DETERMINATION OF GENETIC POLYMORPHISM IN ZHX2 GENE AMONG β°-THALASSEMIA/Hb E

CHIRANYA SANGPRASERT

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE (BIOCHEMISTRY)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2008

Thesis
Entitled

DETERMINATION OF GENETIC POLYMORPHISM IN ZHX2 GENE AMONG β°-THALASSEMIA/Hb E

…….
Miss. Chiranya Sangprasert,
Candidate

…….
Mr. Chayanon Peerapittayamongkol,
M.D. Ph.D.
Major-Advisor

…….
M.L. Saovaros Svasti,
Ph.D.
Co-Advisor

…….
Assist.Prof. Chatchawan Srisawat,
M.D. Ph.D.
Co-Advisor

….……
Prof. Banchong Mahaisavariya,
M.D.
Dean
Faculty of Graduate Studies

…….
Assist.Prof. Vorapan Sirivatanauksorn,
M.D. Ph.D.
Chair
Master of Science Programme in Biochemistry
Faculty of Siriraj Hospital

Thesis
Entitled

DETERMINATION OF GENETIC POLYMORPHISM IN ZHX2 GENE AMONG β°-THALASSEMIA/Hb E

was submitted to the Faculty of Graduate Studies, Mahidol University
for the degree of Master of Science (Biochemistry)
on
31 October, 2008

…….
Miss. Chiranya Sangprasert,
Candidate

…….
Mr. Pisut Katavetin,
M.D. M.Sc
Chair

…….
Mr. Chayanon Peerapittayamongkol,
M.D. Ph.D.
Member

…….
M.L. Saovaros Svasti,
Ph.D.
Member

…….
Assist.Prof. Chatchawan Srisawat,
M.D. Ph.D.
Member

…..……
Prof. Banchong Mahaisavariya,
M.D.
Dean
Faculty of Graduate Studies
Mahidol University

…….
Clin.Prof. Teerawat Kulthanan,
M.D.
Dean
Faculty of Medicine, Siriraj hospital
Mahidol University

ACKNOWLEDGEMENTS

At the outset, the success of this thesis can be attributed to the extensive support and assistance from my major advisor, Dr. Chayanon Peerapittayamongkol. I take the privilege in expressing my deep sense of gratitude to him for his kindness, guidance and generous advice throughout this research.

I would like to sincerely thank you to my co-advisor Dr. M.L. Saovaros Svasti and Assist.Prof. Dr. Chatchawan Srisawat, for their valuable suggestions, guidance and comments to my study.

I am also very grateful to Prof. Suthat Fucharoen, Thalassemia Research Center, Institute of Science and Technology for Research and Development, Mahidol University, for providing samples in this study and for his kindness, helpful suggestions and guidance.

I am profoundly thankful to Mr. Pisut Katavetin, chair of the thesis examination committee, for his valuable comments and kindness.

I appreciate Miss Aree Futrakul, Miss Thongperm Munkongdee and all TRC colleagues for being kind, helpful, introducing me to several techniques in Molecular Biology and co-operative throughout my study.

Special thanks given to my friends, the graduate students in Department of Biochemistry for their friendships and generous helps in many ways that would always be remembered.

Thanks are also due the Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand for the scholarship to support this study.

Last but certainly not the least, I fall short of vocabulary in expressing my special appreciation to my dearest parents and younger brother who have always supported and believed in me.

Chiranya Sangprasert

DETERMINATION OF GENETIC POLYMORPHISM IN ZHX2 GENE AMONG β°-THALASSEMIA/Hb E

CHIRANYA SANGPRASERT 4836208 SIBC/M

M.Sc. (BIOCHEMISTRY)

THESIS ADVISORS: CHAYANON PEERAPITTAYAMONGKOL, M.D. Ph.D. (BIOCHEMISTRY), M.L. SAOVAROS SVASTI, Ph.D. (BIOCHEMISTRY), CHATCHAWAN SRISAWAT, M.D. Ph.D. (BIOCHEMISTRY).

ABSTRACT

β°-Thalassemia/Hb E disease is a common genetic disease that reflects from the defective β-globin chain synthesis. Several lines of evidence suggest that cosegregation together with genetic determinants of high levels of Hb F which alter the ratio of imbalance of α-and β-globin chains should ameliorate the severity of symptoms. ZHX2 gene is among other candidate genes located in chromosome 8q region which is identified as the QTLs influencing Hb F. Moreover, ZHX2 expressions have shown in various tissues. Therefore, this study determined the potential effects of single nucleotide polymorphism (SNP) in the ZHX2 gene (G779A) in influencing percentage of fetal hemoglobin and examined ZHX2 expression altered in cultured erythroid cells of the β°-thalassemia/Hb E patients. The expression level of ZHX2 during erythroid developmental stages in mild and severe cases was analyzed using reverse transcriptase polymerase chain reaction (RT-PCR). The results show that there is ZHX2 expression in cultured erythroid cells at low level. The effect of this polymorphism was studied in 450 β°-thalassemia/Hb E patients by polymerase chain reaction-based restriction enzyme analysis (PCR-RFLP). The results showed that this polymorphism was not associated with Hb F level and severity of β°-thalassemia/Hb E. The study therefore concludes that the actual causative variant should be genetic loci other than those proposed to be associated with Hb F level for β°-thalassemia/Hb E.

KEY WORDS : BETA-THALASSEMIA/Hb E / Hb F / ZHX2 GENE / SNP

119 pp.

การศึกษาความหลากหลายทางพันธุกรรมของยีน ZHX2 ในผู้ป่วยเบต้าธาลัสซีเมีย/ฮีโมโกลบินอี
(DETERMINATION OF GENETIC POLYMORPHISM IN ZHX2 GENE AMONG β°-THALASSEMIA/Hb E)

จิรัญญา แสงประเสริฐ 4836208 SIBC/M

วท.ม.(ชีวเคมี)

คณะกรรมการควบคุมวิทยานิพนธ์: ชยานนท์ พีระพิทยมงคล, M.D.Ph.D.(BIOCHEMISTRY), ม.ล. เสาวรส สวัสดิวัฒน์, Ph.D. (BICHEMMISTRY), ชัชวาลย์ ศรีสวัสดิ์, M.D.Ph.D. (BIOCHEMISTRY)

บทคัดย่อ

โรคเบต้าธาลัสซีเมีย/ฮีโมโกลบินอีเป็นโรคทางพันธุกรรมที่เกิดจากความผิดปกติของการสังเคราะห์สายเบต้าโกลบิน โดยมีการศึกษาก่อนหน้านี้พบว่าพันธุกรรมหรือยีนที่เกี่ยวข้องกับการแสดงออกของปริมาณฮีโมโกลบินเอฟที่สูงจะช่วยลดความรุนแรงของโรคได้ โดยจะไปช่วยลดความไม่สมดุลของสายแอลฟาและเบต้าโกลบิน ซึ่งยีน ZHX2 นั้น เป็นหนึ่งในยีนที่น่าสนใจ คือ อยู่บนโครโมโซม 8 ที่มีการรายงานว่าเป็น QTLs ที่มีผลต่อระดับฮีโมโกลบินเอฟและยังมีรายงานว่ายีน ZHX2 มีการแสดงออกในหลายๆ tissue ด้วย ดังนั้นการศึกษาครั้งนี้มีวัตถุประสงค์เพื่อศึกษาผลของความหลากหลายทางพันธุกรรมชนิดการเปลี่ยนนิวคลีโอไทด์กวานีนเป็นอะดินีน ที่ตำแหน่ง 779 ของยีน ZHX2 ว่ามีผลต่อระดับฮีโมโกลบินเอฟหรือไม่ และศึกษาระดับการแสดงออกของยีน ZHX2 ใน cultured erythroid cells ในผู้ป่วยเบต้าธาลัสซีเมีย/ฮีโมโกลบินอี ว่ามีระดับการแสดงออกของยีนใน erythroid developmental stages เป็นอย่างไรในผู้ป่วยที่มีอาการน้อยและมีอาการรุนแรง โดยวิธี reverse transcriptase polymerase chain reaction (RT-PCR) พบว่าการแสดงออกของยีน ZHX2 ใน cultured erythroid cells อยู่ในระดับต่ำ และการศึกษาผลของความหลากหลายทางพันธุกรรมที่ตำแหน่ง 779 ของยีน ZHX2 โดยวิธี polymerase chain reaction-based restriction enzyme analysis (PCR-RFLP) ในผู้ป่วยเบต้าธาลัสซีเมีย/ฮีโมโกลบินอี 450 คน พบว่าไม่มีความสัมพันธ์ต่อระดับ Hb F และความรุนแรงของอาการของโรค ดังนั้น จึงสรุปได้ว่า น่าจะมี genetic loci อื่นๆที่เกี่ยวข้องกับระดับ Hb F ในผู้ป่วยเบต้าธาลัสซีเมีย/ฮีโมโกลบินอี

119 หน้า

CONTENTS

CONTENTS (continued)

LIST OF TABLES

Page

LIST OF TABLES (continued)

Page

LIST OF FIGURES

Page

LIST OF FIGURES (continued)

Page

LIST OF ABBREVIATIONS

α	alpha
β	beta
γ	gamma
%	percent
°C	Degree Celcius
µg	microgram(s)
µM	micromolar
µl	microliter(s)
aa	amino acid(s)
Abs F	absolute fetal hemoglobin
bp	base pair
cm	centimeter (s)
Cor% F	correct fetal hemoglobin.
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
dGTP	deoxyguanosine 5'-triphosphate
DNA	Deoxyribonucleic Acid
dNTPs	dATP, dCTP, dGTP and dTTP
dTTP	deoxythymidine 5'-triphosphate
EDTA	Ethylene Diamine tetra Acetic Acid
e.g.	*exempli gratia* (for example)
et al.	*et alii* (and other people)
g	gram(s)
g/dl	gram per deci liter
HS	hypersensitive site
IVS	intervening space
kb	kilobase (s)

LIST OF ABBREVIATIONS (continued)

kg	kilogram (s)
LCR	locus controlling region
LD	linkage disequilibrium
Hct	Hematocrit
M	molar
mA	miliampere(s)
MCH	mean corpuscular hemoglobin
MCHC	mean corpuscular hemoglobin concentration
MCV	mean cell volume
mg	milligram(s)
min	minute(s)
ml	milliliter(s)
mM	millimolar
mm	millimeter(s)
$MgCl_2$	magnesium chloride
ng	nanogram(s)
mRNA	messenger ribonucleic acid
PCR	polymerase chain reaction
QTL	quantitative trait
rpm	revolutions per minute
RFLP	restriction fragment length polymorphism
sec	second(s)
SNP	single nucleotide polymorphism
TBE	tris-borate EDTA
U	unit(s)
V	volt(s)

CHAPTER I
INTRODUCTION

Thalassemia is an autosomal recessive inherited blood disease affecting a person's ability to produce hemoglobin, the important protein in red blood cells that carry oxygen to every cell in the body. It is one of the most frequent genetic diseases that are distributed worldwide. Thalassemia is particularly prevalent in areas in which malaria is or was once endemic (Figure 1). Therefore, it occurs throughout the Mediterranean, the Middle East, the India subcontinent, Southeast Asia including Thailand (1). These lead to the hypothesis that the genetic effect of this disorder may confer resistance to malaria parasite. The resulting environmental alterations inside red blood cells do not promote normal growth and survival of malaria parasite (2).

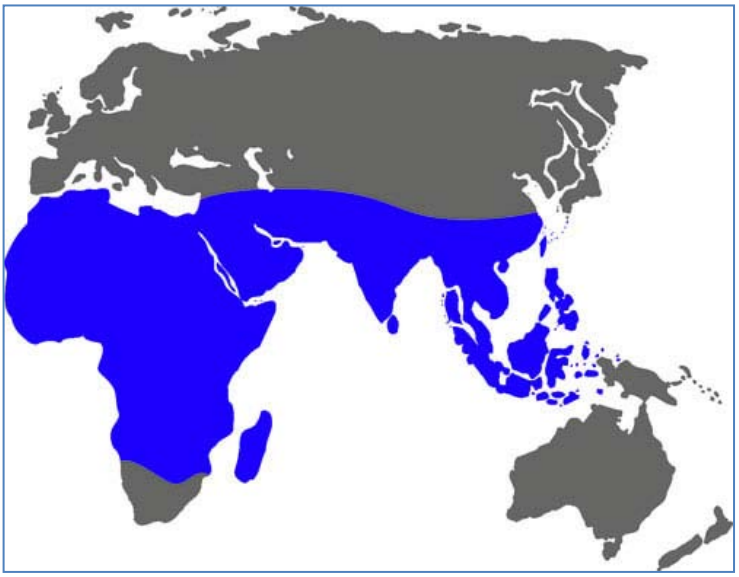


Figure 1 Worldwide distribution of thalassemia, paralleling the distribution of malaria (blue area) (3).

