DIFFERENTIAL GENE EXPRESSION ANALYSES IN HBE/BETA THALASSAEMIA PATIENTS AND THEIR RELATIONSHIP TO DISEASE SEVERITY

HEBA A. A. ALSALEH

UNIVERSITI SAINS MALAYSIA

2022

DIFFERENTIAL GENE EXPRESSION ANALYSES IN HBE/BETA THALASSAEMIA PATIENTS AND THEIR RELATIONSHIP TO DISEASE SEVERITY

by

HEBA A. A. ALSALEH

Thesis submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

August 2022

ACKNOWLEDGEMENT

First of all, praise to the Almighty Allah for giving me the ability to learn and the strength to complete this research.

My first and deepest appreciation goes to Professor Dr Rosline Hassan for her supervision, advice and trust throughout my study who shared her precious time to give me the insights to move forward in the appropriate direction. She has been invaluable in both academic and personal aspects. Special thanks to my co-supervisor, Associate Professor Dr Sarina Sulong for her advice and guidance.

I would like to express my gratitude to the wonderful staff at the Haematology and the Central Research Laboratories, especially Pn. Abdah Karimah Che Md. Nor, Pn. Noor Suhana Azila Mat Yusoff and Mdm. Ang Cheng Yong for their great advice and help regarding my laboratory work.

I would like to thank my beloved family; my parents Sadiyah and Ali for their endless love, trust and support throughout my life, without them, I would not be able to pursue my PhD, Adel my dear brother who was always there when I needed him, and finally, I would like to thank my gorgeous daughter Nareen for her hugs, support and endless love that give me the power to finish my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AHSCT	Autologous Haematopoietic Stem Cell Transplantation
AHSP	Alpha haemoglobin stabilizing protein
AKT	Protein kinase B
APC	Allophycocyanin
ARMS	Amplification Refractory Mutation System
ASPP	Apoptosis-stimulating of p53 protein
BAD	BCL2-associated agonist of cell death
BAG1	BCL2-associated athanogene
BAG3	BCL2-associated athanogene 3
BAK1	BCL2-antagonist/killer 1
BAX	BCL2-associated X protein
BCL	B-cell lymphoma
BCL-XL	B-cell lymphoma-extra large
BFU-E	Burst forming unit-erythroid
BIM	Bcl-2-like protein 11
BM	Bone marrow

BMD	Bone marrow disease
BMM	Microenvironment of the bone marrow
bp	Base pair
BSA	Bovine serum albumin
CAE	Cellulose acetate electrophoresis
CD	Cluster of differentiation
cDNA	Complementary DNA
CE	Capillary Electrophoresis
CFU-E	Colony forming unit-erythroid
CGA	Glycoprotein Hormones, Alpha Polypeptide
CRADD	Caspase-2 and RIPK1 domain-containing adaptor with death domain
G-CSF	Granulocyte colony-stimulating factor
DEGs	Differentially expressed genes
DEPC	Diethylpyrocarbonate
DFO	Desferoxamine
DFP	Deferiprone
DFX	Deferasirox

DIABLO	Direct IAP-binding protein with a low pi
DISC	Death-inducing signalling complex
DNA	Deoxyribonucleic Acid
EMH	Extramedullary haematopoiesis
EPC	Erythroid progenitor cells
EPO	Erythropoietin
ESRF	End stage renal failure
ET	Essential thrombocythemia
FBC	Full blood count
FC	Fold change
FDR	False discovery rate
FISH	Fluorescence *in situ* hybridisation
FITC	Fluorescein isothiocyanate
FSH	Follicular stimulating hormone
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GATA1	GATA Binding Protein 1
GDC	Genomic DNA contamination

GO	Gene Ontology
GPA	Glycophorin A
Hb	Haemoglobin
HBA	Haemoglobin A
HBA2	Haemoglobin A2
HBB	β-globin gene
HBC	Haemoglobin C
HbCS	Haemoglobin Constant Spring
HBE	Haemoglobin E
HbF	Foetal Haemoglobin
HCT	Haematocrit
HIV	Human immunodeficiency virus
HPLC	High-pressure liquid chromatography
HSC	Haematopoietic stem cell
HUSM	Hospital USM
IAP	Inhibitor of apoptosis protein
ICT	Iron chelation therapy

IDA	Iron deficiency anaemia
IE	Ineffective erythropoiesis
IGF1	Insulin-like growth factor 1
IGF1R	Insulin-like growth factor1 receptor
IGF2	insulin-like growth factor 2
IL	Interleukin
ISH	*In situ* hybridisation
JAK-STAT	Janus kinase/signal transducers and activators of transcription
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LCR	Locus control region
LH	Luteinizing hormone
LPO	lipid peroxidation
MAFB	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog B
MARMS	Multiplex amplification refractory mutation system
MCH	Mean corpuscular haemoglobin
MCHC	Mean corpuscular haemoglobin
MCL1	Myeloid cell leukemia-1

MCV	Mean Cell Volume
MDS	Myelodysplastic syndromes
MLPA	multiplex ligation-dependent probe amplification assay
MOMP	Mitochondrial outer membrane permeabilization
MPN	Myeloproliferative disease
mRNA	Messenger RNA
NAIP	NLR Family Apoptosis Inhibitory Protein
NTDT	Non-transfusion dependent
OH^-	Hydroxide
OPG	Osteoprotegerin
OPN	Osteopontin
PBMCs	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PCA	Principal Component Analysis
PCR	Polymerase chain reaction
PE	Phycoerythrin
PI	Propidium iodide

PI3K	phosphatidylinositol 3’-kinase
PIDD	P53-Induced Death Domain Protein
PTH	Parathyroid hormone
PTHLH	Parathyroid hormone like hormone
PUMA	p53 upregulated modulator of apoptosis
PV	Polycythaemia vera
QC	Quality control
QPCR	Quantitative PCR
RANK	Receptor activator of nuclear factor κ B
RANKL	Receptor activator of nuclear factor κ B ligand
RBC	Red blood cell
RDW	Red cell distribution width
RIN	RNA integrity number
RNA	Ribonucleic acid
ROS	Reactive Oxygen species
RPM	Revolutions per minute
RT-PCR	Reverse transcription PCR

Real-Time qRT-PCR	Real-Time Quantitative Reverse Transcription PCR
SCF	Stem cell factor
SEA	Southeast Asia
SFEM-II	Serum-free expansion medium-II
SNP	Single nucleotide pleomorphism
SPSS	Statistical Package for the Social Sciences
SPP1	secreted Phosphoprotein 1
STAT	Signal transducer and activator of transcription
TBD	Thalassaemia bone disease
TDT	Transfusion-dependent
TGF-β	Transforming growth factor beta
TI	Thalassaemia intermedia
TM	Thalassaemia major
TNF	Tumour necrosis factor
TSH	Thyroid stimulation hormone
TYROBP	TYRO protein tyrosine kinase-binding protein
UK	United Kingdom

