T. D., & Tagliabracci, V. S. (2012). Glycogen and its metabolism: Some new developments and old themes. *Biochemical Journal*, *441*(3), 763–787.

Rosenblatt, D. S., Aspler, A. L., Shevell, M. I., Pletcher, B. A., Fenton, W. A., & Seashore, M. R. (1997). Clinical heterogeneity and prognosis in combined methylmalonic aciduria and homocystinuria (cblC). *Journal of Inherited Metabolic Disease*, *20*(4), 528–538.

Rost, S., Fregin, A., Ivaskevicius, V., Conzelmann, E., Hörtnagel, K., Pelz, H.-J., … Oldenburg, J. (2004). Mutations in *VKORC1* cause warfarin resistance and multiple coagulation factor deficiency type 2. *Nature*, *427*(6974), 537–541.

Sakamoto, O., Ogawa, E., Ohura, T., Igarashi, Y., Matsubara, Y., Narisawa, K., & Iinuma, K. (2000). Mutation analysis of the *GLUT2* gene in patients with Fanconi-Bickel syndrome. *Pediatric Research*, *48*(5), 586–589.

Santer, R., Groth, S., Kinner, M., Dombrowski, A., Berry, G. T., Brodehl, J., … Schaub, J. (2002). The mutation spectrum of the facilitative glucose transporter gene *SLC2A2* (*GLUT2*) in patients with Fanconi-Bickel syndrome. *Human Genetics*, *110*(1), 21–29.

Santer, R., Kinner, M., Steuerwald, U., Kjaergaard, S., Skovby, F., Simonsen, H., … Schaub, J. (2001). Molecular genetic basis and prevalence of glycogen storage disease type IIIa in the Faroe Islands. *European Journal of Human Genetics*, *9*(5), 388–391.

Santer, R., Schneppenheim, R., Dombrowski, A., Götze, H., Steinmann, B., & Schaub, J. (1997). Mutations in *GLUT2*, the gene for the liver-type glucose transporter, in patients with Fanconi-Bickel syndrome. *Nature Genetics*, *17*(3), 324–326.

Santer, R., Schneppenheim, R., Suter, D., Schaub, J., & Steinmann, B. (1998). Fanconi-Bickel syndrome - The original patient and his natural history, historical steps leading to the primary defect, and a review of the literature. *European Journal of Pediatrics*, *157*(10), 783–797.

Sati, S. P., Singh, S. K., Kumar, N., & Sharma, A. (2002). Extra terminal residues have a profound effect on the folding and solubility of a *Plasmodium falciparum* sexual stage-specific protein over-expressed in *Escherichia coli*. *European Journal of Biochemistry*, *269*(21), 5259–5263.

Scotti, M. M., & Swanson, M. S. (2016). RNA mis-splicing in disease. *Nature Reviews Genetics*, *17*(1), 19-32.

Sentner, C. P., Hoogeveen, I. J., Weinstein, D. A., Santer, R., Murphy, E., McKiernan, P. J., … Derks, T. G. J. (2016). Glycogen storage disease type III: Diagnosis, genotype, management, clinical course and outcome. *Journal of Inherited Metabolic Disease*, *39*, 697–704.

Shaiu, W.-L., Kishnani, P. S., Shen, J., Liu, H.-M., & Chen, Y.-T. (2000). Genotype-phenotype correlation in two frequent mutations and mutation update in type III glycogen storage disease. *Molecular Genetics and Metabolism*, *69*(1), 16–23.

Shen, J., Bao, Y., & Chen, Y. (1997). A nonsense mutation due to a single base insertion in the 3′-coding region of glycogen debranching enzyme gene associated with a severe phenotype in a patient with glycogen storage disease type IIIa. *Human Mutation*, *40*, 37–40.

Shen, J., Bao, Y., Liu, H. M., Lee, P., Leonard, J. V, & Chen, Y. T. (1996). Mutations in exon 3 of the glycogen debranching enzyme gene are associated with glycogen storage disease type III that is differentially expressed in liver and muscle. *The Journal of Clinical Investigation*, *98*(2), 352–357.

Sherman, M. Y., & Goldberg, A. L. (2001). Cellular defenses against unfolded proteins: A cell biologist thinks about neurodegenerative diseases. *Neuron*, *29*(1), 15–32.

Şimşek, E., Savaş-Erdeve, Ş., Sakamoto, O., Doğanci, T., & Dallar, Y. (2009). A novel mutation of the *GLUT2* gene in a Turkish patient with Fanconi-Bickel syndrome. *The Turkish Journal of Pediatrics*, *51*(2), 166-168.

Skuzeski, J. M., Nichols, L. M., Gesteland, R. F., & Atkins, J. F. (1991). The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. *Journal of Molecular Biology*, *218*(2), 365–373.

Stiebler, A. C., Freitag, J., Schink, K. O., Stehlik, T., Tillmann, B. A. M., Ast, J., & Bilker, M. (2014). Ribosomal readthrough at a short UGA stop codon context triggers dual localization of metabolic enzymes in fungi and animals. *PLoS Genetics*, *10*(10), 1–9.

Stone, E. A., & Sidow, A. (2005). Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. *Genome Research*, *15*(7), 978–986.

Takeda, J., Kayano, T., Fukomoto, H., & Bell, G. I. (1993). Organization of the human *GLUT2* (pancreatic beta-cell and hepatocyte) glucose transporter gene. *Diabetes*, *42*(5), 773–777.

Tan, I. K., Gajra, B., & Lim, M. S. F. (2006). Study of inherited metabolic disorders in Singapore - 13 years experience. *Annals of the Academy of Medicine Singapore*, *35*(11), 804–813.

Tang, H., & Thomas, P. D. (2016). Tools for predicting the functional impact of nonsynonymous genetic variation. *Genetics*, *203*(2), 635–647.

Taylor, R., Magnusson, I., Rothman, D. L., Cline, G. W., Caumo, A, Cobelli, C., & Shulman, G. I. (1996). Direct assessment of liver glycogen storage by 13C nuclear magnetic resonance spectroscopy and regulation of glucose homeostasis after a mixed meal in normal subjects. *The Journal of Clinical Investigation*, *97*(1), 126–132.