Thalassemia is classified into α, β, γ, δβ, δ and εγδβ thalassemias depending on the globin chains which are affected. Defective synthesis of one of the globin chain leads to the imbalance of α/β chains production and consequently precipitation of excessive unmatched normal synthesized globin chains. The excess globin chains precipitate in the bone marrow, lead to red blood cell membrane rigidity, ineffective erythropoiesis and premature red blood cell destruction resulting in anemia. The two important forms of this disorder, α- and β-thalassemia, emanate from the defective synthesis of the α- or β-globin chains of hemoglobin, respectively (4). Gene deletion is a common cause of α-thalassemia (5, 6), whereas base substitutions or small deletions in the β-globin gene are the major cause of β-thalassemia (7-10).

β-thalassemia syndromes encompass a wide variety of clinical phenotypes ranging in severity from clinically silent heterozygous β-thalassemia to severe transfusion-dependent thalassemia major (11). β-thalassemia disease that results from the interaction of β-thalassemia and Hb E (β26 Glu→ Lys), β°-thalassaemia/Hb E, is a common form of β-thalassemia (12) pose an increasingly important health problem in various parts of Asia (13) due to the high gene frequencies for both Hb E and β-thalassemia (14). β°-thalassemia/Hb E patients also show remarkable variability clinical expression, ranging from asymptomatic or mild clinical symptoms with normal growth development and survival without transfusions, to thalassemia major who have marked anemia, transfusion-dependent, growth retardation, severe bone changes, hepatosplenomegaly, and heavy iron overload (15, 16). In Thailand in particular, the frequency of β°- thalassemia reaches 3-9 % and Hb E approaches 13% of total population and 50% in northeastern population in Thailand (14, 17-19). The number of cases with β°-thalassemia/Hb E is 15 times higher than that of β°-thalassemia homozygotes (20).

High Hb F values are known to reduce the severity of symptoms by minimizing the degree of imbalance of α- and β-globin chains (21, 22). In normal twin study (23), genetic factors alone are accounting for 89% of variation in the number of F cells (FC) which is the number of erythrocytes that contain Hb F. Among genetic factors, *Xmn*I-$^{G}\gamma$ polymorphism is accounting for 13% of total variance, thereby implicating the presence of one or more other genetic loci that contribute to continuous phenotype (quantitative trait loci, QTLs) controlling FC levels in adults (24). Linkage studies in

Indian Kindred have identified two more *trans*-acting quantitative trait loci (QTLs) for FC variance have been mapped, one on chromosome 6q23 (25, 26) and other on chromosome 8q (27, 28). Furthermore, linkage analysis in sib pair with sickle cell disease has localized the QTL to chromosome X22.2. Alleles on X22.2 appeared to be codominant and not affected by X inactivation (29, 30).

Epidemiological studies have shown that a DNA sequence variant (C→T) at position -158 upstream of the $^{G}\gamma$-globin gene, referred to as the *Xmn*I-$^{G}\gamma$ polymorphism increased HbF production in adult life (31, 32). The -158 substitution is near a DNase I hypersensitive site located 50 to 150 bp *5'* of the γ gene Cap sites (33). Perhaps the -158 substitution increases the probability that the chromatin of this region will have an "open" structure, more accessible to DNase *in vitro* and to components of the transcription apparatus of the adult erythroid cell *in vivo*. In addition, the $^{G}\gamma$ -158 variant does not always raise Hb F levels in otherwise healthy individuals, suggesting that the effect of the *Xmn*I-$^{G}\gamma$ site is modulated by the presence of an intermediary factor (27). The linkage study of a large Asian Indian kindred revealed that a genetic interaction between the *Xmn*I-$^{G}\gamma$ site and a locus on chromosome 8q is one of the major factors that influence adult F-cell levels (28). Subsequently, Garner C et al. confirmed these results by replication of linkage to chromosome 8q in a sample of European twin pairs. This result provides strong evidence that a quantitative trait locus exists on chromosome 8q that influences the developmental switch from fetal to adult hemoglobin (27).

Zinc-fingers and homeoboxes 2 (ZHX2) gene or formal known as KIAA0854 is the member of the zinc fingers and homeoboxes gene family that introduced as a transcriptional repressor. This gene is located on chromosome 8q24.13. ZHX2 consists of 837 amino acid residues and contains two Cys2-His2-type zinc-finger motifs and five homeodomains (HDs). ZHX2 not only forms homodimer but also forms heterodimer with ZHX1 and ZHX3, although heterodimerization with ZHX1 is not necessary for repressor activity. Therefore, ZHX2 is likely to possess inherent repressor activity or acquires it through an interaction(s) with transcriptional regulators other than ZHX1. The ZHX2 mRNA is detected among various tissues (34, 35). Further analysis revealed that ZHX2 is a transcriptional repressor that is localized in the nuclei (34). For a number of proteins, nuclear localize signal (NLS) is mapped

to a cluster of basic amino acid residues (36). In contrast, ZHX2 may associate with other molecules for translocation to the nuclei, since the NLS of ZHX2 is mapped to the amino acid sequence between residues 317 and 446 including the proline-rich region but not the basic amino acid region. Besides ZHX2 is a transcriptional repressor, it is the novel candidate genes for globin regulation in erythroid cells (37). The study using siRNA inhibition showed that when ZHX2 levels are reduced, α-fetoprotein (AFP) is re-expressed so ZHX2 was recently identified as one factor involved in postnatal repression of fetal expressing genes including AFP (38) In 2007, Shen H et al. indicated that overexpression of a *ZHX2* transgene led to complete silencing of AFP in the adult liver on a BALB/cJ background, confirming that this gene is responsible for hereditary persistence of the α-fetoprotein (*Afp*) which is transcribed at high levels in the mammalian fetal liver but are rapidly repressed postnatally (39). Down regulation of ZHX2 was recently demonstrated in two HPFH-2 subjects by real-time PCR (37). ZHX2 gene coincides on the QTL on chromosome 8q that has been reported to influence the absolute fetal hemoglobin levels (27).

After all, ZHX2 is a good candidate gene for regulating β-globin gene expression. Nowadays, the molecular biology techniques are effective to use in genetic analysis. In this study, the strategies use to investigate the ZHX2 polymorphism is polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) and detection of the ZHX2 expression during erythroid developmental stages by reverse transcriptase polymerase chain reaction (RT-PCR). Moreover, β°-thalassemia/HbE patients were genotyped for ZHX2 polymorphisms in an attempt to evaluate the involvement of polymorphisms in predisposition to severity. The observation of this polymorphism may have an impact of trend to use for therapy and prevention the patients with thalassemia disease.

Objectives

1. To determine the genotype and allele frequencies of ZHX2 polymorphism in β°-thalassemia/Hb E patients.

2. To study the expression pattern of ZHX2 during erythroid developmental stages.

3. To evaluate the association of ZHX2 polymorphism with percentage of Hb F and degree of severity in β°-thalassemia/Hb E patients.

CHAPTER II
LITERATURE REVIEW

1. Red blood cells

Oxygen is required for aerobic organisms' survival. Through the oxidation of nutrients, organisms can generate energy. As organism size and complexity increased through evolution, not all the cells of an organism were directly exposed to available oxygen. To overcome this problem, vertebrates have evolved a circulatory system that actively delivers oxygen to the cells. Red blood cells (RBCs) in the circulating blood are the carriers, taking up oxygen in the lungs and delivering it to internal tissues to maintain the viability of cells. The actual oxygen-binding molecule in the red blood cells is hemoglobin, an iron-containing protein that gives blood its red color. One milliliter of human blood contains approximately 5 billion RBCs that have a lifetime of about three months. Aging cells are removed from the circulation and destroyed by macrophages in the spleen and liver. To replace them, an adult produces about $2x10^{11}$ RBCs daily (40).

2. Hemoglobin

The human hemoglobin (Hb) molecule is the major component of red blood cells. It consists of two duplicate polypeptide chains in combination with four heme groups, which in turn allows the hemoglobin in erythrocytes to bind oxygen (Figure 2.1) (41). The varying oxygen requirements during embryonic, fetal and adult life are reflected in the synthesis of different structural hemoglobins at each stage of human development. They all have the same general tetrameric structure, however, consisting of two different pairs of globin chains, each attached to one heme molecule. Most normal human Hbs have identical α-chains while the non-α-chain β, γ and δ differ from each other. In human, the genes for α–like globins are clustered on chromosome 16 containing one gene for ζ and two genes for α (α_1 and α_2, the proteins of which are identical). The genes encoding β-like globins are clustered on

chromosome 11 that contains genes for ε, β and δ, one gene for each, and two slightly different genes for γ ($^{G}\gamma$ and $^{A}\gamma$, the proteins of which differ in one amino acid) (42). The majority of normal adult human consists of Hb A ($\alpha_2\beta_2$) and a small amount (2-3%) of Hb A_2 ($\alpha_2\delta_2$) and <1% of Hb F ($\alpha_2\gamma_2$) (1). However, during development different Hbs are produced. The structure of human hemoglobin (Hb) changes during embryonic, fetal and adult life. Adult and fetal hemoglobin have α chains combined with β (Hb A, $\alpha_2\beta_2$), δ (Hb A_2, $\alpha_2\delta_2$) or γ chains (Hb F, $\alpha_2\gamma_2$), whereas in the embryo α-like chains (termed ζ chains) combine with γ (Hb Portland, $\zeta_2\gamma_2$) or ε chains (Hb Gower 1, $\zeta_2\varepsilon_2$), and α and ε chains form Hb Gower 2 ($\alpha_2\varepsilon_2$) (Figure 2.2).

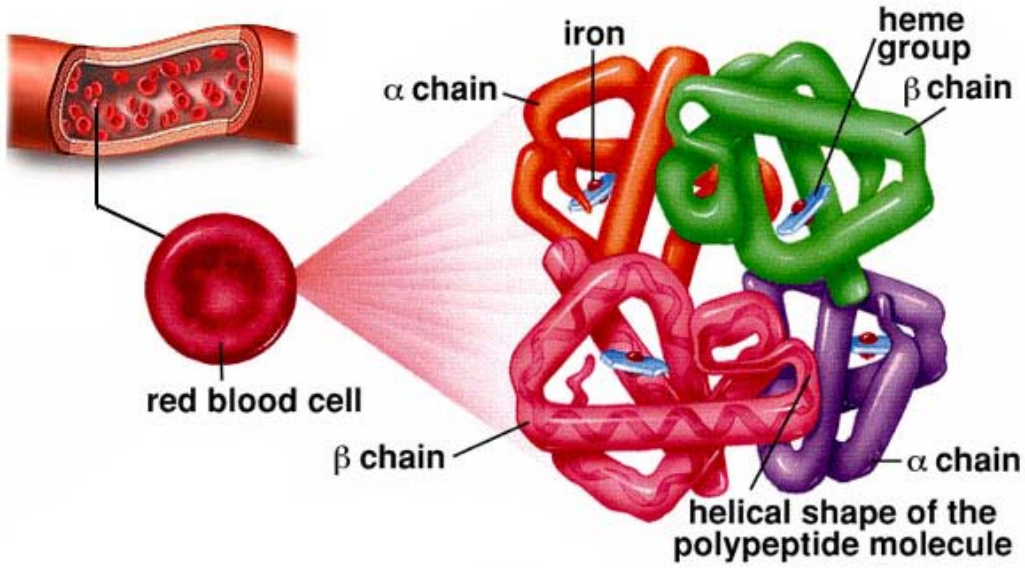


Figure 2.1 Hemoglobin structure (43)

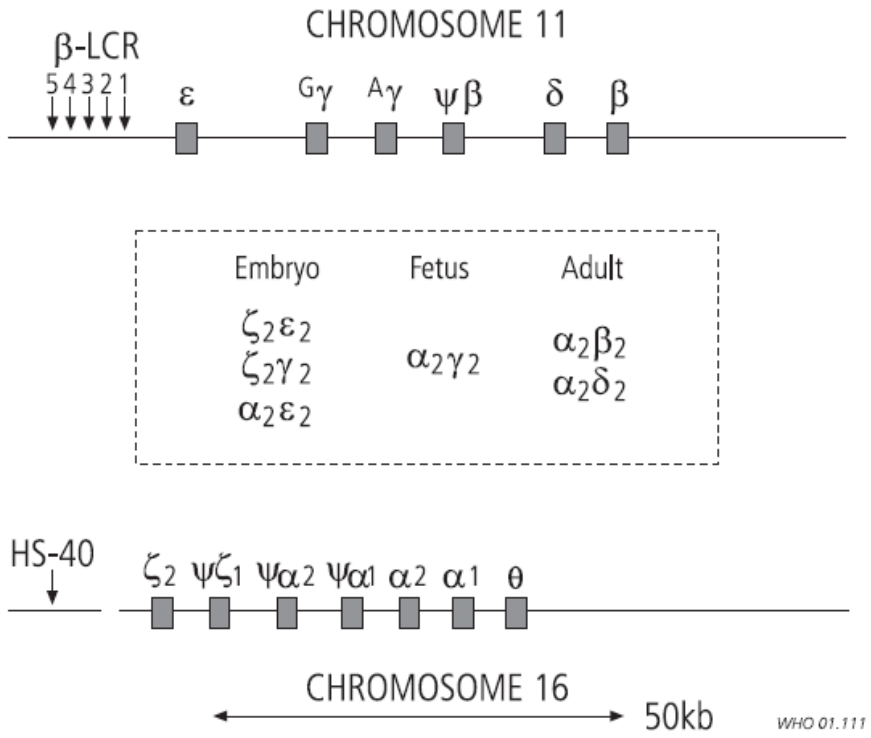


Figure 2.2 The embryonic, fetal and adult hemoglobin coded by different genes during development are shown, together with the gene cluster that regulate their production on chromosome 11 and 16. The β-LCR and HS-40 are the main regulatory regions for these gene cluster (10).