USA	United State of America
USM	Universiti Sains Malaysia
UTR	Untranslated region
UV	Ultraviolet
VAD	Carbobenzoxy-valyl-alanyl-aspartyl-[O-methyl]- fluoromethylketone
VDR	Vitamin D receptor
WHO	World Health Organization

LIST OF SYMBOLS

α	Alpha
β	Beta
^{0}C	Celsius
Δ	Delta
€	Epsilon
γ	Gamma
g	Gram
g/dl	Gram per decilitre
mg/m^2	Milligram per square metre
mL	Millilitre
mm	Millimetre
µg	Microgram
µl	Microlitre
µm	Micron
mg	Milligram
µM	Micromolar

ng	Nanogram
ng/µl	Nanogram per microlitre
nM	Nanometre
%	Percent
U	Unit
ζ	Zeta

LIST OF APPENDICES

ANALISIS PENGEKSPRESAN GEN PEMBEZAAN DALAM HBE/BETA TALASEMIA DAN HUBUNGANNYA DENGAN KETERUKAN PENYAKIT

ABSTRAK

Talasemia haemoglobin E-beta (Hb E/β-thalassaemia) ialah gangguan keturunan penyakit genetik yang paling biasa . Ia meliputi hampir separuh daripada kesemua kes talasemia beta yang teruk di seluruh dunia. Di negeri Kelantan, 50.93% pesakit talasemia mempunyai Hb E/β-talasemia, dan kebanyakan kes biasanya dalam kalangan bangsa Melayu berbanding dengan bangsa Cina dan India. Kepelbagaian klinikal ialah manifestasi yang ketara di kalangan pesakit berkisar dari gejala ringan hingga ke teruk yang mana memerlukan pemindahan darah biasa. Terdapat banyak pengubahsuaian dijumpai yang menjejaskan peranan penyakit. Walaubagaimanapun, sebab yang tepat di sebalik keheretogenan ini tidak difahami sepenuhnya. Penyelidikan ini bertujuan untuk mengkaji perbezaan expresi gen dan kemungkinan peranan mereka menyebabkan penyakit dan komplikasi dalam kedua-dua kumpulan transfusi dependens (TDT) dan transfusi tidak dependens (NTDT) pesakit talasemia HbE/β. Ia telah dijalankan masing-masing dengan bantuan tatasusunan dan microtatasusunan pembukah PCR RT2 yang digunakan dalam kajian ekspresi gen dalam retikulosit dan progenitor erithroid. Tiga kawalan biasa dan sejumlah 20 pesakit telah didaftarkan dalam kajian ini; 10 pesakit (50%) adalah TDT, dan 10 orang (50%) NTDT. Kajian retikulosit menunjukkan peningkatan pengawalseliaan gen *BAX* dan *BAD* dalam pesakit TDT, yang mempunyai peranan dalam induksi apoptosis melalui laluan apoptosis mitokondria. Pengawalseliaan mereka dalam TDT mungkin memainkan peranan dalam apoptosis retikulosit, jangka hayat pendek RBC matang dan eryptosis. Kajian sitometri aliran menunjukkan

apoptosis yang lebih tinggi dalam progenitor erithroid pesakit TDT. Peningkatan apoptosis dalam progenitor erithroid dan pengawalseliaan *BAD* dan *BAX* retikulosit dalam TDT mungkin dikaitkan dengan pengawalseliaan turun gen yang terlibat dalam laluan PI3k / AKT bagi kumpulan pesakit yang sama. Analisis laluan dan ontologi menunjukkan penglibatan osteoporosis dan faktor pengawalseliaan tulang berkait dengan laluan VDR dan pembezaan peraturan negatif osteoklas dalam kumpulan TDT. Gen yang terlibat boleh menjadi sasaran terapeutik adalah seperti *SPP1* dan *MAFB*. Pengaktifan mereka bertindak mengurangkan beban penyakit dengan mengurangkan anemia dan mengurangkan komplikasi sumsum tulang. Secara ringkasnya, kajian kami menunjukkan ekspresi gen yang menarik dan laluan yang mungkin berpotensi mengubah peranan penyakit dan perkembangan komplikasi.

DIFFERENTIAL GENE EXPRESSION ANALYSES IN HBE/BETA THALASSAEMIA PATIENTS AND THEIR RELATIONSHIP TO DISEASE SEVERITY

ABSTRACT

Haemoglobin E-beta thalassaemia (Hb E/β-thalassaemia) is a common inherited genetic disorder. It is responsible for approximately half of all severe beta-thalassaemia cases globally. In the state of Kelantan, 50.93% of thalassaemic patients have Hb E/β-thalassaemia, and most of the cases are commonly seen in Malay compared to Chinese and Indian. Clinical heterogeneity is the most outstanding criteria among these patients ranging from mild to severe clinical courses that need a regular blood transfusion. There are many modifiers found to affect the disease presentation. However, the exact reasons behind this heterogeneity are not fully understood. This research aimed to study the differential gene expression and their possible role in the disease presentation and complications development in both transfusion-dependent (TDT) and non-transfusion dependent (NTDT) HbE/β-thalassaemia patients. It was conducted with the aid of RT2 profiler PCR array and microarray that were used in gene expressional study in reticulocytes and erythroid progenitors, respectively. Three normal controls and a total of 20 patients were enrolled in this study; 10 (50%) were TDT, and 10 (50%) NTDT. The reticulocytes study showed the up-regulation of BAX and BAD genes in TDT patients, which have a role in apoptosis induction through the mitochondrial apoptotic pathway. Their up-regulation in TDT may play a role in the reticulocytes' apoptosis, mature RBCs' short life span and eryptosis. Flow cytometry study showed higher apoptosis in the erythroid progenitors of TDT patients. The increased apoptosis in erythroid progenitors and the up-regulation of BAD and BAX

of reticulocytes in TDT may be linked to the down-regulation of the genes involved in the PI3k/AKT pathway in the same patients' group genes. Pathway and ontology analysis showed the involvement of osteoporosis and bone regulating factors related to the VDR pathway and the negative regulation of osteoclast differentiation in the TDT group. The genes involved can be therapeutic targets like SPP1 and MAFB. Their activation act to reduce the disease burden by reducing anaemia and alleviating bone marrow complications. In summary, our study showed the expression of interesting genes and pathways that may potentially modify the disease presentation and the development of the complications.