Thoma, N. H., & Leadlay, P. E. (1996). Homology modeling of human methylmalonyl-CoA mutase: A structural basis for point mutations causing methylmalonic aciduria. *Protein Science, 5*(9), 1922–1927.

Tozzo, P., Caenazzo, L., & Parker, M. J. (2014). Discovering misattributed paternity in genetic counselling: Different ethical perspectives in two countries. *Journal of Medical Ethics*, *40*(3), 177–181.

Uotani, S., Yamasaki, H., Takino, H., Kawasaki, E., Matsuo, H., Yamasaki, S., … Eguchi, K. (1995). Identification of a 5' splice junction mutation in the debranching enzyme gene in a Japanese patient with glycogen storage disease type IIIa. *Journal of Inherited Metabolic Disease*, *7*(1990), 861–862.

VanderMeer, J. E., Lozano, R., Sun, M., Xue, Y., Daentl, D., Jabs, E. W., … Ahituv, N. (2014). A novel *ZRS* mutation leads to preaxial polydactyly type 2 in a heterozygous form and Werner mesomelic syndrome in a homozygous form. *Human Molecular Genetics*, *35*(8), 945–948.

Veiga-Da-Cunha, M., Gerin, I., Chen, Y.-T., De Barsy, T., De Lonlay, P., Dionisi-Vici, C., … Schaftingen, E. Van. (1998). A gene on chromosome 11q23 coding for a putative glucose-6-phosphate translocase is mutated in glycogen-storage disease types Ib and Ic. *American Journal of Human Genetics*, *63*, 976–983.

Veiga-Da-Cunha, M., Gerin, I., Chen, Y.-T., Lee, P. J., Leonard, J. V, Ene Maire, I., … Schaftingen, E. Van. (1999). The putative glucose 6-phosphate translocase gene is mutated in essentially all cases of glycogen storage disease type I non-a. *European Journal of Human Genetics*, *7*, 717–723.

Veltman, J. A., & Brunner, H. G. (2012). De novo mutations in human genetic disease. *Nature Reviews Genetics, 13*(8), 565-575.

Visser, G., Rake, J.-P., Fernandes, J., Labrune, P., Leonard, J. V., Moses, S., … Smit, G. P. A. (2000). Neutropenia, neutrophil dysfunction, and inflammatory bowel disease in glycogen storage disease type 1b: Results of the European study on glycogen storage disease type I. *The Journal of Pediatrics*, *137*, 187–191.

Visser, G., Rake, J. P., Labrune, P., Leonard, J. V, Moses ', S., Ullrich, K., … Smit, G. P. A. (2002). Granulocyte colony-stimulating factor in glycogen storage disease type lb. Results of the European study on glycogen storage disease type 1. *European Journal of Pediatrics*, *161*, 83–87.

Waldo, G. S., Standish, B. M., Berendzen, J., & Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. *Nature Biotechnology*, *17*(7), 691–695.

Wang, F., Han, L. S., Ye, J., Qiu, W. J., Zhang, Y. F., Gao, X. L., … Gu, X. F. (2009). Analysis of the *MUT* gene mutations in patients with methylmalonic acidemia. *Chinese Journal of Medical Genetics*, *26*(5), 485–489.

Wang, J., Li, E., Wang, L., Wang, Z., Yang, S., Zhou, Q., & Chen, Q. (2015). Genetic analysis of four cases of methylmalonic aciduria and homocystinuria, cblC type. *International Journal of Clinical and Experimental Pathology*, *8*(8), 9337–9341.

Wang, L.-L., Li, Y., & Zhou, S.-F. (2009). A bioinformatics approach for the phenotype prediction of nonsynonymous single nucleotide polymorphisms in human cytochromes P450. *Drug Metabolism and Disposition*, *37*(5), 977–991.

Wang, Z., & Moult, J. (2001). SNPs, protein structure, and disease. *Human Mutation*, *17*(4), 263–270.

Wang, Han, L., Yang, Y., Gu, X., Ye, J., Qiu, W., … Wang, Y. (2010). Clinical, biochemical, and molecular analysis of combined methylmalonic acidemia and hyperhomocysteinemia (cblC type) in China. *Journal of Inherited Metabolic Disease*, *33*(3), 435-442.

Weber-Lehmann, J., Schilling, E., Gradl, G., Richter, D. C., Wiehler, J., & Rolf, B. (2014). Finding the needle in the haystack: Differentiating "identical" twins in paternity testing and forensics by ultra-deep next generation sequencing. *Forensic Science International: Genetics*, *9*, 42–46.

Weisfeld-Adams, J. D., Bender, H. A., Miley-Åkerstedt, A., Frempong, T., Schrager, N. L., Patel, K., … Diaz, G. A. (2013). Neurologic and neurodevelopmental phenotypes in young children with early-treated combined methylmalonic acidemia and homocystinuria, cobalamin C type. *Molecular Genetics and Metabolism*, *110*(3), 241–247.

Weisfeld-Adams, J. D., Morrissey, M. A., Kirmse, B. M., Yu, C., P, W. M., Salveson, B. R., … Sunny, S. (2010). Newborn screening and early biochemical follow-up in combined methylmalonic aciduria and homocystinuria. *Molecular Genetics and Metabolism*, *99*(2), 116–123.

Willard, H. F., Mellman, I. S., & Rosenberg, L. E. (1978). Genetic complementation among inherited deficiencies of methylmalonyl-CoA mutase activity: Evidence for a new class of human cobalamin mutant. *American Journal of Human Genetics*, *30*(1), 1–13.

Worgan, L. C., Niles, K., Tirone, J. C., Hofmann, A., Verner, A., Sammak, A., … Rosenblatt, D. S. (2006). Spectrum of mutations in mut methylmalonic acidemia and identification of a common Hispanic mutation and haplotype. *Human Mutation*, *27*(1), 31–34.

Yang, B. Z., Ding, J. H., Enghild, J. J., Bao, Y., & Chen, Y.-T. (1992). Molecular cloning and nucleotide sequence of cDNA encoding human muscle glycogen debranching enzyme. *Journal of Biological Chemistry*, *267*(13), 9294–9299.