3. Hemoglobin switching

During development of all vertebrate animals, hemoglobin production is characterized by 'switches' in the hemoglobin composition of red cells (42). In human, the α-like globin genes undergo a single switch occurs as development proceeds, from ζ-globin gene expression during primitive erythropoiesis (embryonic) to α-globin gene expression during definitive erythropoiesis (fetal and adult life). While the β-like globin genes switch over one time. First, from embryonic (ε-globin gene) to fetal ($^{G}\gamma$- & $^{A}\gamma$-globin genes) and then around birth, to adults (β-globin gene) globin expression (Figure 2.3). The first position of erythropoiesis occurs in the yolk sac blood islands. Erythroid cells in the yolk sac are formed from embryonic mesoderm at approximately 3 weeks of gestration. These cells produce first hemoglobins in tetramers of two pairs of unalike globin chains. In embryonic state, α-like globin gene called the ζ–globin gene gradually decreasingly expresses and α-globin expression continuously increases over the next few weeks. For the β-like globin gene, the ε-globin gene is expressed during the first 6 weeks of gestation in primitive, nucleated erythroid cells of the yolk sac, and γ-globin gene expression gently enlarge while the β-globin genes are silent (embryonic or primitive erythropoiesis). So, the productions of embryonic Hbs are Gower 1 ($\zeta_2\varepsilon_2$), Gower 2 ($\alpha_2\varepsilon_2$) and Hb Portland ($\zeta_2\gamma_2$), respectively. During the first switch, approximately 5 weeks of development, $^{G}\gamma$- and $^{A}\gamma$-globin gene expression is activated in the definitive hematopoietic cells of the fetal liver and spleen (fetal definitive erythropoiesis); the ε-globin gene is concomitantly silenced. Then these organs become the major site of erythropoiesis. The majority of hemoglobin produced at this stage of development is fetal hemoglobin, Hb F ($\alpha_2\gamma_2$) which takes the place of embryonic hemoglobin. Eventually, the second switch which occurs shortly after birth, the β-globin gene and, to a lesser extent, the δ-globin gene is activated in the bone marrow (adult definitive erythropoiesis). This organ becomes the main position of erythropoiesis and this is a final switch in hemoglobin synthesis. When the adult β- and δ- globin gene is expressed, the γ-globin genes are reciprocally silenced (7) and ensuing in Hb F being replaced Hbs A and A_2, (Figure 2.3). Hence, the adult Hb is Hb A: $\alpha_2\beta_2$ (>98%, adult major), and Hb A_2: $\alpha_2\delta_2$ (<2%, adult minor) (10, 42, 44-46).

However, regulation of each α- or β- locus present on different chromosomes is independent, as the level of expression of one has no effect on the other. In addition, the γ-globin gene still coordinately expressed throughout development. The Hb F in these cells is estimated about 1% of total hemoglobin in adult blood (47).

At birth β-chain synthesis increases sharply whereas γ-chain production falls steadily. It takes about 1 year to reach the level that is characteristic of adult red cell (~1%). Hb F is believed to confine in few erythrocytes call “F cell”. Approximately 3 to 7 percent of erythrocytes are F cells. Both the number of F cells and the amount of Hb F in each F cell may be increased in various acquired and genetic conditions characterized by elevated Hb F levels (48).

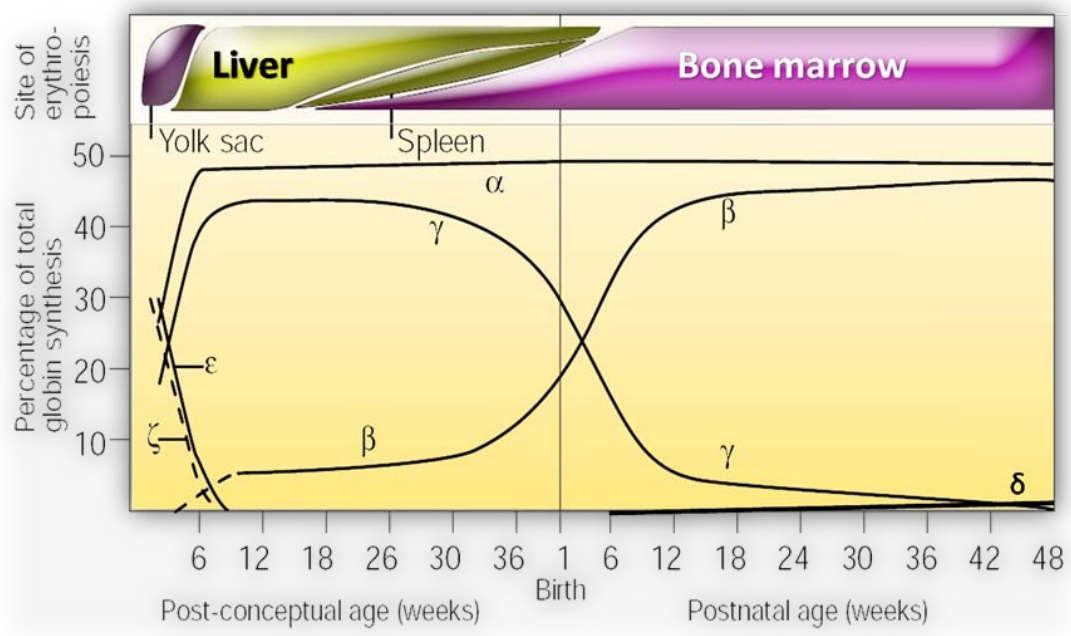


Figure 2.3 Developmental changes in human globin chain production, sites of erythropoiesis, and red cell morphology (49).

4. Globin genes

The human Hb genes exist two separate clusters of related multigene families, a frequent type of organization of mammalian genes. The α-gene cluster is located on the short arm of chromosome 16 (16p13.3) over a 30 kb region while the ε, γ, δ, β family is situated on short arm of chromosome 11 (11p15.5) across a 70 kb region. The α-cluster from 5' to 3' (Figure 2.2) include embryonic ζ gene, the non functional $\psi\zeta_1$, $\psi\alpha_2$ and $\psi\alpha_1$ pseudogenes, the duplicated adult α genes α_2 and α_1 , and θ gene of undetermined function. While the location of the various genes on the β-cluster (Figure 2.2) are, the embryonic ε-gene, two fetal γ genes, $^{G}\gamma$- and $^{A}\gamma$- (Glycine or Alanine at position 136), ψβ gene, and the adult δ and β genes (8). The 5’ to 3’ arrangement of these genes is in order of ontogenetic expression during development. Pseudogenes have DNA sequences that resemble those of their homologues. However, various mutational alterations have inactivated transcription, so that there is no functional expression. The δ-gene, whose gene product comprises only 2-3 % of total non-α-chains, can be seen as a gene in transition to becoming a pseudogene.

Each globin gene of the α- or β-cluster has three exons and two introns (50). The α-like globin chain has 141 amino acids while β-like globin chain has 146 amino acids. Distinct functions have been correlated with each of the exons. Exon 2 encodes for the segment involved in heme binding and α-β dimer formation whereas exon 3 encodes for many of the amino acids involved in globin subunit interactions required for cooperativity of the Hb tetramer in binding oxygen. The intron lengths are variable but the largest intron 1 of the ζ-gene is only 1265 bp (51, 52).

4.1 Molecular Regulation of Globin Genes

The molecular aspects of the expression of individual globin genes involve *cis*–acting and *trans*-acting regulatory mechanisms which have been extensively studied in cell lines and transgenic mice (45). The factors that regulate globin gene are both tissue restricted and ubiquitous with respect to their pattern of expression. At the murine and human β-globin loci, gene transcription is controlled by the complex interactions between:

1. *cis*-acting sequences: besides the local *cis*-acting sequences, like the promoter sequences, the positive enhancer and negative regulatory elements, remote *cis*-acting sequences are involved in regulation of globin gene expression, the Locus Control Region in β-globin gene cluster (βLCR) and downstream globin gene sequences embedded with histones in nucleosomes, in chromatin (Figure 2.4), moreover the hyper sensitive site at position -40 (HS-40) in the α-globin cluster also (53-56), and

2. *trans*-acting factors: transcription factors and chromatin remodeling activities (57-59).

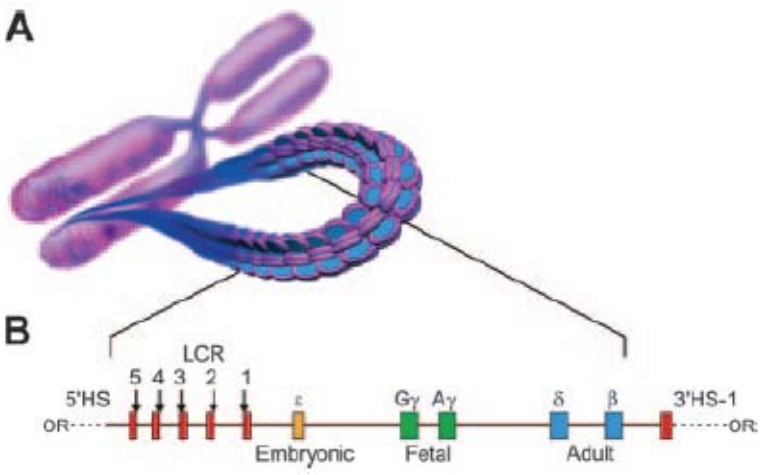


Figure 2.4 The human β-globin locus. (A) The β-globin locus on chromosome 11 embedded in chromatin is shown. (B) A linear map with the globin LCR and its hypersensitive (HS) sites is indicated by the vertical arrows. The structural ε–, γ–, δ–, and β-globin genes as well as the locations of the olfactory receptor (OR) genes are shown (60).

4.2 Regulatory Elements of the α-Globin Gene Cluster

Regulation of expression in the human α-globin locus is dependent, beside the local-acting sequences, on an upstream regulatory element HS-40, which is an erythroid specific DNase 1 hypersensitive site located 40 Kb upstream of the ζ-globin gene cap-site (56). Sequence analysis of the 350 bp element HS-40, denoted as the α Major Regulatory Element (αMRE) shows the presence of regulatory protein binding sites (53) Even though HS-40 has some ability to form open chromatin in transgenic mice, but it's main function is probably to activate and enhance expression from preformed complexes at the α-globin promoters in a constitutively "open" chromatin environment (59, 61).

4.3 Regulatory Elements of the β-Globin Gene Cluster

The promoter, an intragenic enhancer and a downstream enhancer act in concert in controlling β-globin gene expression (62). The promoter exhibits tissue specific expression, it contains elements that interact with LCR located within the erythroid specific DNase 1 hypersensitive sites HS 1, 2, 3, 4 and 5 which are distributed over a region 4-20 Kb upstream of the ε-gene and others that contribute independently to increased gene expression with erythroid maturation. The enhancers also have the potential to establish tissue and developmental stage specificity of β-globin gene expression (63).

In both non erythroid and erythroid cells the TATA, CCAAT and CACC boxes are required for high expression of the β-globin gene promoter (64-66) The CACC box is duplicated, but the proximal CACC box appears to be the most functionally important. Mutations in the TATA and proximal CACC boxes cause the β^+-thal phenotype. There is a 10 bp direct repeated sequence between the CCAAT and TATA boxes called the direct repeat element (DRE). Mutations in one of the β-DRE repeats have no effect on promoter function, but corresponding mutations in both repeats do reduce promoter function. Furthermore, there is an enhancer within the β-globin gene near the junction of intron 2 and exon 3. Another enhancer is located just downstream from the β-globin gene polyadenylation site. The intragenic and downstream regions contain three and four binding motifs for GATA-1 respectively, perhaps accounting for their enhancing activity. The intragenic enhancer is

indispensable for high level β-globin gene expression because of an influence on the efficiency of polyadenylation of the β-globin gene transcript (67).