CHAPTER 1

INTRODUCTION

1.1 Study background

Thalassaemia and hemoglobinopathies are the most common single-gene disorders all over the world (Farashi and Harteveld, 2018; Kountouris et al., 2016; Williams and Weatherall, 2012). In southeast Asia, β-thalassaemia and HbE with α-thalassaemia and Hb CS are very prevalent, and the combination of these genes results in very variable presentation and severity (Organization, 2021). Around half of severe β-thalassaemia worldwide are HbE/β-thalassaemia patients (Olivieri et al., 2011; Yusof et al., 2020). The disease resulted in anaemia with ineffective erythropoiesis and iron overload, and the variable degree of these abnormalities give rise to very heterogeneous presentation ranging from asymptomatic to severe anaemia that needs a regular blood transfusion (Yusof et al., 2020).

HbE/ β-thalassaemia is found at high rates in India, Bangladesh, and throughout Southeast Asia (Olivieri et al., 2011; Yusof et al., 2020). In Malaysia, Hb E β-thalassaemia is common in Orang Asli and Malays of peninsular Malaysia (George, 2013). According to the annual report of the Malaysian thalassaemia registry for 2019, the state of Kelantan has 485 thalassaemic patients, most (460) of them are of Malay descent. HbE/β-thalassaemia constitutes the highest per cent among these patients (50.93%). More than half (51.55%) of these patients are under 25 years. Figures 1.1 and 1.2 illustrate the distribution of thalassaemia patients by diagnosis and age group, respectively (Ibrahim et al., 2019).

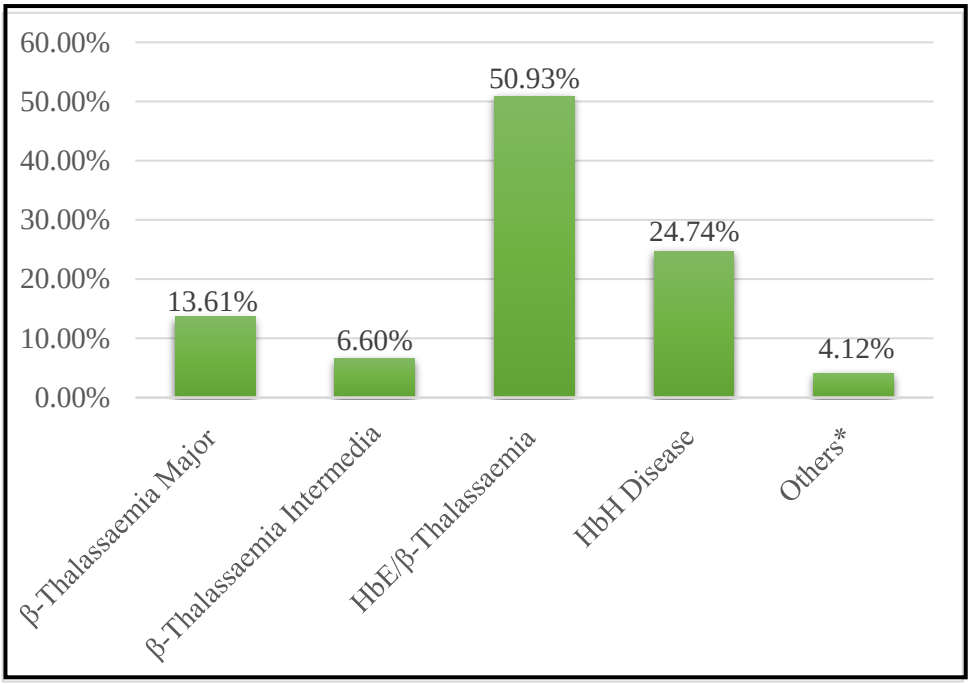

Figure 1.1 Distribution of thalassaemia patients in Kelantan according to diagnosis

*Other types of α-thalassaemia

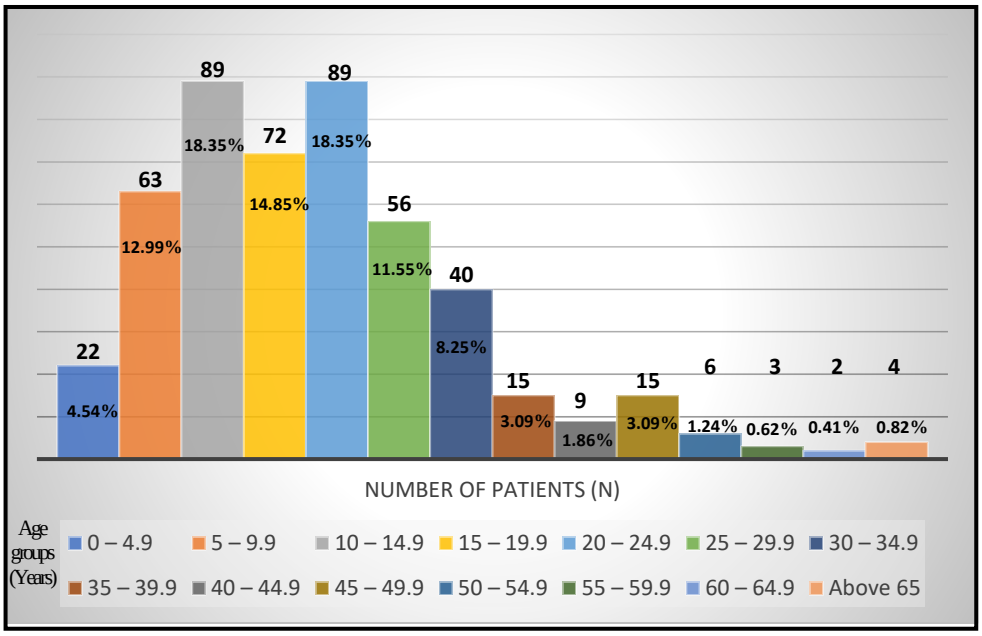

Figure 1.2 Thalassaemia patients' distribution according to age groups in Kelantan

The HbE/β-thalassaemia is very heterogenous on the clinical level. Clinically it is classified into two main categories: transfusion-dependent (TDT) and non-transfusion-dependent thalassaemia (NTDT). Those with TDT usually develop severe anaemia that requires lifelong blood transfusion and probably ends up with iron overload (similar to β-thalassaemia major). NTDT patients, on the other hand, do not need a regular blood transfusion (Taher et al., 2014; Tatu, 2020).

Heterogeneity of thalassaemia is also seen at the molecular level. Interaction between different types of genetic defects can modify and change clinical presentation. The defects directly involve globin synthesis, and others affect the disease process. Genetic modifiers are classified into primary, secondary, and tertiary modifiers. Primary modifiers encompass different mutations of beta-globin genes. More than 350 different mutations have been identified (Jaing et al., 2021), which have a major effect on determining the severity and phenotype of the disease. Secondary modifiers abnormalities are genes that affect the degree of α/non α-chain imbalance. Increased α/non α chain imbalance will result in precipitation of the excessive α-globin in erythroid precursors and peripheral blood red cells resulting in premature destruction of the red blood cells, anaemia and erythroid hyperplasia. Factors that reduce α /non-α chain imbalance can ameliorate or aggravate the disease. Amelioration has been associated with concurrent inheritance of alpha-thalassaemia, increased HbF synthesis and AHSP. Excessive production of chain reduces non-imbalance by binding the extra a to produce HbF (Thein, 2005).