Yie, J., Hecht, R., Patel, J., Stevens, J., Wang, W., Hawkins, N., … Xu, J. (2009). FGF21 N- and C-termini play different roles in receptor interaction and activation. *FEBS Letters*, *583*(1), 19–24.

Youssoufian, H., & Pyeritz, R. E. (2002). Mechanisms and consequences of somatic mosaicism in humans. *Nature Reviews Genetics*, *3*(10), 748–758.

Yunus, Z. M., Rahman, S. A., Choy, Y. S., Keng, W. T., & Ngu, L. H. (2016). Pilot study of newborn screening of inborn error of metabolism using tandem mass spectrometry in Malaysia: Outcome and challenges. *Journal of Pediatric Endocrinology and Metabolism*, *29*(9), 1031-1039.

Zhai, L., Feng, L., Xia, L., Yin, H., & Xiang, S. (2016). Crystal structure of glycogen debranching enzyme and insights into its catalysis and disease-causing mutations. *Nature Communications*, *7*, 11229.

Zhao, F.-Q., & Keating, A. F. (2007). Functional properties and genomics of glucose transporters. *Current Genomics*, *8*(2), 113–128.

Zhou, S.-F., Liu, J.-P., & Chowbay, B. (2009). Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metabolism Reviews*, *41*(2), 89–295.

Zong, Y., Liu, N., Zhao, Z., & Kong, X. (2015). Prenatal diagnosis using genetic sequencing and identification of a novel mutation in *MMACHC*. *BMC Medical Genetics*, *16*, 48.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

1) Publication

Yap, S., Leong, H. Y., Abdul Aziz, F., Hassim, H., Sthaneshwar, P., Teh, S. H., … Mohamed, Z. (2016). N-carbamylglutamate is an effective treatment for acute neonatal hyperammonaemia in a patient with methylmalonic aciduria. *Neonatology*, *109*(4), 303-307.

Yap, S., **Abdullah, I. S**., Ngu, L. H., Teh, S. H., Ong, S. Y., Chong, S. Y., … Mohamed, Z. (2012). Molecular spectrum of glycogen storage disease type III in Malaysian patients. *Journal of Inherited Metabolic Disease*, *35*(1), 82.

2) Paper presentation

Abdullah, I. S. (2015, November). *Functional analysis of six novel exonic and splicing variations of the AGL gene as potential causes of glycogen storage disease type III.* Paper presented at the meeting of the 2nd IAMS-UM Research Collaboration Meeting & Worksop, Kuala Lumpur, Malaysia.

Abdullah, I. S. (2012, December). *Novel mutations of glycogen storage disease type III (GSD III) in Malaysian patients.* Paper presented at the meeting of the 17th Biological Sciences Graduate Congress (BSGC), Bangkok, Thailand.

Abdullah, I. S. (2012, October). *Investigating the mutation spectrum of glycogen storage disease type III in Malaysian population.* Paper presented at the meeting of the 19th Malaysian Society of Molecular Biology and Biotechnology (MSMBB), Kuala Lumpur, Malaysia.

Novel Insights from Clinical Practice

Neonatology

Neonatology 2016;109:303–307
DOI: 10.1159/000443630

Received: October 22, 2015
Accepted after revision: December 23, 2015
Published online: February 24, 2016

N-Carbamylglutamate Is an Effective Treatment for Acute Neonatal Hyperammonaemia in a Patient with Methylmalonic Aciduria

Sufin Yap[a, b] Huey Yin Leong[a, c] Fadzlina Abdul Aziz[b] Haszlin Hassim[d] Pavai Sthaneshwar[e] Ser Huy Teh[f] Ili Syazwana Abdullah[f] Lock Hock Ngu[c] Zulqarnain Mohamed[f]

[a]Department of Inherited Metabolic Diseases, Sheffield Children's Hospital, NHS Foundation Trust, Sheffield, UK; [b]Department of Paediatrics, Faculty of Medicine and University of Malaya Medical Center, University of Malaya, [c]Department of Clinical Genetics, Hospital Kuala Lumpur, [d]Department of Dietetics, University of Malaya Medical Center, [e]Department of Laboratory Medicine, Faculty of Medicine, University of Malaya, and [f]Genetics and Molecular Biology Division, Institute of Biological Sciences, University of Malaya, Kuala Lumpur, Malaysia

Established Facts

- Hyperammonaemia is a medical emergency that requires immediate treatment to minimise neurocognitive damage. Secondary hyperammonaemia is a common presenting feature of methylmalonic aciduria.
- Oral N-carbamylglutamate has recently been used as a multiple treatment strategy with conventional ammonia scavengers (sodium benzoate, sodium phenylbutyrate) and dialysis, in conjunction with dietary manipulation to treat acute neonatal hyperammonaemia.

Novel Insights

- N-carbamylglutamate was successfully used as the sole ammonia-lowering medication, together with aggressive nutritional management, in a neonate presenting with severe hyperammonaemia secondary to methylmalonic aciduria.
- The rapid fall in ammonia levels allowed for early recommencement of feeds after minimal duration of catabolism and toxicity, and averted all possible complications associated with dialysis.

Key Words
N-carbamylglutamate · Neonatal hyperammonaemia · Methylmalonic aciduria · Rapid therapeutic response

Abstract
N-carbamylglutamate (NCG) has been used in combination with ammonia scavengers (sodium benzoate, sodium phenylbutyrate) and dialysis to treat hyperammonaemia in methylmalonic aciduria (MMA). The sole use of NCG for acute neonatal hyperammonaemia secondary to MMA is demonstrated in a neonate presenting at day 9 with encephalopathy, severe metabolic acidosis, hyperammonaemia (1,089 μmol/l), ketonuria and urinary methylmalonic acids. Emergency treatment included discontinuing protein feeds, providing high calories, carnitine and hydroxocobalamin. NCG 200 mg given at 0 and 90 min decreased plasma ammonia dramatically from 1,089 to 567 μmol/l at 90 min and further