Expression of the adult β-globin gene is controlled by a process entirely different from that turns off the embryonic globin genes. A competitive model for the γ to β switch has been proposed. The model states that the probability of LCR interaction with either the γ or the β promoter is primarily determined by the *trans*-acting environment. If the gene that successfully competes for the LCR is expressed then the unsuccessful gene is switched off. This model is expanded to explain the ε to γ to β switches. In the embryonic stage, the LCR interacts with the ε-globin gene. In the fetus, ε is silenced, and the LCR interacts with the $^{G}\gamma$ and $^{A}\gamma$ genes. In the adult, the γ gene is silenced, and the LCR now recognizes β the last available gene such a competition model was originally proposed to explain switching in the chicken globin system by Choi and Engel in 1988 (68). Experiment using combinations of tandem γ or β globin genes suggested that gene order (relative to LCR) contributes relatively less to developmental regulation than does the *trans*-acting environment (67).

4.3.1 β-globin locus control region

The β-globin locus control region (βLCR) upstream (5') of the globin structural genes is the major structural component of the murine and human β-globin loci (60). It was not only to be responsible for stimulating high-level transcription of the β-globin genes but also required for the formation of the active β-globin gene domain as defined by propensity to DNase I digestion (69). The murine and human βLCRs contain 5 critical DNAase 1 hypersensitive (HS) sites, namely HS1-5 (Figure 2.4), that are formed in regions devoid of nucleosomes and are more accessible in comparison with other regions of chromatin to interactions with transcription factors and downstream gene sequences (69, 70). Another HS site, 3' HS1, is 3' to the β-globin structural globin gene, and olfactory receptor (OR) sequences are at both the 5' and 3' borders of the β-globin locus (Figure 2.4) (69, 71).

High level globin gene expression at all stages of development is achieved at least in part by the locus control region (LCR). The LCR acts as an enhancer, resides 6- 20 kb upstream of the ε-globin gene (72).

4.4 Erythroid cell-specific Trans-acting Factor

There are several important transcription factors for erythroid cell development and function which play critical roles in murine globin gene expression. GATA-1 is a member of zing finger family (73, 74) which binds to βLCR sequences (75, 76) and it is needed for early stage of erythroid cell differentiation (77). Mice deficient in GATA-1 cannot produce mature erythroid cells (78-80) and erythropoiesis can be normalized by restoring GATA-1 function (78, 79). Erythroid Krupple-like factor (EKLF) is necessary for activation of the adult β-globin gene. It is able to interact with the sequence CCA CAC CCT, an essential element of the β-globin promoter (βLCR) through its zing finger motifs (81-83). Mice deficient in EKLF die in utero with a disease resembling severe β–thalassemia (84). The transcription factor NF-E2 is a member of the basic-leucine zipper (BZIP) family (85) which is important in its function at the human β-globin locus although mice devoid of in NF-E2 have only mild erythroid defects (86, 87). NF-E4 is a critical transcription factor in chicken globin switching (88). More recently, a human homolog of chicken NF-E4, NF-E4p22, has been shown to be active in human fetal globin gene activation (89, 90).

5. Genetic polymorphisms

The crucial keys that underly human phenotypic differences are genetic and environmental factors (91). There are numerous variations in the DNA sequence at many points in every part of the genome when the genomic sequence on similar chromosome of any two individual is aligned and compared. These genetic variations have many forms. The simplest type results from a single base mutation which replaces one nucleotide by another, called single nucleotide polymorphism (SNP). Many other types result from the insertion or deletion polymorphism. Insertion/deletion tends to occur in repetitive sequence element, where the repeated nucleotide patterns or variable number of tandem repeat polymorphisms (VNTRs) expand or contact as a result from insertion or deletion (92, 93).

DNA sequence variations are sometimes described as mutations and sometimes as polymorphisms. A mutation is defined as any changes in a DNA sequence resulting

in functional deviation in comparison to the normal allele. In contrast, a polymorphism is a DNA sequence variation that is common in the population (1 percent). That is, to be considered as a polymorphism, the variation must have a frequency of 1 percent or greater in the given population (94). Generally sequence variants that directly and overly cause human diseases are rare in the population because they reduce fitness. However, if a rare allele in one population confers an advantage and increases in frequency, it could turn out to be a polymorphism in another. For example the allele of thalassemia, which is a rare sequence variant of the globin gene in Caucasian populations, is polymorphic in Southeast Asian populations because it renders resistance against the blood-borne parasite that causes malaria (95).

Although more than 99 percent of human DNA sequences are identical across the population, variations in DNA sequence contribute a major impact on how humans respond to disease, bacteria, viruses, toxins, chemicals, drugs, and other therapies (96). Clearly, many clinical phenotypes do have a considerable genetic component. The existence of a specific variation allele can be implicated as a causative factor in human genetic disorders. Therefore, screening for such an allele in an individual enable the detection of a genetic predisposition to disease. The sequence variants can be imagined to merely modifying risk for some phenotypes. Many may be found within genes, but may influence characteristics such as height and hair color rather than those of medical importance while some does contribute to disease susceptibility and can also influence drug responses. However, many polymorphisms are found outside of genes and are completely neutral in effect (93).

5.1 Single Nucleotide Polymorphisms (SNPs)

SNPs are genetic variation that occurs when a single nucleotide: adenine (A), thymine (T), cytosine (C) or guanine (G) in the genome sequence is transformed. SNPs are mostly biallelic polymorphisms, that is, the nucleotide identified at these polymorphic positions is generally constrained to one of two possibilities in human, rather than the three or four nucleotide possibilities that could occur, in principle (94). SNPs are the most common form of genetic variation accounting for 90 percent of all human genetic variations. Their density is estimated that occur about every 1,000 bases in the entire human DNA sequences, leading to a total of several million SNPs in human population (97).

SNPs can occur in both coding and non-coding regions of the genome. However, because only about 3 to 5 percent of a human DNA sequence codes for the production of proteins, changes in non-coding sequence is more common. SNPs found within a coding sequence are of particular interest to researchers because they are more likely to alter the biological function of a protein. Occasionally, a SNP may actually cause a disease, and, therefore, can be used to search for and isolate the disease-causing gene. However, most SNPs are not responsible for a disease state. Instead, they serve as biological markers for pinpointing a disease on the human genome map, because they are usually located close to a gene underlying a certain disease. Many SNPs have no effect on cell function, but it is believed that they could predispose people to disease or influence their response to drug. However, SNPs are not absolute indicators of disease development. A good example is the genes associated with the late onset Alzheimer's disease, apolipoprotein E or ApoE (98). This gene contains two SNPs that result in three possible alleles: ε2, ε3 and ε4. Each allele differs by one DNA base, and the protein product of each gene differs by one amino acid. Recent study has shown that an individual who inherits at least one ε4 allele will have a greater chance of getting Alzheimer's disease (98). Apparently, one amino acid change in the ε4 protein alters its structure and function enough to make disease development more likely. Interestingly, the ε2 allele, on the other hand, seems to indicate that an individual is less developing Alzheimer's disease. Genetic studies of other complex disorders such as heart disease, diabetes, or cancer are complicated by numerous factors due to multiple gene interactions (perhaps influenced by environmental factors), with variable clinical expression of the disease phenotype, but a good SNP map would make such experiments more feasible (99).

The significance of SNPs in genetic studies stems from at least three different considerations. First, SNPs can be used to construct the history of genome. This is due to their abundance and most are inherited from one generation to the next, evolutionary stable, making them easier to follow in population studies. Studying the frequency and distribution of SNPs can lead to information on the evolution of the species. Second, SNPs can be directly responsible for genetic diseases since they may alter the sequence of gene or of a regulatory region. Finally, SNPs can be used as markers to build the high-density genetic maps needed to perform association studies.

A map of 100,000 or more SNPs has been proposed as an ultimate goal to enable effective genetic-mapping studies in large populations (100). The conclusion of such studies is that the polymorphism being tested either affects risk of disease directly or is a marker for some nearby genetic variant (101).

6. Hemoglobinopathies

Inherited abnormalities of the hemoglobin tetramer may be divided into two categories: those that are characterized by structural anomalies of the hemoglobin chains, and others that result from an array of molecular defects that either reduce or completely abolish the synthesis of one or more of the polypeptide chains of the hemoglobin molecule (102). The term 'hemoglobinopathy' refers to the former disorders, whereas the latter defines the term 'thalassemia' (103).

To date, nearly 700 mutant alleles have been characterized (http://globin.cse.psu.edu) (13). These structural alterations may include amino acid substitutions, deletions, and insertions. Many of these mutations are functionally normal and, therefore considered as clinically silent.

6.1 Thalassemia

The name 'thalassemia' is derived from the Greek word 'thalassa' ('sea'), because it is well recognized that there is a high incidence of people suffering from this hemoglobinopathy near the Mediterranean Sea. The thalassemia syndromes are a diverse group of inherited disorders that can be characterized according to their insufficient synthesis or absent production of one or more of the globin chains. They are classified into α, β, γ, δβ, δ and εγδβ thalassemias depending on the globin chains which are affected. Defective synthesis of one of the globin chain leads to the imbalance of α/β chains production and consequent precipitation of excessive unmatched normal synthesized globin chains. The excess globin chains precipitate in the bone marrow, causing red blood cell membrane rigidity, following ineffective erythropoiesis and short red cell survival results in anemia (104).

Thalassemias can be classified at three levels.

1. Clinical
2. Genetic
3. Molecular

Clinical

The thalassemias are divided according to the severity of the illness.

a) Thal-major: Patients are severe and transfusion dependent, results either from the compound heterozygous state of two different β-globin mutations or homozygous state. The subsistence without blood transfusion has to live with troubled deformities. They feature all signs and symptoms associated with severe anemia such as growth retardation, hepatosplenomegaly and thalassemic faces. The disease is usually fatal early in life (105).

b) Thal-intermedia: Thalassemia intermediate patients are not as severe as the major forms and associated with a more severe degree of anemia in the intermediate. They are able to survive without blood transfusion. The differences can be distinguished by the degree of anemia endurance and the threshold of the physician to transfuse patients with thalassemia in association with intercurrent illness. Growth and development during childhood is relatively uncompromised, pubescence takes place normally and fertility is preserved (106, 107).

c) Thal-minor: A person has symptomless

d) Silent carrier state: Some forms of thal-trait, which are clinically and hematologically completely silent.

Genetic

The thalassemias are classified according to their genetic basis by describing the globin chain that is absent (0) or produced at the reduced rate (+) and in turn reflects the structure of the globin genes involved in their synthesis. Accordingly they are α, β, γ, δβ, δ and εγδβ varieties, depending on which chain or chains are absent or synthesized at reduced rate.

Molecular

Thalassemias can be more accurately classified and it is now possible to elucidate the genotype of a patient with the clinical feature of β-thal major according to the particular mutations at the homozygous or compound heterozygote condition like βIVS-1-5 G→C/βIVS-1-5 G→C or βIVS-1-5 G→C/β8/9 (+G) (108, 109).

Anyhow the most two important subdivisions of this disorder, α- and β-thalassemia, resulting from the defective synthesis of the α- and β–globin chains of hemoglobin, respectively.

6.1.1 α-Thalassemia

α-Thalassemia was first identified in 1925 when Detroit pediatrician Dr. Thomas Cooley described a syndrome he observed in Italian children characterized by extreme anemia, splenomegaly, and bone deformities (110).

α-Thalassemias which result in reduced α-chain synthesis are common in Southeast Asia and Southern China. Since the α genes are duplicated, the genetics of α-thalassemia is more complicated than that of β-thalassemia. The genetic makeup of normal individuals can be written αα/αα. Loss of both α genes on a chromosome is called α° thalassemia (α-thalassemia 1), and is represented --/αα. Loss of one of the linked pair of a globin gene is called α^+ thalassemia (α-thalassemia 2), -α/αα. Usually these α gene is lost by deletion, though sometimes they are inactivated by a point mutation, as is the case in the β thalassemias (6).

Heterozygotes for α^+- and α°-thalassemia show minor hematological abnormalities but are clinically unaffected while compound heterozygotes for α°- and α^+-thalassemia, --/-α, have Hb H disease, a moderately severe hemolytic anaemia. Homozygotes for α°-thalassemia have a lethal condition known as the Hb Bart's hydrops fetalis syndrome. Furthermore, the unequal crossing over on α-globin cluster results in the triplicated α-thalassemia (ααα/) with increasing α-globin chain synthesis.

Molecular analysis has shown that α-thalassemia is caused by a large variety of genetic defects. α^+-Thalassemia can result either from deletions which remove one of the α genes ($-\alpha^{3.7}$ and $-\alpha^{4.2}$, where the superscript indicates the size of the deletion in kb) or a number of non-deletion defects (written as $\alpha\alpha^T$). α°-Thalassemia results from deletions which involve both a genes ($--^{SEA}$, $--^{MED}$, $(\alpha)^{20.5}$ and $(\alpha)^{5.2}$ where the superscripts SEA and MED refer to previously characterised α° defects from Southeast Asia and Mediterranean subjects, and the superscripts 20.5 and 5.2 refer to the size of the deletions in two less common α° defects) (5).