A study showed an improved life quality among TDT children in Malaysia (Shafie et al., 2020). However, despite all the progress achieved in terms of appropriate therapy provided, the overall patients' survival and quality of life, thalassaemia is still

an incurable disease, except for bone marrow transplantation, which is suitable only for selected cases and taking into account that all the other curative modalities of treatment is still under trial or applied at a limited level, further studies and researches are in need hoping for the amelioration of β- thalassaemia burden (Shah et al., 2020)

1.2 Justification of the study

Beta-thalassaemia is still an incurable disease and represents the commonest single-gene disorder worldwide. It is genetically heterogeneous, with many factors playing a role in phenotype variability. β - Thalassaemia is a disease characterized by anaemia and associated with ineffective erythropoiesis. This ineffective erythropoiesis is defined as a suboptimal production of mature erythrocytes originating from a proliferating pool of immature erythroblasts. It is characterized by (1) accelerated erythroid differentiation, (2) maturation blockade at the polychromatophilic stage and (3) erythroid precursors apoptosis. Despite extensive knowledge on molecular defects causing β-thalassaemia, less is known about the mechanisms responsible for the disease severity (Ribeil et al., 2013).

So, this study is conducted to study the differential gene expression in HbE/β thalassaemia and the factors contributing to the severity of disease that can potentially provide information about potential therapy and management options toward better outlook and better quality of life for these patients diagnosed with this disease in the future.

1.3 Research questions

1. What are the possible pathways and genes responsible for apoptosis in reticulocytes and shortened erythrocyte lifespan?
2. What are the differentially expressed genes (DEGs) in HbE/β-thalassaemia patients?
3. What are the pathways associated with the DEGs?
4. What is the potential role of the DEGs and the pathways in the disease severity?

1.4 Objective(s) of the Research

1.4.1 General Objective

This study was designed to investigate the differential gene expression and their potential role in disease severity among HbE/β-thalassaemia patients.

1.4.2 Specific Objectives

1. To characterize the differential expression of apoptosis genes among NTDT and TDT HbE/β-thalassaemia patients.
2. To determine the possible pathways and genes responsible for apoptosis in reticulocytes of the patients by RT2 Profiler PCR array.
3. To generate erythroid progenitors from peripheral blood mononuclear cells of patients and control.
4. To determine the *in vitro* apoptosis among NTDT and TDT HbE/β-thalassaemia patients.
5. To characterize the differential expression of genes in erythroid progenitors among patients and control groups.

6. To validate the differentially expressed gene results from microarray using qPCR expressional study for the genes of interest.
7. To identify the potential role of the differentially expressed genes and the pathways involved in disease severity of HbE/β thalassaemia.

CHAPTER 2

LITERATURE REVIEW

2.1 Thalassaemia

The term thalassaemia originates from two Greek words: Thalassa, which means sea, and “haima”, which means blood (Peters et al., 2012). It is a genetic disorder that results from abnormal haemoglobin synthesis. Normal adult haemoglobin contains alpha and beta chains. In patients with thalassaemia, there is a defect in one or more of these chains (Helmi et al., 2017). This defect can be a quantitative defect seen mainly in α-thalassaemia and β-thalassaemia or a qualitative one as in sickle cell disease. The interactions of these globin defects cause a wide range of syndromes with different clinical presentations and severity (Kohne, 2011; Viprakasit and Ekwattanakit, 2018).

Thalassaemia is inherited in an autosomal recessive pattern. It can be transmitted from one or both parents, and at least one of them is a carrier of the disease. α- thalassaemia commonly occurs due to the deletion of the α-globin gene, while β- is mainly caused by mutations (Bajwa and Basit, 2019), and more than 350 mutations have been identified so far (Jaing et al., 2021).

Thalassaemia is classified according to transfusion frequency into TDT and NTDT thalassaemia. TDT patients are commonly presented early in life with severe anaemia that requires lifelong regular blood transfusion to keep the Hb level between 9-10g/dL, which necessitate chelation therapy to avoid iron overload and its

complications (Musallam et al., 2021). Blood transfusion is usually given every two to five weeks, depending on the transfusion needs (Cappellini *et al.*, 2014; Langhi *et al.*, 2016). On the other hand, NTDT is composed of those who do not need a regular transfusion, although they may need occasional transfusions (Musallam *et al.*, 2013; Shafie *et al.*, 2021; Taher *et al.*, 2021; Taher *et al.*, 2014).

Despite being a monogenic disorder (Saeed and Piracha, 2016) and with all the advances in terms of understanding of its pathophysiology, diagnosis and treatment, thalassaemia is still a very complex disease that requires good comprehension of all the disease's aspects.

2.1.1 The haemoglobin gene

Humans carry six globin genes arranged on two chromosomes: 16 and 11. Five globin genes are on chromosome 11 (Das and Sharma, 2019), collectively called β-like genes. β and δ are adult genes while γ (Gγ and Aγ) and ε are embryonic genes (Ohls, 2017). The remaining two genes are α and ζ globin genes on chromosome 16 and are called α–like genes (Figure 2.1) (Higgs, 2013). Except for α and γ which have two gene copies in each chromatid (i.e., Four genes per diploid cell), all the other globin genes have one gene copy in each chromatid (i.e., two genes per diploid cell) (Randolph, 2020).

The genes of the α-like globin are clustered near to the telomere of chromosome 16 at around 150 kb from it (16p13.3). This gene cluster is controlled by four noncoding multispecies conserved sequences (MCS-R1 to MCS-R4) and other

conserved cis-acting regulatory sequences (Farashi and Harteveld, 2018; Mettananda et al., 2015a). α-globin is encoded by alpha-1 and alpha-2 genes that produce two types of α-globin, HBA1 and HBA2, respectively. The alpha-1 and alpha-2 are almost identical, with a minor difference at the introns and 5' untranslated regions (Denton et al., 2021; Kalle Kwaifa et al., 2020).

Alpha-thalassaemia develops due to the deletion of one or more of the α- genes, and the disease severity is directly proportionated with the number of genes involved. Mutational α-thalassaemia also occurs but less frequently (Farashi and Harteveld, 2018).

Beta-globin (HBB) gene is located on the short arm of chromosome 11 (11p15.5); it is composed of two introns, three exons encoding 147 amino acids with around 1.6 kilobases molecular weight (Mashi et al., 2017; Poon and Tan, 2021). The β-globin Locus Control Region (LCR) is a cis-regulatory element composed of a set of DNAse I hypersensitive (HS) sites; each is separated by two to four kbp DNA and around 200 to 400 bp in size. LCR is mainly in charge of regulating the expression of all the β- like globin genes and is located upstream of the β-like globin that is arranged in the same order of their expression (Liang et al., 2008).