KARGER

© 2016 S. Karger AG, Basel
1661–7800/16/1094–0303$39.50/0

E-Mail karger@karger.com
www.karger.com/neo

Prof. Sufin Yap
Department of Inherited Metabolic Diseases
Sheffield Children's Hospital, NHS Foundation Trust
Western Bank, Sheffield S10 2TH (UK)
E-Mail Sufin.Yap@sch.nhs.uk

P-194

MULTIPLEX LIGATION-DEPENDENT PROBE AMPLIFICATION (MLPA) ANALYSIS IN HEREDITARY FRUCTOSE INTOLERANCE (HFI)

Santer L[1] , Steglich C[1] , Bergmann J[2] , Tsiakas K[2] , Schneppenheim R[3] , Santer R[2]
[1]*Inst Hum Genet, Univ Med Cent Eppendorf, Hamburg, Germany*
[2]*Dept Pediatr, Univ Med Cent Eppendorf, Hamburg, Germany*
[3]*Dept Pediatr Hematol Oncol, Univ Med Cen, Hamburg, Germany*

Congenital defects of ALDOB are the molecular basis of HFI. While certain point mutations (p.A150P, p.A175D) are frequent, the role of long-range deletions has remained unclear. Here, we report our experience with MLPA analysis, a method which can detect both known mutations but also copy number variations.
Patients/Methods: A total of 88 HFI patients were enrolled in the study (72 previously published, Hum Mutat 2006). Among them were 7 cases suspicious for homo- or heterozygosity for a long-range deletion. MLPA analysis was performed with the SALSA MLPA P255 ALDOB probemix (MRC Holland) and by separation of exon- and mutation-specific PCR products by capillary electrophoresis. Results were confirmed by conventional Sanger sequencing of junction fragments.
Results: The two common point mutations were reliably detected. In 5/7 cases, long-range deletions were found in exons 1, 6, and 2-6, respectively, and in 4/5 cases breakpoints were determined (del c.540+567_c.651 insATCCGTT [del 1315 ins 7 bp], del -2085_c.624+63 [del 6489 bp], del -1747_c.624+41 [del 6131 bp]). Furthermore, a novel 2-bp-deletion, c.893_4delTA (p.L298HfsX7X) was detectable. Detection rates for **two** at least one ALDOB mutation in our cohort were compared: sequencing **92.0/**97.7, MLPA **75.9/**92.0, MLPA (extended for p.N335K) **83.9/**97.7 %.
Conclusion: MLPA is a reliable first diagnostic step in HFI.

P-195

COULD GSD TYPE I EXPAND THE SPECTRUM OF DISORDERS WITH ELEVATED PLASMA CHITOTRIOSIDASE ACTIVITY?

Turner L[1] , Kasapkara C S[1] , Ilkeroglu G[1] , Hasanoglu A[1]
[1]*Div Ped Metab Dis, Gazi Univ Hosp, Ankara, Turkey*

Glycogen storage disease type I (GSDI) is characterized by accumulation of glycogen and fat in the liver and kidneys, resulting in hepatomegaly and renomegaly. Human chitotriosidase is a recently described fully active chitinase expressed by activated macrophages. Marked elevation of chitotriosidase activity was initially observed in plasma of patients with gaucher disease. Subsequently, elevation was also observed in various lysosomal storage disorders such as fucosidosis, galactosialidosis and glycogen storage disease type IV. The aim of the present study was to evaluate plasma chitotriosidase activity in 19 children with Glycogen storage disease type I. Plasma chitotriosidase levels were found to be significantly higher in children with GSD type I than healthy age-matched controls (21.5±16.4 vs. 12.5±6.9 nmol/h/ml, p=0.04). All the patients reported here presented with hepatomegaly. Irrespective of the mechanisms involved, our report expands the spectrum of disorders that should be included in the differential diagnosis of patients with increased plasma chitotriosidase activity.

P-196

CLINICAL APPLICATION OF NEXT GENERATION SEQUENCING TO MOLECULAR DIAGNOSIS OF GENETICALLY HETEROGENEOUS GLYCOGEN STORAGE DISEASE

Wang J[1] , Cui H[1] , Lee NC[2] , Craigen WJ[1] , Zhang VW[1] , Wong LJ[1]
[1]*Mol Hum Genet, Baylor College Medicine, Houston, Texas, United States*
[2]*Pediatr, national Taiwan Univ Hospital, Taipei, Taiwan*

Background: Glycogen storage diseases (GSDs) are a group of genetic disorders of glycogen metabolism. Clinical features include hypoglycemia, hepatomegaly, developmental delay and muscle cramps. Untreated patients may suffer devastating outcome, therefore, prompt and accurate diagnosis is critical. Here we report the utility of a cost effective next generation sequencing (NGS)-based molecular analysis for GSDs.
Method: The coding exons and their flanking 20 bp intron regions of 16 genes; GYS2, GYS1, G6PC, SLC37A4, GAA, AGL, GBE1, PYGM, PYGL, PFKM, PHKA2, PHKB, PHKG2, PHKA1, PGAM2, and PGM1, are captured using DNA probes, followed by fragmentation, template library preparation, and massively parallel sequencing (MPS).
Results: All bases in the target regions have been unbiasedly enriched with deep coverage. Validation was performed by parallel Sanger sequencing. MPS demonstrated 100 % sensitivity and specificity compared to Sanger sequencing. This method correctly identified all types of mutations, including single nucleotide substitutions, small deletions and duplications, and large deletions involving one or more exons. Furthermore, novel mutations have been identified in previously undiagnosed patients.
Conclusion: This report demonstrated the clinical utility of MPS-based technology in molecular diagnosis of a group of clinically overlapping and genetically heterogeneous disorders such as GSD, in a cost and time efficient manner.