6.1.2 β-thalassemia

The β-thalassemia refer to a group of inherited hemoglobin disorders which are characterized by a reduced synthesis (β^+-thalassemia) or absence (β°-thalassemia) of β-globin chain production (13). This leads to an imbalanced ratio of α/non α-globin synthesis, which is the major factor in determining the severity of the disease in the β-thalassemia syndromes (111). Over 200 different mutations of the β-globin genes have been found in patients with β-thalassemia (7, 10, 112). In Thailand, more than 20 different mutations have been described with, the four most common mutations are: 4 bp deletion (-TCTT) in codon 41-42; C→T, amber mutation in codon 17; C→T in position 654 of the IVS 2 and A→G at position -28 in ATA box (113). Homozygous β°-thalassemia has a clinical feature of thalassemia major which is severe form of β-thalassemia and most of the patients die in the pediatric age group. Whereas β^+-thalassemia is a milder form of β-thalassemia since the patients can produce certain amount of Hb A.

6.2 Hemoglobin variant

Hemoglobin variants are abnormal hemoglobin resulting from single amino acid substitution in one of the globin chains. These changes may affect the structure of the hemoglobin, its behavior, its production rate, and/or its stability. Usually there are no alterations in amount of globin chain production in abnormal hemoglobin. The most common abnormal hemoglobins which also have thalassemia-like defects are Hb E and Hb Constant Spring (Hb CS).

6.2.1 Hemoglobin E

Hemoglobin E, the commonest structural hemoglobin variant in Thailand and wouldwide, is innocuous in its heterozygous and homozygous states. This variant results from a mutation in the hemoglobin beta chain which normal glutamic acid residue is substituted by lysine in the 26^{th} amino acid of β chain (β26 Glu → Lys). It is synthesized at a reduced rate and can interact with β°-thalassemia to produce a condition called β°-thalassemia/Hb E, the common form of β-thalassemia presenting an increasingly important health problem in various parts of Asia (13) due to the high gene frequencies for both Hb E and β-thalassemia (14). β°-thalassemia/Hb E patients can suffer from a moderate to severe anemia and require regular blood transfusion (18).

6.2.2 Hemoglobin Constant Spring (Hb CS)

Hemoglobin Constant Spring is a variant in which a mutation in the alpha globin gene produces an alpha globin chain that is abnormally long. It occurs from the termination codon (UAA→ CAA, Gln) of the α_2-globin gene. This abnormality results in a translation of a longer α-globin chain with an extra 31 amino acids (114). The quantity of hemoglobin in the cells is low for two reasons. First, the messenger RNA for hemoglobin Constant Spring is unstable (115). Some is degraded prior to protein synthesis. Second, the Constant Spring alpha chain protein is itself unstable. The result is a thalassemic phenotype (115).

In Southeast (SE) Asia, α-thalassemia , β-thalassemia, hemoglobin (Hb) E, Hb Constant Spring (Hb CS) are common (111). In Thailand in particular, the frequency of α-thalassemia reaches 25%, 3-9 % for β°- thalassemia, Hb E approaches 60% in frequency in many regions of Thailand, Laos and Cambodia. Hb CS is found between 1%–10% of the population in these areas (14, 17, 18).

7. β°-Thalassemia/Hb E disease

β°-Thalassemia/Hb E disease is the most common form of β-thalassemia in many Asian countries (116, 117) because there are higher chances that an individual could carry on genes for both thalassemia and structural hemoglobin variants condition. It is a compound heterozygous state of β°-thalassemia and Hb E causes β°-thalassemia/Hb E disease which generally has a wide clinical spectrum of the disease ranging from mild to severe condition. In Thailand, approximately 3,000 children are born with this condition each year, and there are some 100,000 patients with the average life expectancy is about 30 years. It accounts for well over 50% of cases of severe β-thalassemia in Indonesia and Bangladesh and is also very common in Vietnam, Cambodia, Laos, and Malaysia. It is found occurs frequently on the eastern side of Indian subcontinent, including Sri Lanka (117, 118). Nowadays we know few factors that affect the severity is Hb F level. However we still do not know all of them.

8. Thalassemia with increased HbF levels

There are many types of thalassemia that increase Hb F production during adult life. These mutations are clinically relevant, since increased synthesis of Hb F in individuals with sickle cell anemia or β-thalassemia reduces disease severity. These mutations are either deletion type, that is a portion of the β-globin gene cluster is deleted, or non-deletion type in which point mutations are within or outside the cluster. The major clinical significance of mutations that increase Hb F is found in their interactions either the β-thalassemia or sickle cell anemia. Compound heterozygotes have a much milder clinical syndromes like individuals with $^{G}\gamma^{A}\gamma$ HPFH and either a β-thalassemia or an Hb S allele are asymptomatic and detected by chance during population screening (45)

Thalassemia syndromes individuals who are homozygous for β-thalassemia have striking increase in the proportion of Hb F that is of diagnostic significance. The proportion of Hb F in the blood may range from 10 to 98%, depending on whether the patient has inherited thalassemia mutations of the β^{+} or β° variety. Interaction of the β-thal genes with mutations that increase Hb F in compound heterozygote may give rise to thalassemia syndromes of moderate severity (119). Hb F and F cell number are also moderately increased in abou t 3 0 % of β-thal heterozygotes, although the mechanism for this increase is obviously unknown.

9. Factors affecting Hb F level of β-thalassemia.

High Hb F values are known to reduce the severity of symptoms by altering the degree of imbalance of α and β-globin chain (21, 22). Genetic factors involved in the control of FC production have been reported including co-inheritance of α-thalassemia, *Xmn*I polymorphism on chromosome 11 (24), QTL on chromosome 6q23 (25), the X-linked locus at Xp22.2-p22.3 (29, 30), and QTL on chromosome 8q (27, 28)

9.1 *Xmn*I polymorphism

A common variant (C→T) at position -158 upstream of the $^{G}\gamma$-globin gene have been associated with increased HbF levels in normal individuals, in β-thalassemia and in sickle cell anemia (120, 121), namely the "*Xmn*I-$^{G}\gamma$ polymorphism." The *Xmn*I-$^{G}\gamma$ site is believed to be involved in the expression of the

$^{G}\gamma$-globin gene through interaction with transcription factors and polymorphisms in the transcription factors could influence fetal Hb production (28). Were the molecular events in hemoglobin switching better understood and Hb F could be more fully reactivated in adult cells and might lead to a cure for these disorders.

9.2 Quantitative trait locus on chromosome 6q23

A quantitative trait locus (QTL) controlling HbF levels has previously been mapped to chromosome 6q23 in an Asian-Indian kindred with β-thalassemia and heterocellular hereditary persistence of fetal hemoglobin (HPFH) (30). This interval contains three (*HBS1L*, *MYB*, and *AHI1*) protein-coding genes which expressed in erythroid progenitor cells (122). Common sequence variants situated between the *HBS1L* and *MYB* genes on chromosome 6q23.3 that may influence the proportion of F cells. *cMYB* and *HBS1L* have been demonstrated to simultaneously transcriptional down-regulated in individuals with elevated Hb F levels (123, 124).

9.3 The X-linked locus at Xp22.2-p22.3

In 1992, G.J. Dover et al. observed that F-cell levels were significantly higher in nonanemic females than males (mean ± SD, 3.8% ± 3.2% v 2.7% ± 2.3%) so they tested the hypothesis that F-cell production in both normal and anemic SS individuals was controlled by an X-linked gene (29). Furthermore, Miyoshi et a1. reported evidence suggesting that F-cell production in heterocellular HPFH within the Japanese population is controlled by an X-linked locus (125). A dominant locus on the X chromosome has been shown to influence F-cell levels in nonanemic Japanese blood donors (126).

9.4 Quantitative trait locus on chromosome 8q

Epidemiological studies have shown that a DNA sequence variant (C→T) at position -158 upstream of the $^{G}\gamma$-globin gene, referred to as the *Xmn*I-$^{G}\gamma$ polymorphism increased Hb F production in adult life (31, 32). On the other hand, the $^{G}\gamma$-158 variant does not always raise Hb F levels in otherwise healthy individuals, suggesting that the effect of the *Xmn*I-$^{G}\gamma$ site is modulated by the presence of an intermediary factor (27). The genome-wide linkage study of a large Asian Indian kindred revealed that a genetic interaction between the *Xmn*I-$^{G}\gamma$ site and a locus on chromosome 8q is one of the major factors that influence adult F-cell levels (28). After that, Garner C et al. confirmed these results by replication of linkage to

chromosome 8q in a sample of European twin pairs. This result provides strong evidence that a quantitative trait locus exists on chromosome 8q that influences the developmental switch from fetal to adult hemoglobin.

Based on the finding that there is an interaction between the *Xmn*I-$^{G}\gamma$ site and a QTL on chromosome 8q in influencing the production of Hb F. Zinc-fingers and homeoboxes 2 gene (ZHX2), a transcription repressor, is a one of novel candidate genes that regulates globin gene in erythroid cells and also found downregulated of ZHX2 in HPFH and δβ-thalassemia (37). It plays an important role in the postnatal repression of AFP and Glypican 3 (Gpc3) gene which are fetal genes (38, 39, 127, 128). Therefore, the polymorphic variations are displayed within the ZHX2 genes which have an important field of investigation in this study.

10. ZHX2

10.1 History

The zinc-fingers and homeoboxes 2 was initially discovered in a fishing expedition for novel transcriptional repressor. It is an outgrowth from the analysis of molecular mechanism by which ZHX1 functions as a transcriptional factor. The researcher firstly demonstrated the interactions of human ZHX1 with other transcription factors by using the yeast two-hybrid system. The fusion protein between whole coding sequence of the human ZHX1 and the DBD of the GAL4 transcription factor was conducted as a bait to screen a rat liver cDNA library. Indeed, some 15 million independent clones were screened and 16 showed reproducible *HIS3*-, *ADE2*-, *MEL1*-positive properties, and *β*-galactosidase activity respectively. After a determination of their nucleotide sequences, they were compared with the Genbank database using the BLAST search program. One of these clones exhibited a similarity to the nucleotide sequence of the human KIAA0854 protein. Very noticeably, the deduced amino acid sequence of the KIAA0854 protein has an open reading frame of 837 amino acid residues and contains two Cys2-His2-type Znf motifs and five HDs as well as ZHX1, indicating that it also belongs to the Znf class of the homeobox protein superfamily (129). Thereafter, they refer to the KIAA0854 as ZHX2 and report on its further characterization.

The cDNA encodes a protein of approximately 92 kDa and a pI of 6.42 (34). Initially the mRNA expression was found in various tissues at different levels e.g. bone marrow and spleen, the two hematopoietic organs of human. Further studies showed that the expression of ZHX2 may implicate in the globin gene regulation in erythroid cells (37).

10.2 Zinc fingers and homeoboxes gene family

Most transcription factors are grouped into families on the basis of common DNA- or protein-binding domains. ZHX2 has been classified to the zinc fingers and homeoboxes genes family of transcription factors. These transcription factors bind with the A subunit of the ubiquitous transcription factor nuclear factor-Y (34, 130) which acts through the Y box sequence, an inverted CCAAT box, 5'-ATTGG-3' (131), that are present in promoters, enhancers and locus control regions of numerous genes. Besides that, they can form homodimers with themselves as well (34, 132, 133). A combination of two conserved Cys_2His_2-type zinc finger motifs found close to the C-terminus and five homeodomains (HDs) from the N-terminus, defines the transcription factors of the zinc-fingers and homeoboxes (ZHX) family. The name of this family of proteins derives from the observation that a similar arrangement of zinc fingers and homeoboxes was found in the *Drosophila* regulatory gene, called homeotic genes (134). Of which mutations cause developmental abnormalities. In human genome, 3 genes have been classified as belonging to this family of transcription factors, ZHX1, ZHX2 and ZHX3. All these proteins function as transcriptional repressors and are localized in the nuclei of cells (34, 133, 135).

10.3 Domains in ZHX2

The gene coding for ZHX2 or KIAA0854 in human is located on chromosome 8q24.13 (Figure 2.5) in the position 123863082-124055936. Four exons code for 837 amino acid protein of approximately 92 kDa. The position of exon1 is 123863082-123863366, exon 2 is 123944897-123944959, exon 3 is 124032713-124035449, exon 4 is 124054663- 124055936 but exon3 is the only one coding exon. The coding position in this exon starts at 124032932 to 124035445 (Table 2.1). In the literature, the amino acids are commonly numbered from the ATG site as in Figure 2.6. zhx2 contains two Cys2-His2-type zinc-finger motifs and five homeodomains

(HDs) (Figure 2.6). In addition to two Znf motifs and five HDs, zhx2 contains a unique proline-rich region between HD1 and HD2.

The genomic structure is highly conserved between rat, mouse and human (34, 132). The DNA binding domain, consisting of two Cys2His2 zinc fingers, is the most conserved part. The homology of the amino acid sequence between rat and human zhx2 and between mouse and human zhx2 domain among the genes are about 93.3% and 87%, respectively. Consequently, the proteins are believed to recognize the same DNA sequences with very similar affinity. In human, the zinc fingers span the region from residues 78-101 and 110-139.