Unlike α-thalassaemia, β-thalassaemia can be caused by point mutations, insertions or small deletions in the coding region and exon-intron junction (Murad et al., 2021; Traivaree et al., 2014).

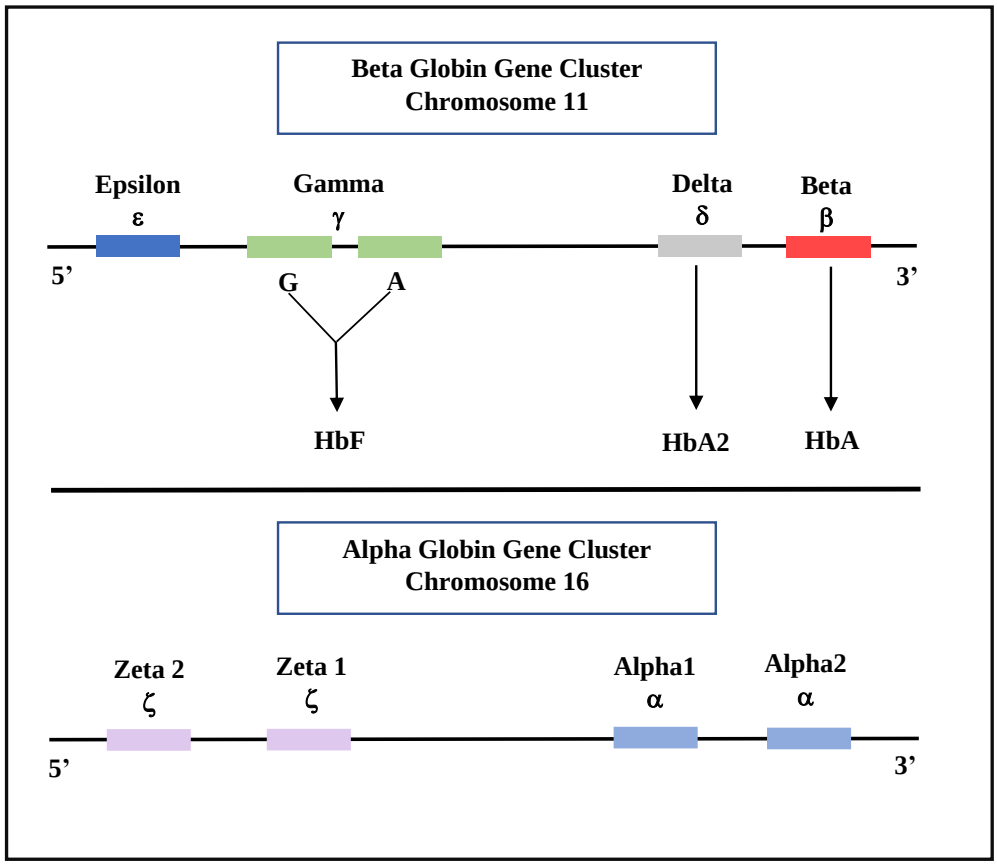

Figure 2.1 Schematic representation of the globin gene loci.

2.1.2 The haemoglobin protein

Haemoglobin (Hb) in the red blood cells is responsible for transporting oxygen from the lungs to the tissues and transporting carbon dioxide from the tissues back to the lungs. Hb shows a high oxygen (O2) affinity and a low carbon dioxide (CO2) affinity in arterial blood with reverse affinities in venous blood (Fucharoen and Winichagoon, 2002; Marengo-Rowe, 2006). Hb is composed of heme and globin parts that bind together to form a tetramer. The heme has a protoporphyrin ring with iron ion (Fe2+) in the centre, the globin part in adult Hb has two α-globin and two β-globin. They bind together with the heme to form HBA (Farid et al., 2021).

The process of binding and releasing O2 are well-adjusted mechanisms. The Hb molecule is adapted to handle the issues associated with exposure to free O2. For a long time, the Hb has been an interesting topic of study. In 1961, it was one of the earliest proteins to be sequenced, and in 1976, globin genes had been cloned (Thom et al., 2013).

2.1.3 The haemoglobin development

Primitive erythropoiesis is detected as early as 14 to 19 days of conception in the yolk sac, which continues until the ninth gestational week. In the yolk sac, the primitive erythroblasts experience Hb switch: primary Hb Gower I is produced in the embryo before five weeks of gestation, then at week 6 to 8, erythroblasts start to synthesise Gower II (Qiu et al., 2008). Hb Gower I (ζ2ε2) is unstable and easily being broken down. It is composed of two zeta (ζ) chains and two epsilon (ε) chains. Hb Gower II (α2ε2), composed of two (α) alpha and two (ε)epsilon globin chains, is more stable than Hb Gower I (Jacob, 2016).

Definitive erythropoiesis is detected during the sixth week of gestation for the first time in the foetal liver. α, γ, ζ, ε and small amounts of β-globin are expressed in foetal liver erythroblasts. ζ, and ε -globin genes are quickly silenced, but α and γ, globin genes continue expressing until birth. At around the eleventh week of gestation, erythropoiesis in bone marrow is detected and will become the main source of erythropoiesis (figure 2.2) (Qiu et al., 2008).

After the embryonic globin expression, Hb switch, which is a complex process of globin gene expression, will occur during foetal development and the first six months after birth. First, the γ-chain of foetal haemoglobin (HbF), which is composed of two α and two γ globin chain is expressed predominantly then is replaced by the β-globin chain that binds to α-globin to form the adult haemoglobin (HbA) (Forget, 2011).

In full-term new-borns, HbF represents 60–80% of total haemoglobin, which starts to reduce around 6-12 months after birth to reach <1% of haemoglobin in normal adults (Sotoudeh and Sotoudeh, 2020). HbA (α2β2) constitutes around 97% of total Hb in healthy adults, and HbA2 constitutes around 2-3% (Weatherall, 2001).