P-197

MOLECULAR SPECTRUM OF GLYCOGEN STORAGE DISEASE TYPE III IN MALAYSIAN PATIENTS

Yap S[1] , Abdullah IS[2] , Ngu LH[3] , Teh SH[2] , Ong SY[3] , Chong SY[4] , Boey CCM[4] , Lee WS[4] , Mohamed Z[2]
[1]*Div IMD, Paeds Dept & UMMC, Univ Malaya, Kuala Lumpur, Malaysia*
[2]*Inst Biol Sc, Univ of Malaya, Kuala Lumpur, Malaysia*
[3]*Dept Genet & Metab, Hosp Kuala Lumpur, Kuala Lumpur, Malaysia*
[4]*Div Hep & GE, Paeds Dept & UMMC, UM, Kuala Lumpur, Malaysia*

Glycogen storage disease type III (GSD III) is a rare disorder of glycogen debranching enzyme (GDE) deficiency which is caused by ethnically dependent mutations in the AGL gene. There is biochemical and clinical heterogeneity reflecting the severity of the AGL mutations. Clinically, it affects the liver, cardiac and skeletal muscles. GSD IIIa is the most common (85 %) and shows both liver and muscle involvement, while GSD IIIb affects the liver only. This study reports on the molecular spectrum of the AGL gene, which is as yet unknown, in an ethnically diverse Malaysian population. Nine patients (7 males; 2 females) from 8 unrelated families were studied; seven Malays and two Chinese. The mean age was 8 years (range: 2.4 -12.5); Mean age of diagnosis is at 2.8 years (range: 1.1-7.0). Six had GDE enzymology confirmation. Of the 16 independent alleles studied, 8 different mutations were found; six novel (c.4490 G>A, c.1423+1 G>T, c.3814_3815del AG, c.2914_2915delAA, c.4333 T>G, c.4531_4534delTGTC) and 2 previously reported (c.2681+1G>A, c.99C>T). The commonest mutation c.2681+1G>A (n=7 alleles; 70 %) was found in five Malay patients.
Conclusion: The Malaysian AGL gene spectrum shows heterogeneity. The novel c.4531_4534delTGTC mutation is associated with a milder phenotype with relatively good response to carbohydrate and protein supplementation.

APPENDIX

Appendix A: List of primers used for mutation screening

***AGL* gene**		
Exons	Primer	Sequences
1-2	1-3F	5'-TCCCCAGGGCAAGGAGAAAG-3'
	2R	5'-GCAATACACTTAGGAAAGCCTGAAG-3'
3	Seq3F	5'-TTCGAACATGTAAGTGCCGC-3'
	3R	5'-GGAGAACACAGCACCATCTTTGC-3'
4a	4aF	5'-AGTGCTGCCTCAACTTCCTG-3'
	4aR	5'-GGTGGGAATACATGGAGTGATTC-3'
4-5	4-5F	5'-TAATTCAGGGGTGAGGATGG-3'
	4-5R	5'-CGTAGCACCTCCCACTTTGT-3'
6	6F	5'-TGAACCCAAGTGTTTGACCTC-3'
	6R	5'-AAACAGAAGGACACCTAATCATCA-3'
7-8	7-8F	5'-AAAGTGCTGGGGTTGAGCTAC-3'
	7-8R	5'-GGAGCCCCTCTACGTCAGTT-3'
9-11	9-11F	5'-TGTTTCTTTCTCCCCCACAC-3'
	9-11R	5'-AGGAGCCAGATGCTTGCTTA-3'
12	12F	5'-AGCCAGATATGCTGTCATGTT-3'
	12R	5'-TTTCCCCATCCACAGAGAAG-3'
13	13F	5'-ATGGCAGAAATGATCAAAGC-3'
	13R	5'-CCAAATCAATGCTTGTGTCC-3'
14	14F	5'-CAAATCATGCCTCCTTTTGTC-3'
	14R	5'-TGAGTCTTCAGTTCATCCTACTGG-3'
15-18	15-18F	5'-TACATGAGCCATTTCTCCAGTTAAG-3'
	15-18R	5'-TTTCTCAAGTACGTCCACAGTAGC-3'
19-21	19-21F	5'-GAAAATGCCAGGGGTAGTGA-3'
	19-21R	5'-CCGTACTTTGAGTAGCAAGG-3'
22	22F	5'-TGAAGGAAGGAAGGAAAACACT-3'
	22R	5'-AGCCATAATTGCAACCCAAG-3'
23-25	23-25F	5'-GCAGAATAGGGACTAGAGGATATAGG-3'
	23-25R	5'-CACACATAAATGCCTGGTGCACAGTT-3'
26	26F	5'-TGCCTTGTACCCCAGGTTTA-3'
	26R	5'-AGCAATGCCTTCACTATGGA-3'
27	27F	5'-GCAAGAGAGAAAACGCATTCA-3'
	27R	5'-CAGCAGGTGCCAAATCAATA-3'

Appendix A, continued.

28	Seq28F	5'-TCACATTACTTCAGTTGTCGG-3'
	E29R	5'-TGCAAGGTTTAATTTCATCACC-3'
29	29F	5'-AACTGAGCTTTAGAGTGGTTGTCC-3'
	29R	5'-AGGCAGGGAAATTTTGATTG-3'
30-31	30F	5'-CATTACAATTGTTTACCGAATGCC-3'
	31R	5'-CAAATGGGAATAACCAACTAAGCA-3'
32-34	32F	5'-GGTAGGAACTAATTCTTCTGTG-3'
	33-34R	5'-AAAGGCTTCACTTTTCATTGGG-3'
35	35F	5'-CACCAGGTCTTGCCTATTTTG-3'
	35R	5'-GCTTTTACAATCTTACCTAATTGAGC-3'
		***SLC37A4* gene**
Exons	Primer	Sequences
3	G6PT-1F	5'-AGGCTGTGCGTCTTGGCTGGTAGGG-3'
	G6PT-1R	5'-TTCGTGTCCCCAGGTCCACCA-3'
5	G6PT-2F	5'-CCTTCTTTCATTGCTCCTGTGTTT-3'
	G6PT-2R	5'-CTCTATGACAATCCAAACAGGCTC-3'
6	G6PT-3F	5'-CTGCCCCATCTGACCCCACCCTCA-3'
	G6PT-3R	5'-AGTGGTCGGTCTGGGTGGGGGCTC-3'
7	G6PT-4F	5'-GGGAGAGCAGTCAGGGCAGAGCCT-3'
	G6PT-4R	5'-CTGCTCCTTATGCCCACCCTTGTC-3'
8	G6PT-5F	5'-TCCCACCACAACTCCCTACTGCAG-3'
	G6PT-5R	5'-CCCTTCTCCTTCCTGTCCCTTCTG-3'
9	G6PT-6F	5'-TGTTCTGAGGACGTGACATTGCCG-3'
	G6PT-6R	5'-CCTTGTGCCCTGCCGTGAGCC-3'
10	G6PT-7F	5'-TCTGGGCCTGGTTTTCTTTTCTTC-3'
	G6PT-7R	5'-GTGAGACAGACCAGGAGAAAAACC-3'
11	G6PT-8F	5'-CTCTGAATGCCACTCCACTCTGCCC-3'
	G6PT-8R	5'-ACAGGTGGGGGTGAGGGAGAGACT-3'
12	G6PT-9F	5'-GCTTAGGTTCTTCCCTTTCCCCCTG-3'
	G6PT-9R	5'-AGAGCGTGCAGGGGGAAGGCCACCG-3'
		***SLC2A2* gene**
Exons	Primer	Sequences
1	SLC2A2-1F	5'-CTTGCTCCTCCTCCTGCAATGC-3'
	SLC2A2-1R	5'-CCCAAATCCTCTGTATGACTTGAC-3'
2	SLC2A2-2F	5'-CTAAAATTGCCTGGTGTTTCCC-3'
	SLC2A2-2R	5'-CGTGGACACCCTTTATCTCTG-3'