Since zhx2 is a nuclear protein, it has to be imported from the cytosol to the nucleus. Precisely residues 317-446 contain one Nuclear Localize Signals (NLS) (34). For a number of proteins, NLS is mapped to a cluster of basic amino acid residues (36). In contrast, the NLS of zhx2 is mapped to the amino acid sequence between residues 317 and 446 including the proline-rich region but not the basic amino acid region so zhx2 may associate with other molecules for nuclear translocation. However, the most efficient nuclear localisation is found with the entire domain.

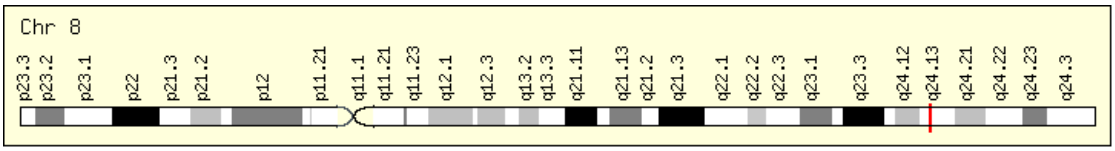


Figure 2.5 Location of ZHX2 on chromosome 8

http://www.genecards.org/cgi-bin/carddisp.pl?gene=ZHX2

Synonyms	Id	Region	Chr.	start	end	length
AFR1	OMIM : 609185	Transcription	8	123863082	124055936	192,855
KIAA0854	Entrez Gene : 22882	Coding seq.	8	124032932	124035445	2,514
RAF	SWISSPROT : Q9Y6X8	Exon 1	8	123863082	123863366	285
	Ensembl : ENSG00000178764	Exon 2	8	123944897	123944959	63
	MRNAACC : NM_014943	Exon3	8	124032713	124035449	2,737
		Exon4	8	124054663	124055936	1,274

Table 2.1 Information of ZHX2 gene

http://www.ensembl.org/Homo_sapiens/exonview?db=core;transcript=ENST00000314393

10.4 ZHX2 functions

Zhx2 not only forms homodimer but also forms heterodimer with Zhx1 and Zhx3, although heterodimerization with Zhx1 is not necessary for repressor activity. Therefore, Zhx2 is likely to possess inherent repressor activity or acquires it through an interaction(s) with transcriptional regulators other than Zhx1. The dimerization domain with that of zhx1 is confined in the region containing HD1, between amino acid residues 195 and 358. The domain that interacts with NF-YA is the HD1 to HD2 region, via the region between 263 and 497 of Zhx2. Therefore, the dimerization domain of zhx proteins also overlaps with the interaction domain with the AD of NF-YA. The repressor domain of Zhx2 is mapped to the amino acid sequence between residues 263 and 446, which contains the HD1 and a proline-rich region, overlapping with the dimerization domain. The ZHX2 mRNA is expressed among various tissues. Further analysis revealed that Zhx2 is a transcriptional repressor that is localized in the nuclei (34). Zhx2 was recently identified as one factor involved in postnatal repression of fetal expressing genes (38). In summary, ZHX2 is the novel candidate genes for globin regulation in erythroid cells (37).

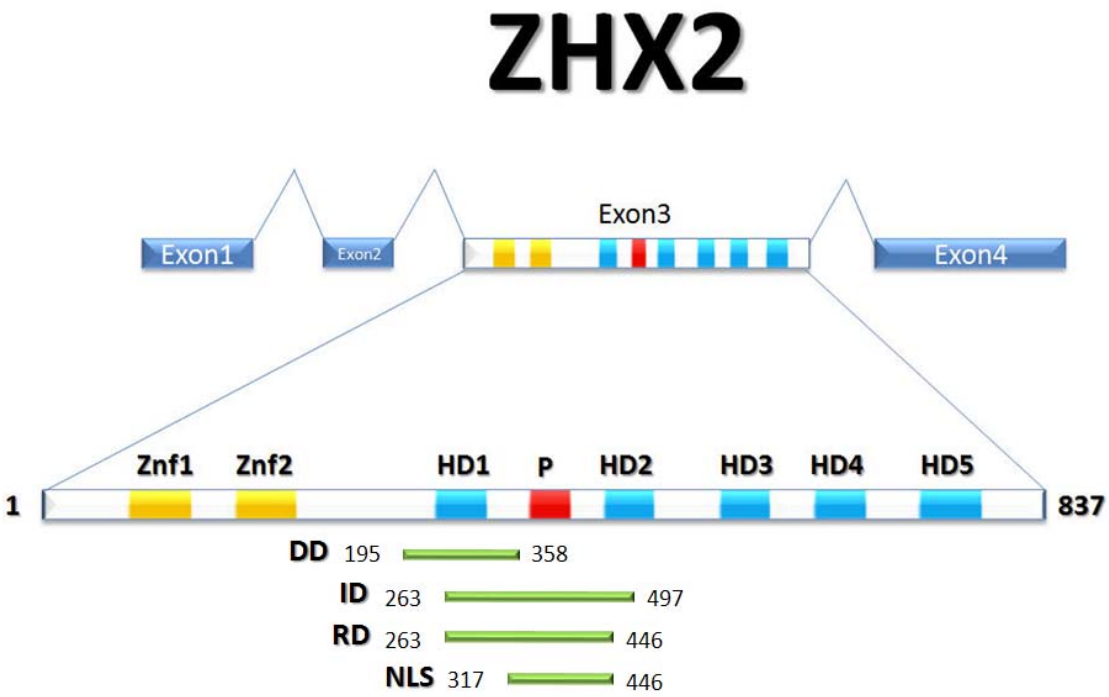


Figure 2.6 Schematic diagram of the functional domains of human ZHX2
DD, dimerization domain; ID, interaction domain with NF-YA; RD, repressor domain; NLS, nuclear localize signal (34).

In 2005, Perincheri S et al. indicated that overexpression of a *Zhx2* transgene restores *H19* repression on a BALB/cJ background, confirming that this gene is responsible for hereditary persistence of the α-fetoprotein (*Afp*) and *H19* which are transcribed at high levels in the mammalian fetal liver but are rapidly repressed postnatally (39). Down regulation of ZHX2 was recently demonstrated in two HPFH-2 subjects by real-time PCR (37).

ZHX2 gene coincides on the QTL on chromosome 8q that has been reported to influence the absolute fetal hemoglobin levels (27). Taken together, ZHX2 is a good candidate gene for regulating γ-globin gene expression.

Table 2.2 Data for SNP rs3802264

SNP ID	Chr 8 position	DNA Chg	Amino acid Chg	Amino acid position	Locus ID	mRNA Accession
rs3802264	124035266	G/A	G/S	779	22882	NM_014943

Several polymorphisms of human ZHX2 gene have been reported. Total SNPs on this gene are 535, 26 of which are in exon by which 5 SNPs are nonsynonymous. In this study, rs3802264 at chromosome position 124035266 in the third exon is selected as:

1. This SNP occurs in coding regions (cSNPs) and change nucleotide from G to A resulting in an amino acid sequence change from glycine to serine at position 779.

2. The information from the International HapMap project demonstrates that this SNP exhibits high heterozygosity and very polymorphic among Asian population.

3. The comparison in multiple alignments of zhx2 protein indicated the highly conserved amino acid at position 779 among various species (Appendix A).

4. The identification of PhosphoMotif Finder search tool predicts that serine position 779 is a putative phosphorylation site by serine kinase. However, it is not known yet whether zhx2 protein is phosphorylated at Ser779 in addition to Ser16 or

not. Phosphorylation of transcription factor may affect protein-protein interaction, modulate its DNA binding affinity and nuclear translocation etc.

For these reasons, it is interesting to study the association of ZHX2 gene and fetal hemoglobin levels in the β-thalassemia/Hb E patient.

CHAPTER III
MATERIALS AND METHODS

Materials

1. Enzymes

Enzymes used for this study were as followings:

a. *Bgl* I (10 U/µl); Fermentas, USA
b. *Taq* DNA polymerase (5 U/µl); Fermentas, USA
c. eAMVTMReverse Transcriptase; Sigma, USA
d. JumpStart AccuTaq LA DNA Polymerase; Sigma, USA

2. Oligonucleotide primers

Oligonucleotide primers used for this study were synthesized by Bio Basic Inc., Canada that listed in the Table 3.1 and 3.2.

Table 3.1 Sequences of primers used for the ZHX2 genotyping

Gene	Primer sequence	Size (bp)	Tm (°C)
ZHX2 (G779A)			
Forward	5’ – AGCTCTGCGAAGAGGACTTG – 3’	20	59.9
Reverse	5’ – CGGAGTCTGATTCAGCCAGT – 3’	20	59.9

Table 3.2 Sequences of primers used for RT-PCR*

Gene	Primer sequence	Size (bp)	Tm (°C)
ZHX2 for RT			
Forward	5’ – AGCTCTGCGAAGAGGACTTG – 3’	20	59.9
Reverse	5’ – ACTTCCCTGTCTAGGCCTG – 3’	19	59.7

Table 3.2 Sequences of primers used for the RT-PCR (continued)

Gene	Primer sequence	Size (bp)	Tm (°C)
GAPDH (136)			
Forward	5′- TCATCCCTGAGCTGAACGGG -3′	20	61.9
Reverse	5′- TTACTCCTTGGAGGCCATGTG -3′	21	60.0

* RT-primers include exon-exon junction.

3. Chemical Substances

The chemical substances that used in this study were analytical and molecular biological grades.

Chemical Substances	Company
Absolute ethanol (C_2H_5OH, MW = 46.07)	Merck, Germany
Agarose	Biorad, USA
Boric acid (H_3BO_3, MW = 61.85)	promega, USA
Bromophenol blue ($C_{19}H_9Br_4O_5SNa$, MW = 691.9)	BDH, England
Deoxyribonucleotide triphosphate (dNTP)	Fermentas, USA
Ethidium bromide ($C_{12}H_{20}N_2Br$, MW = 394.31)	Amersharm,United Kingdom
Ethylene diamine tetraacetic acid disodium salt dihydrate ($C_{10}H_4N_2Na_2O_8.2H_2O$, MW = 372.24)	BDH, England
Glycerol	Sigma, USA
50 bp DNA Ladder	Fermentas, USA
Sodium hydroxide (NaOH, MW = 40)	Sigma, USA
Tris [Hydroxymethyl] aminomethane ($C_4H_{11}NO_3$, MW = 121.1)	Sigma, USA
Tris [Hydroxymethyl] aminomethane hydrochloride ($C_4H_{11}NO_3$.HCI, MW = 157.6)	Sigma, USA
Xylene cyanol	Sigma, USA

4. Equipments

Equipments	Company
Autoclave	Huxley, Taiwan
Automatic pipette	Gilson, France
Bench top centrifuge	Heraeus Sepatech, USA
PTC-200 DNA Engine Thermal Cycler	Biorad, USA
Hotplate stirrer	Barnstead International, USA
Incubator	Heraeus Sepatech, USA
Personal microcentrifuge	Daihan labtech, Korea
pH meter	Beckman, USA
Power supply	Biorad, USA
Submarine agarose gel electrophoresis	Hoefer, USA
Vortex mixer	Scientific Industries Inc., USA
Water bath	Precision pacific Inc., USA
Universal Hood Π and Gel Doc EQ system	Biorad, USA

5. Miscellaneous

Plastic wares	Company
Microcentrifuge tube (0.5, 1.5 ml)	Axygen Scientific Inc., USA
Pipette tip (0.2 µl, 200 µl, 1000 µl)	Axygen Scientific Inc., USA

6. Reagents

6.1 Reagents for DNA amplification by polymerase chain reaction (PCR)

a. Deoxyribonucleotide triphosphates, dNTPs mixture that contained 2 mM in each dNTPs.

b. 10 x PCR buffer (Fermentas Inc., USA) consisted of 750 mM Tris-HCl (pH 8.8), 200 mM $(NH_4)_2SO_4$, 0.1% Tween 20.

c. 25 mM $MgCl_2$

d. Primers of ZHX2 gene as in Table 3.1.

e. *Taq* DNA polymerase

6.2 Reagents for restriction enzyme digestion

a. 10 x buffer O for enzyme *Bgl* I (Fermentas Inc., USA) composed of 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 100 mM NaCl and 0.1 mg/ml BSA.

b. Enzyme *Bgl* I

6.3 Reagents for agarose gel electrophoresis.

a. 1.5% and 3% agarose gel :

Weighed (w/v) in each percentage of agarose gel with 1X TBE buffer.

b. Ethidium bromide :

Dissolved 1 g ethidium bromide in 100 ml of sterilized RO water

c. Gel loading buffer :

Dissolved 0.125 g each of bromophenol blue and xylene cyanol in 50 ml RO water. 15 ml glycerol was added to this reagent.

d. 10X Tris-borate buffer (TBE) :

Dissolved 108 g Trizma base, 54 g boric acid and 40 ml of 0.5 M EDTA (pH 8.0) in 1 L RO water and autoclaved. The stock solution was diluted 10 times before used.

e. Ladder marker :

- 50 bp DNA ladder

6.4 Reagents for reverse transcriptase polymerase chain reaction (RT-PCR)

a. Enhanced Avian HS RT-PCR Kit (Sigma-Aldrich, USA)

i. eAMVTMReverse Transcriptase

ii. JumpStart AccuTaq LA DNA Polymerase

iii. 10x reaction buffers consisted of 50 mM Tris-HCl, pH 8.3, 40 mM KCl, 8 mM MgCl2, 1 mM DTT.

iv. 10 mM dNTP mix

v. Ribonuclease inhibitor

vi. Nuclease-free water

b. Primers (as in Table 3.2) were used.

i. Primers of ZHX2 gene

ii. Primers of GAPDH gene

Methods

1. Study subjects

All of the subjects used for this study were classified mainly based on symptomatic level as followings:

The patients taken for this study were diagnosed by hematological data and DNA analysis. They were categorized into mild, moderate and severe groups according to the severity score, which depended on hemoglobin level, age onset of thalassemic symptom, age as which patients receive their first blood transfusion, size of spleen and degree of growth retardation (Table 3.3) (39). In general, patients whose total severity score was less than 3.5 were categorized in mild cases, patients whose total severity score ranged from 3.5 to 7.5 were considered to be the moderate case and the patients whose total severity score was more than 7.5 were positioned in severe group of β-thalassemia.