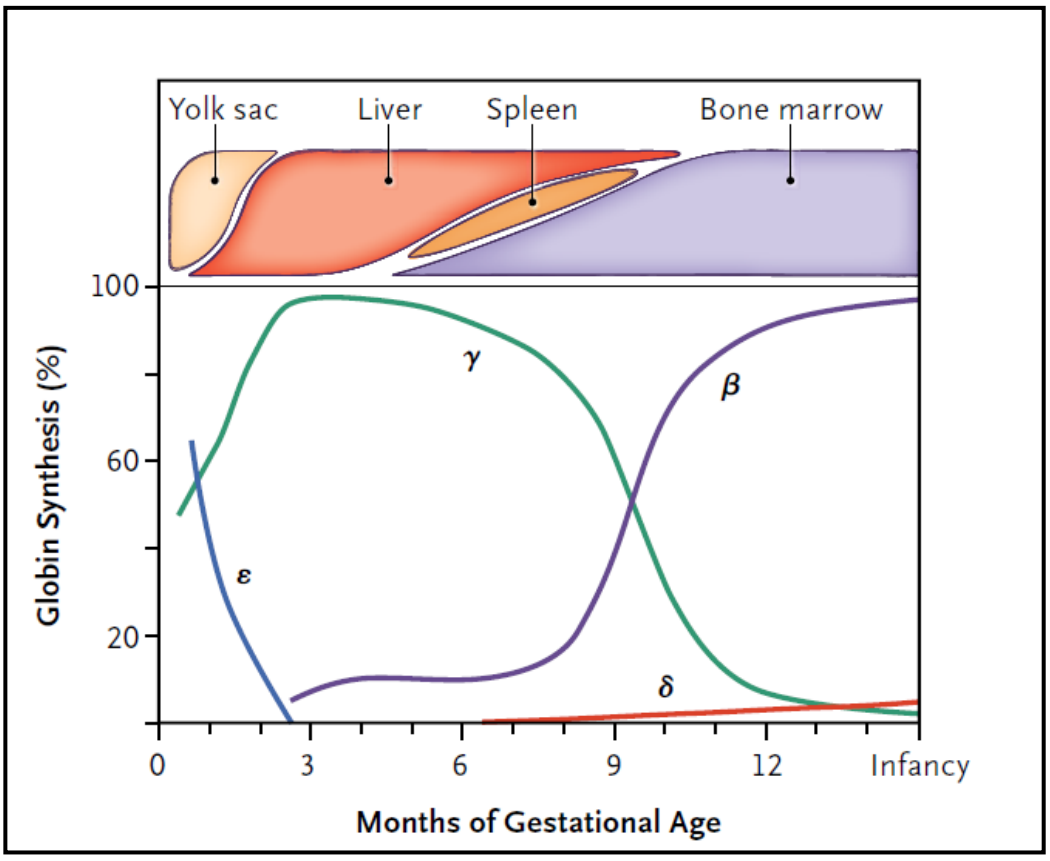

Figure 2.2 The expression pattern of different globin genes and erythropoiesis sites during pregnancy adapted from (Sankaran and Nathan, 2010).

2.1.4 Epidemiology

Thalassaemia nowadays is widely distributed all over the world. The World Health Organization (WHO) in 2018 stated in their report that around 5.2% of people worldwide were carriers of thalassaemia, and 1.1% of couples at risk of having a child with hemoglobinopathy(Apidechkul et al., 2021). Thalassaemia shows high incidence mainly in the Middle East, Southeast Asia (SEA), the Indian subcontinent, and parts of Africa (Sanctis et al., 2017). In SEA α– and β-thalassaemia, Hb Constant Spring and HbE are predominant. The different combinations of these hemoglobinopathies make more than 60 different types of thalassaemia syndromes making SEA a place of very complex genotypes of thalassaemia. (Fucharoen and Winichagoon, 2011).

In a recent meta-analysis to study the α-thalassaemia prevalence between 2010 and 2020 in SEA, Vietnamese showed the highest rates with 51.5% prevalence. Cambodia is coming in second place with a prevalence rate of 39.5%. The remaining prevalence rates were 17.3%, 20.1% and 26.8% in Malaysia, Thailand and Laos, respectively (Goh et al., 2020).

According to the Ministry of Public Health in Thailand, around 18-24 million people are carriers of thalassaemia (30-40%). Six hundred thousand are under the health system and need regular treatment, and more than 12000 born infants are affected by thalassaemia. A study conducted among 1,796 Thai females in the north-eastern part of the country showed the highest prevalence of thalassaemia,

homozygous HbE, α-thalassaemia one trait and β-thalassaemia trait in the country that reach 30.2%, 5.4%, 3.0% and 0.6%, respectively (Apidechkul et al., 2021).

In Indonesia, 1035 individuals have β- thalassaemia and around 43.0% of them are registered with West Java Province (Panigoro et al., 2019). Records also show a high prevalence of thalassaemia carriers, with around 3-20% are carriers of α-thalassaemia, 3% and 1-33% were carriers of β-thalassaemia and HbE, respectively (Husna et al., 2017).

In Singapore, a cord blood genotype study showed that α-thalassaemia presented in 6.4% of Chinese, 5.2% of Indians and 4.8% of Malays. β-thalassaemia was highest in Malays (6.3%) followed by Chinese (2.7%) and Indians (0.7%). The HbH and Barts hydrops fetalis in Chinese, HbE/ β thalassaemia in Malays and β-thalassaemia major among all groups are the main concerns of thalassaemia screening (Lee et al., 2019). Around 49% of Hb E/β-thalassaemia are on regular transfusion and iron chelation therapy (Tan et al., 2014).

In Malaysia, it is estimated that around 6.8% of Malaysians are thalassaemia carriers, and 7984 patients in the Malaysian Thalassaemia Registry were reported alive. Figure 2.3 shows the cumulative numbers of patients registered in each state of Malaysia, with the highest number of registrations are in Sabah, followed by Selangor, Kedah, Johor and Perak. The majority of patients registered were Malay (63.95%), followed by Chinese (11.75%) and Kadazan-Dusuns (11.36%). HbE/β-thalassaemia represents the majority of the cases (34.37%), followed by thalassaemia major (33.52%) and HbH disease (18.26%). The majority (88.96%) of HbE/β-thalassaemia were Malay patients, and the highest proportions of HbE/β-thalassaemia were found

in Selangor, Kuala Lumpur, Kelantan and Kedah. This is reasonably explained in Selangor and Kuala Lumpur by being the centre of thalassaemia referral, while in Kedah and Kelantan, their closeness to Thailand promotes the interaction between the two populations of both countries (Hishamshah Mohd Ibrahim et al., 2020)