Appendix A, continued.

3	SLC2A2-3F	5’-CCAACTGTAAACAAATCTATCC-3’
	SLC2A2-3R	5’-CAACTCTAAAGCTATTCCACAA-3’
4a	SLC2A2-4aF2	5’-ATTCCCCAACTAGCTTTTAC-3’
	SLC2A2-4aR	5’-AAAATATCCCTGAGTGCTAC-3’
4b	SLC2A2-4bF	5’-TTATTGAGATAGTCCTGGTTGC-3’
	SLC2A2-4bR	5’-CAGTAGGGGATGCAATAGTAG-3’
5-6	SLC2A2-5F	5’-AACTAGTGTAAAAGGTAGATCC-3’
	SLC2A2-6R2	5’-AGACCCATGATGCCAATACC-3’
7	SLC2A2-7F	5’-TCTGTACGAAGTTCTAGGG-3’
	SLC2A2-7R	5’-TAGAGCATGTGTACAATAAGTC-3’
8	SLC2A2-8F	5’-CTCCCCACCTTGATCTCACTCC-3’
	SLC2A2-8R	5’-GCACTTTAACCTGGACCACAGAGG-3’
9-10	SLC2A2-9F	5’-AGCAGCCAAATGATCTCAGTCC-3’
	SLC2A2-10R	5’-GGATATAAAATGCTCAAGGAATCATC-3’

***MUT* gene**		
Exons	Primer	Sequences
1	Mut-1F	5’-CTCACCACTGATTCTTCTTG-3’
	Mut-1R	5’-CCTAAGGAGAAGGCAGAG-3’
2-3	Mut-2F	5’-TATGAGTAGCTCCTATTTCCC-3’
	Mut-3R	5’-GCAAGTAACGACAGAAC-3’
4	Mut-4F	5’-CAGTCCTGATGATGGTTC-3’
	Mut-4R	5’-GCTGGAGACAAGATATTCC-3’
5	Mut-5F	5’-AGGGCATAGGAAGAGATTC-3’
	Mut-5R	5’-CTGCTTGTGCCACATTG-3’
6	Mut-6F	5’-CACTGAACTCTGACTCTTC-3’
	Mut-6R	5’-TGCTGTTCTTTGTATGAGC-3’
7-8	Mut-7F	5’-CCCAAGACTTAAGAGGTT-3’
	Mut-8R	5’-CAAGTTTCTCAATGCCTTATC-3’
9	Mut-9F	5’-GGTCTAATCTCTTGATCTCTG-3’
	Mut-9R	5’-ATGGTTTACAGGATCAACC-3’
10-11	Mut-10F	5’-AGAATTGGATGCATAAAGGC-3’
	Mut-11R	5’-TACCAGTTACCAGGAGATG-3’
12	Mut-12F	5’-TTGCCCATTAGTATGTTCTG-3’
	Mut-12R	5’-ACACTGTCCACTTTTAGAC-3’
13	Mut-13F	5’-TAGCACTCCCAATCAAGG-3’
	Mut-13R	5’-CATACTTATAGCATGACACCAG-3’

Appendix A, continued.

***MMACHC* gene**		
Exon	Primer	Sequence
1	MMACHC_1F	5'- GGGATACCGTGATGATACGC-3'
	MMACHC_1R	5'- GAACCCAGGAGGATCAGAGG-3'
2	MMACHC_2F	5'- TGCATCACATAGCGTCAGTG-3'
	MMACHC_2R	5'- AGCCTGGCTTTAGGGTATCA-3'
3	MMACHC_3F	5'- TCATGTTTTCCCTTCTGAGGA-3'
	MMACHC_3R	5'- CAAAGCTAATTTGTTCTGGGTTG-3'
4	MMACHC_4F	5'- AGGCCTAGCTTGCAATGATG-3'
	MMACHC_4R	5'- GAAGGCAGATGGGAATTCTG-3'