A total of 450 DNA samples, 150 mild, 150 moderate, and 150 severe β°-thalassemia/Hb E with normal α-globin genes were recruited in this study.

Table 3.3 Clinical severity criteria and scoring for classifying β°-thalassemia/Hb E patients (137).

Criteria	Status	Score	Status	Score	Status	Score
Hb at steady state (g/dL)	≥7.5	0	6.0–7.5	1	<6	2
Age at onset (year)	>10	0	2–10	0.5	<2	1
Age at first transfusion (year)	>10	0	4–10	1	<4	2
Requirement of transfusion	Rare/none	0	Occasional	1	Regular	2
Spleen condition (cm)	<3	0	3–10	1	>10	2
Splenectomy	No	0	Yes	2		
Growth development	Normal	0	±	0.5	Retarded	1

Mild case Severity Score < 3.5
Moderate case Severity Score 3.5-7.5
Severe case Severity Score > 7.5

2. DNA extraction

DNA samples were provided by Prof. Dr. Suthat Fucharoen (Thalassemia Research Center, Institute of Science and Technology for Research and Development, Mahidol University). DNA was extracted from peripheral blood by Puregene kit (Gentra Systems, Minneapolis, MN, USA). Two volume of red blood cell lysis solution was added to one volume of whole blood in centrifuge tube and mixed rigorously until homogenous, then let stand at room temperature for 10 minutes. Cell pellets was collected after centrifugation at 12,000 rpm, 4 ºC for 5 minutes. This step was repeated until all red blood cells were lyzed. The cell pellets were packed at the bottom and then one volume of cell lysis solution was added to cell pellets and mixed well until the solution is homogenous. RNA-free DNA solution was obtained by addition of 15 µL of RNase A solution. Then 200 µl protein precipitate solution was added, gently mixed for precipitate proteins. After mixing, the sample was kept on ice for 10 minutes. The sample was centrifuged at 12,000 rpm, 4 ºC for 5 minutes. Six hundred µl of 100% isopropanol was pipetted to fresh centrifuge tube and the supernatant from previous step was gently added. DNA would be visible as threads after gently inverting the tube several times. A small white pellet of DNA was collected after centrifugation at 12,000 rpm, 4 ºC for 5 minutes. The pellet was washed once or twice to remove salt by adding 70% ethanol, mixed by inversion and centrifuged at 12,000 rpm for 5 minutes. Air-dried DNA was entirely redissolved with appropriate volume of TE buffer depends on clump of pellet and kept at -20 ºC until used.

3. RNA extraction

Total RNA samples were provided by Prof. Dr. Suthat Fucharoen (Thalassemia Research Center, Institute of Science and Technology for Research and Development, Mahidol University). Total RNAs were isolated from cultured erythroid cells by TRIZOL reagent (Invitrogen, Carlsbad, CA, USA). The sample was added with 1 ml of TRIZOL reagent and mixed homogenously. The homogenized sample was incubated at room temperature for 5 minutes. Next, 200 µl of chloroform was added and shaken for 15 seconds, then left stand at room temperature for an additional 2-3 minutes, followed by centrifuged at 12,000 rpm, 4 ºC for 5 minutes. Centrifugation

separates the sample into three layers. The aqueous phase (upper layer) was transferred to new 1.5 ml eppendrof. The RNA was recovered by precipitation with 500 μl isopropyl alcohol, inverted twice then incubated at room temperature for 10 minutes. This is the RNA precipitation step. The sample was centrifuged at 12,000 rpm, 4 ºC for 10 minutes. The supernatant was discarded and 1 ml of 70% ethanol was added to wash the RNA pellet to remove residual contaminating chemicals, the tube was inverted twice then was centrifuged at 7,500 rpm, 4 ºC for 5 minutes. The RNA pellet was dried at room temperature. The pelleted RNA was dissolved in DEPC-treated water and stored at -20 ºC until used**.**

Total RNA of 2 patients with mild and severe β°-thalassemia/Hb E was extracted from days 7, 10, 12 of cultured erythroid cells which cultured by Fibach method (138). Afterwards, the expression of ZHX2 and GAPDH, served as a control, in the different stages of erythroid maturation at serial days in culture were quantified by reverse transcriptase polymerase chain reaction.

4. Polymerase chain reaction (PCR)

4.1 Principle (139, 140)

The polymerase chain reaction (PCR) is a powerful and clever procedure that takes advantage of DNA polymerase enzymes and synthetic oligonucleotides to make the rapid production of multiple copies of a specific target DNA sequence. The principle of the PCR was first reported in 1971, but it was only after the discovery of thermostable *Taq* DNA polymerase that this technology becomes easy to use. PCR was invented by Kary Mullis and his colleagues in the 1980. PCR is basically an *in vitro* method for the enzymatic amplification that its method based on a cycling reaction in which template DNA is denatured by heating to separate the strands of the molecule. Primer, 20-30 bases fragment of DNA complementary to a region of the template, is annealed to single-strand templates. The cycle ends as the primer molecules are elongated by the action of DNA polymerase to produce molecules that are identical copies of the original template. Initially, thermal cycling was handled manually by transferring samples to be amplified from one water bath to another with the addition of fresh enzyme per cycle after the denaturation step. The first publication of PCR used Klenow polymerase as the elongation enzyme. Due to the

limitation of Klenow polymerase, this enzyme did not stand high temperatures (over 90 ºC) needed to separate the strands of DNA before each round of replication, new enzyme needed to be added for every new cycle. However, special heat–stable polymerase that can endure the heat are now available. One of these, Taq polymerase, comes from *Thermophilus aquaticus*, a bacterium that lives in hot springs and therefore has heat – stable enzyme. This polymerase greatly enhanced the value of PCR lead to invention of automated PCR machine or Thermal cycler. All one has to do is to mix the *Taq* polymerase with the primers, template DNA and DNA precursors (deoxynucleotide triphosphates, dNTPs) in a cap tube, then place it in a thermal cycler. The thermal cycler is programmed to cycle over and over again among three different temperatures: first a high temperature (about 95 ºC) to separate the DNA strands; then a relatively low temperature (about 50 ºC – 70 ºC) to allow the primers to anneal to the template DNA strands; then a medium temperature (about 72 ºC) to allow DNA synthesis. Each cycle takes as little as a few minutes, and it usually takes fewer than 40 cycles to produce as much amplified DNA as one need. Because two primers are used, only the sequence between the two primers will be amplified. Since the cycle is carried out multiple times with twofold increase in the amount of DNA each time, a geometric amplification results such as after 20 cycles would result in a 2^{20} increase in the DNA concentration as shown in figure 3.1. Thus PCR is capable of producing large amounts of DNA fragments from a single piece of template DNA as the amplification increase the amount of fragments produced exponentially. In theory, it is possible to detect a single copy of template DNA by PCR using simple methods. For this reason, PCR is used in medical and biological research labs for a variety of tasks, such as the detection of hereditary disease, the identification of genetic fingerprints, the diagnosis of infectious disease, the cloning of genes and paternity testing.

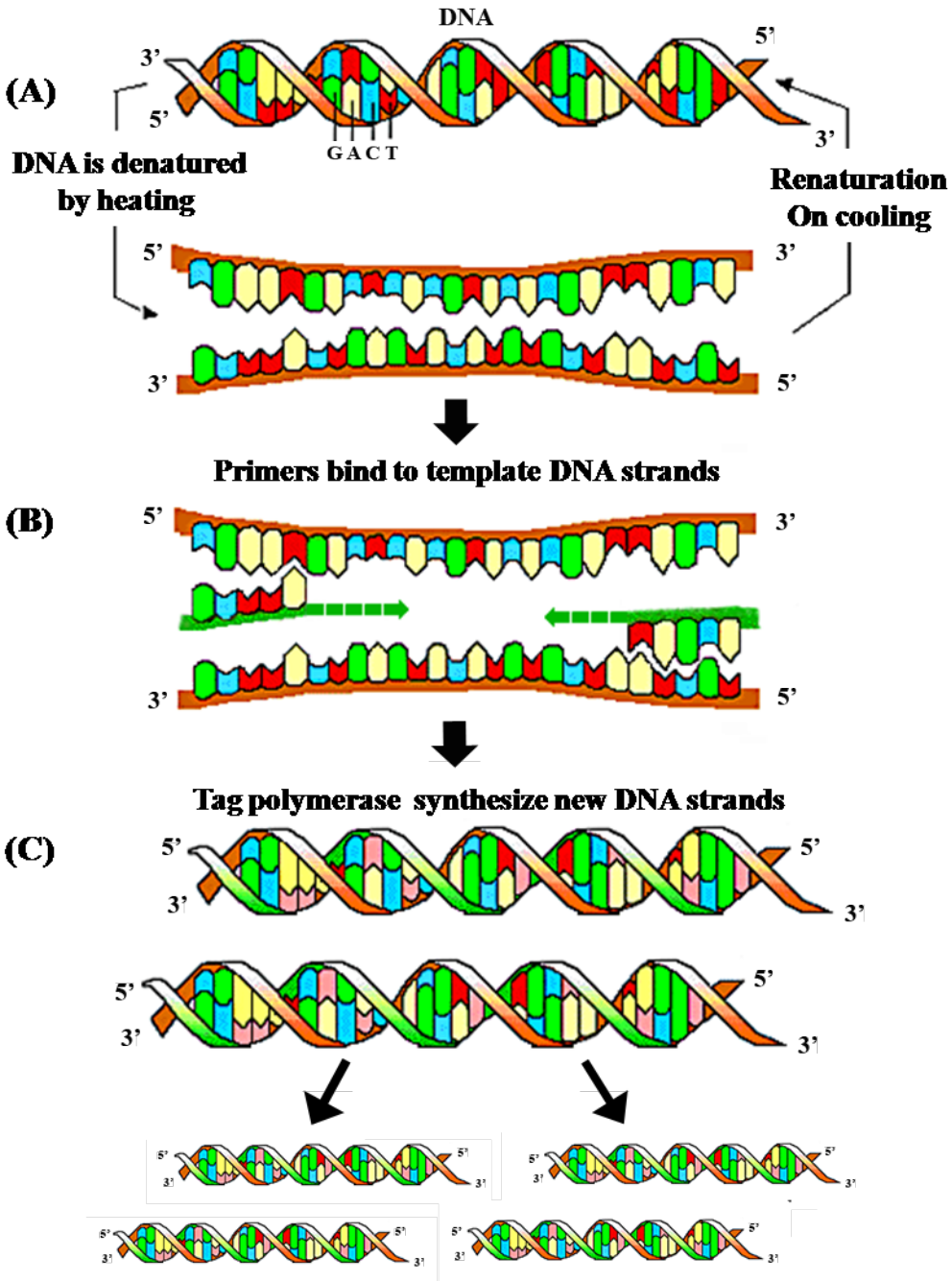


Figure 3.1 An illustration of the polymerase chain reaction (PCR). Step 1: Solution is heated to 95°C to denature the two strands of the target DNA (A). Step 2: Solution is cooled to ~60°C to allow the primers to anneal to the ends of the DNA strands (B). Step 3: Solution is reheated to ~72°C to allow Taq polymerase to synthesize complementary copies of each strand (C). Step 1-3 will repeat 30-35 cycles.