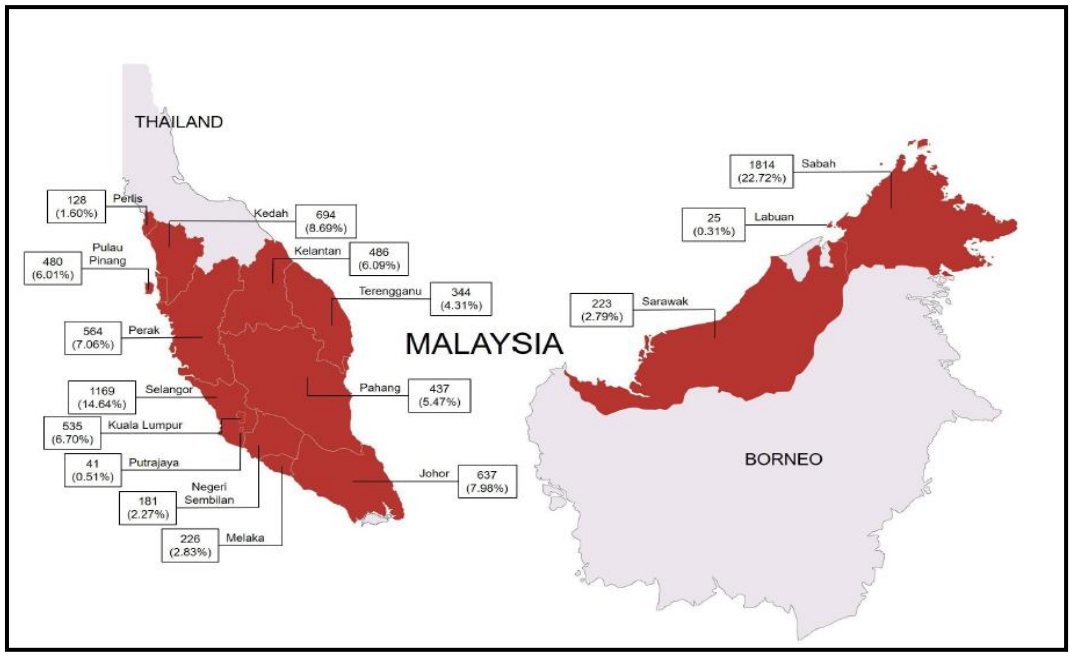

Figure 2.3 Registered patients with thalassaemia across Malaysia adapted from (Hishamshah Mohd Ibrahim et al., 2020)

2.2 Classification of thalassaemia

Thalassaemia is mainly classified into α- and β-thalassaemia, characterized by reduction or absence of α- and β-globin production, respectively. There are other rare types of thalassaemia where the gamma (γ)-globin or delta (δ)- production is affected as well as combined deficiencies of two or more globin subunit production (Forget and Franklin Bunn, 2013). Table 2.1 shows the genotype and phenotype classification of thalassaemia.

Table 2.1 Classification of thalassaemia. Adapted from (Viprakasit et al., 2014)

Type of thalassaemia	**Variant**	**Genotype**	**Phenotype**
α-thalassaemia	Normal	αα/αα	Normal
	Silent Carrier	-α/αα	Haematologically silent or mild reduction of MCH/MCV
	α-thalassaemia minor	-α/-α --/αα	Borderline anaemia or normal, as well as microcytic and hypochromic red blood cells
	HbH disease	--/-α	Majority have mild to moderate anaemia and marked microcytosis and hypochromia; few have a phenotype similar to thalassaemia major
	Barts Hydrops Fetalis	--/--	Most develop hydrops fetalis syndrome and die *in utero* or shortly after birth
β-thalassaemia	Normal	β/β	Normal
	β-thalassaemia minor	β/β^{++} β/β^{0}	Borderline anaemia or normal as well as microcytic and hypochromic red blood cells
	β-thalassaemia intermedia	β^{+}/β^{+} β^{0}/β^{+} β^{+}/β β^{0}/β β^{0}/β^{0}	Severity is very variable. The clinical picture ranges from mild to moderate NTDT
	β-thalassaemia major	$\beta^{0}\beta^{0}$ β^{0}/β^{+}	Severe anaemia requiring regular transfusions (TDT)
HbE	HbE trait	β^{E}/β	Asymptomatic condition with no clinical relevance
	Homozygous HbE	β^{E}/β^{E}	Usually asymptomatic with borderline asymptomatic anaemia and no haemolysis
	HbE/ β-thalassaemia	β^{E}/β^{+} β^{E}/β^{0}	Severity is very variable. The clinical picture ranges from NTDT to TDT
	HbE/HbS	β^{E}/β^{S}	Similar to sickle cell disease, usually with rare vaso-occlusive crisis

2.2.1 Alpha- thalassaemia

Unlike β-thalassaemia, α-thalassaemia results in most cases from a deletion of one or more α-globin genes, although rarely it can be caused by mutation. α-thalassaemia is subclassified into; α^{+}-thalassaemia: this type is caused by one α-gene deletion and is usually asymptomatic. The patient who has this type of thalassaemia is called the silent carrier, α^{0}-thalassaemia caused by two genes deletion and causes mild hypochromic microcytic anaemia phenotype, HbH disease results from three gene deletions, and finally hydrops fetalis with Hb Bart's results from four α-gene deletions (Forget and Franklin Bunn, 2013). The last two types are clinically significant: HbH disease has a wide range of clinical presentations, it is usually presented early in life. However, it can delay until adulthood. Patients usually develop jaundice, splenomegaly and infection that induces acute haemolysis. Hb Bart's is incompatible with life, and it is usually started in the womb. The affected foetus develops massive hepatosplenomegaly, generalized oedema, pericardial and pleural effusion with extramedullary erythropoiesis (Tamary and Dgany, 2020).

The non-deletional α-thalassaemia Hb Constant Spring (HbCS) is common in SEA. A chain termination mutation results in abnormally elongated α-globin that deposits in the red cells. It is clinically significant when combined with α^{0}-cis deletion on the other chromosome as it causes severe HbH disease (Higgs, 2013)

2.2.2 β-thalassaemia

β-thalassaemia results from either inadequate (β^{+}or β^{++}) or absence (β^{0}) of β-globin chains. Most β-thalassaemia is caused by mutation and rarely by gene deletion (Kohne, 2011; Lee et al., 2019).

2.2.2(a) β-thalassaemia trait

β-thalassaemia trait (also called minor or carrier status) occurs due to heterozygous type of inheritance with a prevalence of 1.5% worldwide (Needs et al., 2020). It is caused by a mutation of a single β-globin gene, which reduces HbA (α2β2) (El-Beshlawy et al., 2005; Vasileiadis et al., 2009). This reduction usually causes a slight reduction of the Hb level, and patients are usually asymptomatic or have a mild hypochromic microcytic anaemia, with low mean corpuscular volume (MCV) and elevated Hb A2 and reduced HbA by high-performance liquid chromatography (HPLC) (Surapon, 2011). Genetic counselling and prenatal diagnosis are important when carriers are detected (Needs et al., 2020), as the risk of having a child with homozygous beta thalassaemia is 25% when both parents are carriers (Galanello and Origa, 2010). Normal HBB is assigned as 'β'. Patients with thalassaemia minor are either (β/β^{+}) or (β/β^{o}) (Hanafi et al., 2020).

2.2.2(b) β-thalassaemia intermedia (β-TMi)

This type of β-thalassaemia lies between the asymptomatic thalassaemia trait and the severe thalassaemia major category, usually having a milder presentation and

no need for regular transfusion as the patients can maintain a Hb of 7 g/dL (Surapon, 2011).