Appendix B: List of primers used for DNA sequencing of the *AGL* gene

Fragment	Exon	Forward primer	Reverse primer
1	1	1-3F 5'-TCCCCAGGGCAAGGAGAAAG-3'	2R 5'-GCAATACACTTAGGAAAGCCTGAAG-3'
	2		
2	3	Seq3F 5'-TTCGAACATGTAAGTGCCGC-3'	3R 5'-GGAGAACACAGCACCATCTTTGC-3'
3	4a		4aR 5'-GGTGGGAATACATGGAGTGATTC-3'
4	4	4-5F 5'-TAATTCAGGGGTGAGGATGG-3'	Seq4R 5'-TACAGTAACATTTACAAGTTAGG-3'
	5		4-5R 5'-CGTAGCACCTCCCACTTTGT-3'
5	6	6F 5'-TGAACCCAAGTGTTTGACCTC-3'	6R 5'- AAACAGAAGGACACCTAATCATCA-3'
6	7	E8F 5'- GCTGAAGCGAATGATAGGATACC-3'	Seq7R 5'-ACAGGTTCTAAGTAATTTTCAAC-3'
	8	Seq8F 5'-CCTGTATTTTAATATGATAAACAG-3'	7-8R 5'- GGAGCCCCTCTACGTCAGTT-3'
7	9	9-11F 5'- TGTTTCTTTCTCCCCCACAC-3'	9-11R 5'- AGGAGCCAGATGCTTGCTTA-3'
	10		
	11		
8	12	12F 5'- AGCCAGATATGCTGTCATGTT-3'	12R 5'- TTTCCCCATCCACAGAGAAG-3'
9	13	13F 5'- ATGGCAGAAATGATCAAAGC-3'	13R 5' CCAAATCAATGCTTGTGTCC-3'
10	14	14F 5'- CAAATCATGCCTCCTTTTGTC-3'	14R 5'- TGAGTCTTCAGTTCATCCTACTGG-3'
11	15	15-18F 5'- TACATGAGCCATTTCTCCAGTTAAG-3'	
	16		15-18R 5'- TTTCTCAAGTACGTCCACAGTAGC-3'
	17	Seq16F 5'- ATGCTATAGAATAGCACTTTGC3'	
	18		

Appendix B, continued.

12	19	19-21F 5'- GAAAATGCCAGGGGTAGTGA-3'	Seq19R 5'- TCCAGCTTGTTTAACAATTTTAC-3'
	20		19-21R 5'- CCGTACTTTGAGTAGCAAGG-3'
	21		
13	22	22F 5'- TGAAGGAAGGAAGGAAAACACT-3'	22R 5'- AGCCATAATTGCAACCCAAG-3'
14	23	23-25F 5'- GCAGAATAGGGACTAGAGGATATAGG-3'	Seq24R 5'-AAGAACATGTATGAAGTGACTC-3'
	24		
	25		23-25R 5'- CACACATAAATGCCTGGTGCACAGTT-3'
15	26	26F 5'- TGCCTTGTACCCCAGGTTTA-3'	26R 5'- AGCAATGCCTTCACTATGGA-3'
16	27	27F 5'- GCAAGAGAGAAAACGCATTCA-3'	
17	28	Seq28 5'- TCACATTACTTCAGTTGTCGG-3'	E29R 5'- TGCAAGGTTTAATTTCATCACC3'
18	29	29F 5'- AACTGAGCTTTAGAGTGGTTGTCC-3'	29R 5'- AGGCAGGGAAATTTTGATTG-3'
19	30	30F 5'- CATTACAATTGTTTACCGAATGCC-3'	E30R
	31	E32F 5'-CAGACTGGCCACATTTTCCAA-3'	31R 5'- CAAATGGGAATAACCAACTAAGCA-3'
20	32	32F 5'- GGTAGGAACTAATTCTTCTGTG-3'	E33R 5'-TGATGTCAATGTATTTGGATT-3'
	33	E34F 5'-TGTACTAATGCCGAGCTTATTCTG-3'	33-34R 5'- AAAGGCTTCACTTTTCATTGGG-3'
	34		
21	35	35F 5'- CACCAGGTCTTGCCTATTTTG-3'	35R 5'- GCTTTTACAATCTTACCTAATTGAGC-3'

Appendix C: Chemicals and reagents for Western blotting

Transfer buffer

3.03 g Tris base, 14.4 g glycine and 200 mL of 100 % methanol

* Top up to 1000 mL with dH_2O

Blocking buffer
2.5 g skim milk (non-fat or low-fat) in 50 mL of 1 X TBS

5 X Tris-buffered saline (TBS)
12.11 g Tris base and 48.85 g NaCl
* Top up to 1000 mL with dH_2O, pH adjusted to 7.5 with NaOH or HCl

Probing buffer
4 mL of Blocking buffer and 6 mL of 1 X TBS

Washing buffer
1 mL of Tween-20 in 500 mL of 1 X TBS

TBST
1 X TBS and 0.2 % Tween-20

Appendix D: *In silico* result for PCR Primer Stats from Sequence Manipulation Suite software showing the properties of selected *AGL* gene primers.

```
------------------------------------------------------------
General properties:
-------------------
                    Primer name: 1-3F
                Primer sequence: TCCCCAGGGCAAGGAGAAAG
                Sequence length: 20
                    Base counts: G=7; A=7; T=1; C=5; Other=0;
                 GC content (%): 60.00
     Molecular weight (Daltons): 6185.08
                       nmol/A260: 4.91
                micrograms/A260: 30.38
            Basic Tm (degrees C): 56
   Salt adjusted Tm (degrees C): 51
Nearest neighbor Tm (degrees C): 66.68

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                        Length: Pass
                    Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                      GC clamp: Pass
                 Self-annealing: Pass
             Hairpin formation: Pass
------------------------------------------------------------

------------------------------------------------------------
General properties:
-------------------
                    Primer name: 2R
                Primer sequence: GCAATACACTTAGGAAAGCCTGAAG
                Sequence length: 25
                    Base counts: G=6; A=10; T=4; C=5; Other=0;
                 GC content (%): 44.00
     Molecular weight (Daltons): 7708.10
                       nmol/A260: 3.92
                micrograms/A260: 30.20
            Basic Tm (degrees C): 56
   Salt adjusted Tm (degrees C): 51
Nearest neighbor Tm (degrees C): 64.42

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                        Length: Pass
                    Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                      GC clamp: Pass
                 Self-annealing: Pass
             Hairpin formation: Pass
------------------------------------------------------------
```