β-TMi individuals have either (β^o /β^+) or (β^+ /$\beta^{+)}$ mutation (Hanafi et al., 2020). Clinically, this type can be divided into two subgroups: a milder form that is slightly affected by the anaemia, their Hb level usually range 7-11g/dL and rarely need blood transfusion; the second type has more severe anaemia presented at 2-6 years of age with Hb level of 7g/dL, and they require more frequent blood transfusion and are still at risk of developing skeletal deformity osteoporosis, pathologic fractures of long bones and growth retardation secondary to bone marrow hypertrophy and extramedullary erythropoiesis. Careful assessment is required to differentiate them from β-thalassaemia major (Nasiri et al., 2020).

2.2.2(c) β-thalassaemia major (β-TM)

Patients with β-TM are carrying (β^o /β^o) mutation (Hanafi et al., 2020). They usually present between the age of six months and two years and require a regular blood transfusion to survive. Infants with β-TM start to develop pallor and fail to thrive. They also suffer from diarrhoea, irritability, feeding problems and probably splenomegaly that will cause abdominal enlargement. Poorly managed or untreated patients usually develop many complications: jaundice, poor musculature growth retardation, hepatosplenomegaly, leg ulcers and skeletal changes like typical craniofacial changes, long bones deformity and osteoporosis. Death from heart failure is the main cause of death among those who did not receive a blood transfusion. With blood transfusion, patients' lives can be saved and maintaining the Hb level between

9-10.5 g/dL; the patient can develop normal growth and development until 10 to 12 years. However, blood transfusion is not side effect free; iron overload usually develops if the patient does not receive optimum chelation therapy (Origa, 2017a). In order to reduce the burden of iron overload and consequent complications, iron chelators are given to these patients (Darvishi Khezri et al., 2016). Deferoxamine is the standard chelator used, but because of the side effects of these drugs, the use of natural iron has been studied (Darvishi-Khezri et al., 2018; Hagag et al., 2015; Nasiri et al., 2020).

2.2.3 HbE / β-thalassaemia

HbE is a β-chain haemoglobin variant. Those with homozygous HbE are clinically normal with a mild phenotype of β-thalassaemia. However, co-inheritance of HbE with β- thalassaemia will result in compound heterozygous thalassaemia (HbE/β^{+} or HbE/β^{0}) that may be presented with severe clinical manifestation (Olivieri et al., 2011). HbE results from a change of Glutamic acid to Lysine caused by the base substitution of G to A base at codon 26 on exon 1. This will produce abnormal Hb structure and cryptic splice site activation, leading to abnormal mRNA processing. The β^{E} globin gene dramatically decreased the production of β^{E} -globin mRNA and β^{E}-globin chain. (Hanafi et al., 2020; Olivieri et al., 2011).

One of the most outstanding features of HbE β thalassaemia is the clinical heterogeneity of the disease. On the one hand, some patients presented with a severe clinical course that is difficult to differentiate from severe β-thalassaemia major: On the other hand, some patients can develop and grow normally without the need for blood transfusion, although they commonly have low haemoglobin levels (Fucharoen

and Weatherall., 2012). the reasons behind this clinical heterogeneity are not fully understood (Olivieri, 2012). HbE/β-thalassaemia pathophysiology is linked to numerous factors, including globin chain imbalance, oxidative damage, apoptosis, ineffective erythropoiesis, and reduced RBCs survival (Datta et al., 2006; Olivieri, 2012; Pootrakul et al., 2000). HbE β thalassaemia has three categories.

2.2.3(a) Mild Hb-E/β-thalassaemia

The patients in this category usually have no symptoms. They represent 15% of all southeast Asia cases (Surapon, 2011). Although the heterogeneity in terms of clinical presentation of this disease is not well understood, the investigators from Thailand proposed that the coinheritance of mild β-thalassaemia allele will give rise to mild disease; however, the coinheritance of the same severe β-thalassaemia mutation is seen in mild and severe patients (Olivieri et al., 2010)

2.2.3(b) Moderately severe HbE/ β-thalassaemia

The Hb level is between 6-7g/dL, and the anaemia is tolerable by these patients, and no transfusion is required unless the anaemia worsens secondary to infection (Jha and Jha, 2014). Most of the Hb-E/β-thalassaemia patients fall into this group (Galanello and Origa, 2010).

2.2.3(c) Severe HbE/ b-thalassaemia

The patients in this group show clinical presentation similar to that seen in β-thalassaemia major patients and are treated with regular blood transfusion. The Hb levels range between 4-5g/dL (Jha and Jha, 2014; Surapon, 2011)

2.3 Genotype-phenotype modifiers

Phenotype is the observed characteristics of an individual, while the genotype represents the collection sets of genetic material gained from parents. Understanding the complex genotype-phenotype relationship is still a big challenge, how genotypes affect phenotype and how these genes interact with the environment to produce the phenotype (Orgogozo et al., 2015). Despite being a single gene disorder, β-thalassaemia shows a great range of clinical variability; patients can be presented with severe anaemia that requires regular blood transfusion or presented with mild anaemia or even an asymptomatic form of thalassaemia (Danjou et al., 2011).

The clinical diversity of β- thalassaemia is affected by many genetic factors named modifiers. The phenotype of the disease can be changed by co- inheritance of one or more of these modifiers. Genetic modifiers can be divided into three main types depending on the type of the gene involved:

2.3.1 Primary modifiers

The main genetic factor that determines β-thalassaemia severity is the type of inherited β allele. More than 350 mutations cause β -thalassaemia all over the world, while deletions are responsible for rare cases. Some mutations will produce no β chain

(β^0), others produce β chain but less than normal level (β^+) (Jaing et al., 2021). The location of the mutation plays a role in the phenotypic picture of the disease. Promoter and 5' UTR mutations usually resulted in a transcription-code defect and presented as a mild decrease in β -globin production. Heterozygote inheritance may cause no symptoms at all. Consensus sequence and splice-junction mutations, polyadenylation, and other 3' UTR mutations usually cause a shortage of mRNA and complete β -globin deficiency (β0 phenotype). And finally, nonsense, frameshift, and initiation codon mutations may result in variable degrees of β -globin reduction giving rise to a wide range of severity (Pavlovic et al., 2015)

2.3.2 Secondary modifiers

Those genetic factors modulate the globin chain imbalance by affecting globin chain production other than β-chain. The factors that reduce α globin production will ameliorate α/β-globin ration imbalance and reduce the toxic effect of free α globin. A common example is seen in β- thalassaemia with α-thalassaemia coinheritance.

Genetic factors that affect α-globin chains stability are also found to play a role in the disease severity, e.g., alpha haemoglobin stabilizing protein (AHSP). High γ chains production also plays a role in the disease severity as the excess γ chains will help to compensate for the lack of β chain by binding the excess α globin, so inheriting genes that can change γ chains production like BCL11A, HBG2 and others may participate in the phenotypic variability (Galanello, 2012; Pavlovic et al., 2015).