Appendix E: *In silico* result for PCR Primer Stats from Sequence Manipulation Suite software showing the properties of selected *SLC37A4* gene primers.

```
------------------------------------------------------------
General properties:
-------------------
                    Primer name: G6PT-2F
                Primer sequence: CCTTCTTTCATTGCTCCTGTGTTT
                Sequence length: 24
                    Base counts: G=3; A=1; T=13; C=7; Other=0;
                 GC content (%): 41.67
     Molecular weight (Daltons): 7217.74
                      nmol/A260: 5.02
                micrograms/A260: 36.22
           Basic Tm (degrees C): 54
   Salt adjusted Tm (degrees C): 49
Nearest neighbor Tm (degrees C): 64.60

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                        Length: Pass
                    Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                      GC clamp: Pass
                 Self-annealing: Pass
              Hairpin formation: Pass
------------------------------------------------------------

------------------------------------------------------------
General properties:
-------------------
                    Primer name: G6PT-2R
                Primer sequence: CTCTATGACAATCCAAACAGGCTC
                Sequence length: 24
                    Base counts: G=3; A=8; T=5; C=8; Other=0;
                 GC content (%): 45.83
     Molecular weight (Daltons): 7265.79
                      nmol/A260: 4.37
                micrograms/A260: 31.76
           Basic Tm (degrees C): 56
   Salt adjusted Tm (degrees C): 51
Nearest neighbor Tm (degrees C): 64.02

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                        Length: Pass
                    Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                      GC clamp: Warning: There are more than 3 G's or
C's in the last 5 bases;
                 Self-annealing: Pass
              Hairpin formation: Pass
------------------------------------------------------------
```

Appendix F: *In silico* result for PCR Primer Stats from Sequence Manipulation Suite software showing the properties of selected *SLC2A2* gene primers.

```
------------------------------------------------------------
General properties:
-------------------
                    Primer name: SLC2A2-1F
                Primer sequence: CTTGCTCCTCCTCCTGCAATGC
                Sequence length: 22
                    Base counts: G=3; A=2; T=7; C=10; Other=0;
                 GC content (%): 59.09
     Molecular weight (Daltons): 6573.29
                      nmol/A260: 5.52
                micrograms/A260: 36.32
            Basic Tm (degrees C): 59
   Salt adjusted Tm (degrees C): 53
Nearest neighbor Tm (degrees C): 67.97

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                        Length: Pass
                    Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                     GC clamp: Pass
                Self-annealing: Pass
            Hairpin formation: Pass
------------------------------------------------------------

------------------------------------------------------------
General properties:
-------------------
                    Primer name: SLC2A2-1R
                Primer sequence: CCCAAATCCTCTGTATGACTTGAC
                Sequence length: 24
                    Base counts: G=3; A=6; T=7; C=8; Other=0;
                 GC content (%): 45.83
     Molecular weight (Daltons): 7247.77
                      nmol/A260: 4.52
                micrograms/A260: 32.77
            Basic Tm (degrees C): 56
   Salt adjusted Tm (degrees C): 51
Nearest neighbor Tm (degrees C): 64.00

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                        Length: Pass
                    Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                     GC clamp: Pass
                Self-annealing: Pass
            Hairpin formation: Pass
------------------------------------------------------------
```

Appendix G: *In silico* result for PCR Primer Stats from Sequence Manipulation Suite software showing the properties of selected *MUT* gene primers.

```
------------------------------------------------------------
General properties:
-------------------
                    Primer name: Mut-1F
                Primer sequence: CTCACCACTGATTCTTCTTG
                Sequence length: 20
                    Base counts: G=2; A=3; T=8; C=7; Other=0;
                 GC content (%): 45.00
     Molecular weight (Daltons): 5993.95
                       nmol/A260: 5.76
                micrograms/A260: 34.55
           Basic Tm (degrees C): 50
   Salt adjusted Tm (degrees C): 45
Nearest neighbor Tm (degrees C): 58.85

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                         Length: Pass
                     Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                       GC clamp: Pass
                  Self-annealing: Pass
              Hairpin formation: Pass
------------------------------------------------------------

------------------------------------------------------------
General properties:
-------------------
                    Primer name: Mut-1R
                Primer sequence: CCTAAGGAGAAGGCAGAG
                Sequence length: 18
                    Base counts: G=7; A=7; T=1; C=3; Other=0;
                 GC content (%): 55.56
     Molecular weight (Daltons): 5606.72
                       nmol/A260: 5.25
                micrograms/A260: 29.46
           Basic Tm (degrees C): 50
   Salt adjusted Tm (degrees C): 45
Nearest neighbor Tm (degrees C): 58.53

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                         Length: Pass
                     Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                       GC clamp: Pass
                  Self-annealing: Pass
              Hairpin formation: Pass
------------------------------------------------------------
```

Appendix H: *In silico* result for PCR Primer Stats from Sequence Manipulation Suite software showing the properties of selected *MUT* gene primers.

```
------------------------------------------------------------
General properties:
-------------------
                    Primer name: MMACHC_1F
                Primer sequence: GGGATACCGTGATGATACGC
                Sequence length: 20
                    Base counts: G=7; A=5; T=4; C=4; Other=0;
                 GC content (%): 55.00
     Molecular weight (Daltons): 6182.08
                       nmol/A260: 5.01
                micrograms/A260: 30.97
           Basic Tm (degrees C): 54
   Salt adjusted Tm (degrees C): 49
Nearest neighbor Tm (degrees C): 62.91

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                        Length: Pass
                    Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                      GC clamp: Pass
                 Self-annealing: Pass
              Hairpin formation: Pass
------------------------------------------------------------

------------------------------------------------------------
General properties:
-------------------
                    Primer name: MMACHC_1R
                Primer sequence: GAACCCAGGAGGATCAGAGG
                Sequence length: 20
                    Base counts: G=8; A=7; T=1; C=4; Other=0;
                 GC content (%): 60.00
     Molecular weight (Daltons): 6225.11
                       nmol/A260: 4.77
                micrograms/A260: 29.69
           Basic Tm (degrees C): 56
   Salt adjusted Tm (degrees C): 51
Nearest neighbor Tm (degrees C): 64.28

PCR suitability tests (Pass / Warning):
-----------------------------------
               Single base runs: Pass
         Dinucleotide base runs: Pass
                        Length: Pass
                    Percent GC: Pass
          Tm (Nearest neighbor): Warning: Tm is greater than 58;
                      GC clamp: Pass
                 Self-annealing: Pass
              Hairpin formation: Pass
------------------------------------------------------------
```