4.2 ZHX2 amplification by PCR

For the analysis of the ZHX2 (G779A) polymorphism the genomic DNA with its primers was used to amplify by polymerase chain reaction (PCR). PCR condition of ZHX2 (G779A) polymorphism was created. The forward and reverse oligonucleotide primers are listed in Table 3.1. Twenty five microliter total volume of each PCR reaction in 0.5 ml PCR tube which consisted of 1 µl genomic DNA, 1 unit *Taq* DNA polymerase, 1X PCR buffer, 200 µM dNTPs, 1.5 mM of $MgCl_2$ and 0.2 µM of each oligonucleotide primers. The reaction was mixed. The optimized PCR was conducted in an automated PTC-200 DNA Engine Thermal cycle. After an initial denaturation for 5 minute at 95 ºC, and then followed by 30 cycles of 45 second 95 ºC for denaturation, annealing at 58 ºC for 30 second, and primer extension at 72 ºC for 45 second. Final extension at 72 ºC for 5 minutes was performed after final cycle to promote completion of partial extension products and complete the annealing of single-stranded complementary products, and followed by chilling at 4 ºC for stopped reaction. The PCR–amplified DNA products of ZHX2 (G779A) polymorphism were analyzed by agarose gel electrophoresis. A specific PCR product was then analyzed by restriction fragment length polymorphism.

5. Restriction fragment length polymorphism analysis

5.1 Principle (141, 142)

This is done with the help of special class of enzymes called restriction endonucleases. Restriction endonucleases or restriction enzymes have been found to play a key role in all aspects of molecular biology. Most of the enzymes used today are type Π enzymes from three types (I, Π and Ш), which have the simplest mode of action. These enzymes are nucleases, and as they cut at an internal position in a DNA strand with specific recognition sites, usually 6-4 basepairs in length, and cleave them in a defined manner. The sequences recognized are palindromic or of an inverted repeat nature; that is, they read the same in both direction on each strand (Figure 3.2). Restriction enzymes can be multiple restriction sites for a single endonuclease within a given piece of DNA, there can be only one (a unique restriction site), or there can be none. It all depends on the sequence of the specific piece of DNA in question. The combination of PCR and restriction fragment analysis offers a very simple, quick and

highly sensitive detection and typing strategy. The restriction endonucleases digestion of each polymorphic PCR amplification product will yield a distinct, unique banding pattern. This approach can also be employed in the diagnosis and detection of allelic polymorphism or mutation associated with genetic influenced disease.

5.2 Restriction enzyme digestion reaction

The amplified DNA or PCR product digestion was performed in a final reaction volume of 15 µl. The reaction was contained with 10X buffer O (Fermentas Inc., USA), 5 unit of *Bgl*I restriction enzyme (Fermentas Inc., USA) and 7 µl of each PCR product. The mixture was incubated at 37 ºC for 16 hours. After that, the restriction DNA fragments were analyzed by agarose gel electrophoresis.

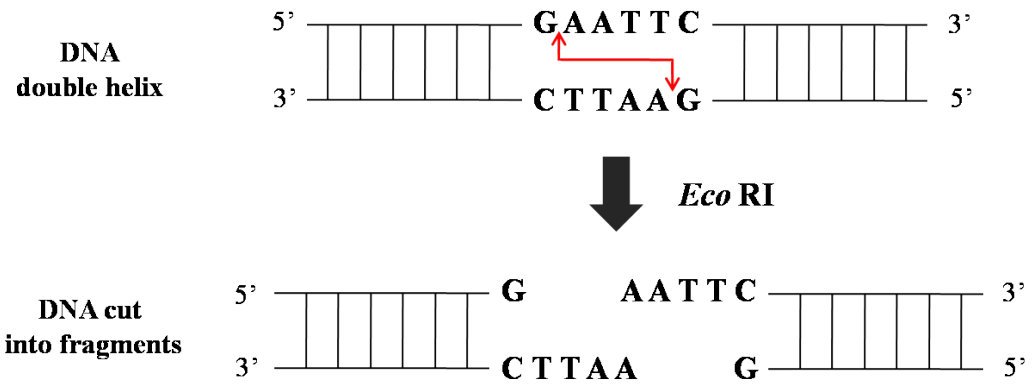


Figure 3.2 The schematic diagram of restriction enzyme digestion. Each restriction enzyme will cut the DNA only when a certain sequence of bases occurs e.g. the enzyme *Eco*RI cuts the DNA between bases G and A only when the sequence GAATTC is present in the DNA. The other restriction enzyme used cuts the DNA at a different sequence of bases. Thus, each restriction enzyme is specific.

6. Electrophoresis (143)

Electrophoresis is the process of moving charged molecules in solution by applying an electric field across the mixture. Because molecules in an electric field move with a speed dependent on their charge, shape, and size, electrophoresis has been extensively developed for molecular separations. As an analytical tool, electrophoresis is simple one relatively rapid. The method relies on the fact that nucleic acids are polyanionic at neutral pH, i.e. they carry multiple negative charges due to the phosphate groups on the phosphodiester backbone of the nucleic acid strands. This means that the molecules will migrate towards the positive electrode when placed in an electric field. The technique is carried out using a gel matrix which separates the nucleic acid molecules according to size. There are two types of matrix that commonly used for electrophoresis, agarose and polyacrylamide, and have important consequences for the degree of separation achieved, which is dependent on the porosity of the matrix.

6.1 Agarose gel electrophoresis (144, 145)

Gels are indispensable tools for the molecular biologist. Agarose is one of material that can be formed into hydrophilic polymers as well as hydrated gels in aqueous solution or water. Agarose gels are more porous and have a larger pore size as compared to polyacrylamide gels and are, therefore, used to fractionate large macromolecules such as nucleic acids that cannot readily penetrate into and move through other types of supporting materials. Agarose used for electrophoresis is a more purified form of the agar that used to make bacterial culture plates. It is a linear polymer of D-galactose and 3, 6-anhydro-L-galactose. Agarose gels are casted by boiling agarose in presence of a buffer, then poured into a mold and allowed to solidify to form a matrix. Porosity of the gel is determined by concentration of agarose. Higher the agarose concentration, smaller the pore size and lower the agarose concentration, larger the pore size. When an electric field is applied to an agrose gel in the presence of a buffer solution which will conduct electricity, DNA fragments, highly negatively charged at neutral pH move through the gel towards the positive electrode at rates determined by their molecular size and conformation. Since charge/mass ratio in nucleic acids is equal, rate of migration of DNA molecules is inversely proportional to $\log_{10}$ of their molecular weights, i.e. small linear

fragments move more quickly than large ones, which are retarded by entanglement with the network of agarose fibers forming the gel. The DNA samples are placed in wells in the gel surface, the power supply is switched on and the DNA is allowed to migrate through the gel in separate lanes or tracks. The added dye also migrates, and is used to follow the progress of electrophoresis. The DNA is stained by the inclusion of ethidium bromide in the gel, or by soaking the gel in a solution of ethidium bromide after electrophoresis. The DNA shows up as an orange band on illumination by UV light (Figure 3.3).

6.1.1 Agarose gel preparation

The gel tray was fixed into gel setting block and placed the slot – forming combs at the position. Agarose powder was dissolved at the desired concentration (w/v) with 1X TBE buffer and boiled on the hot – plate or microwave oven. When the gel was completely dissolved, the agarose solution was let to be cool down at room temperature at approximately 50 ºC, then poured into the block with about 3 mm thickness. Once the block was tilted, the gel flowed behind the comb and air bubbles were removed by a pipette tip. When gel was solidified at room temperature, the comb was removed from the polymerized gel and placed the gel in the electrophoresis apparatus. The chamber was added with 1X TBE buffer to fill the electrode chambers and covered the gel with this buffer in a depth of about 1 mm.

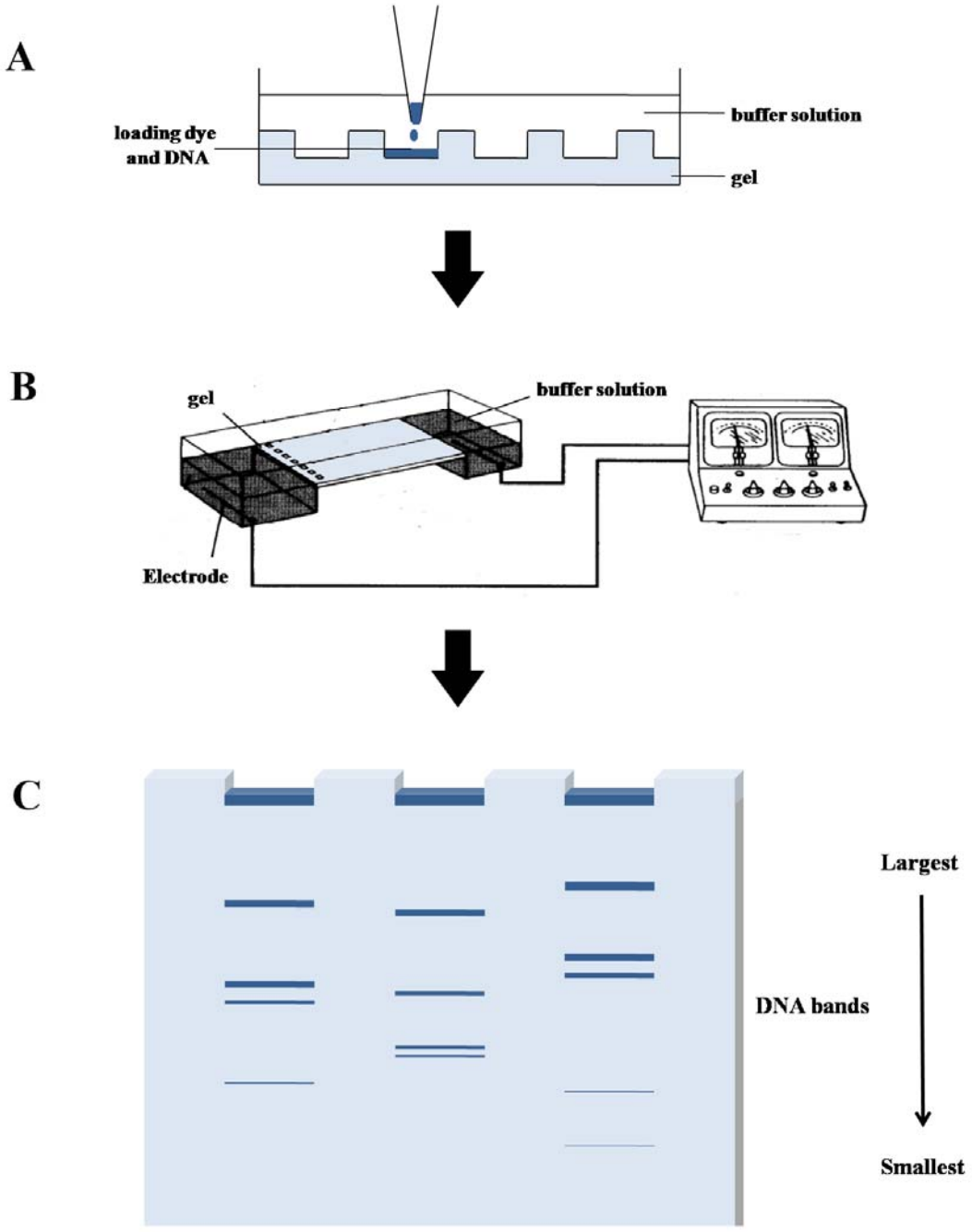


Figure 3.3 An illustration of agarose gel electrophoresis. (A) load dye and DNA into agarose gel, (B) an electric field is applied to an agrose gel in the presence of a buffer solution which will conduct electricity, the DNA migrate from cathode to anode, (C) The DNA is stained by ethidium bromide.

6.1.2 Detection of amplified ZHX2 (G779A) polymorphism and PCR product of ZHX2

The PCR product of ZHX2 was detected by 1.5% agarose gel electrophoresis. Five µl of amplified product of ZHX2 gene was mixed with 2 µl of loading dye and loaded in each well of agarose gel compared with 50 bp DNA marker. The agarose gel was run for 30 minutes at 90 volts. ZHX2 fragment was stained by ethidium bromide and visualized on a uv – transilluminator as fragment of 276 bp.

The restriction DNA fragments of ZHX2 (G779A) polymorphism was detected by 3% agarose gel electrophoresis. Five µl of amplified product of ACE gene was mixed with 2 µl of loading dye and loaded in each well of agarose gel compared with 50 bp DNA marker. The agarose gel was run for 35 minutes at 90 volts, then the 2 two fragments with 149 and 127 bp in the presence of G allele and 276 bp in the presence of A allele were observed by submerged in ethidium bromide solution for 5 minute. The stained gel was destained in distilled water for 5 minutes and visualized on UV – transilluminator. Photograph was taken by Gel Doc EQ system to collected data for analysis.

7. Reverse transcriptase polymerase chain reaction (RT-PCR) (146, 147)

7.1 Principle (146, 147)

Reverse transcriptase polymerase chain reaction is the common techniques used to detect or quantify the expression of mRNA. It is the process which allows RNA sequences to be amplified indirectly by converting RNA to DNA and then subsequent amplifying the DNA that has been reversely transcribed. The enzyme reverse transcriptase is used to copy the RNA into complementary DNA (cDNA); the cDNA is then amplified by PCR. In the first step of RT-PCR, which is called the first strand reaction, the complementary DNA (cDNA) is produced from mRNA template that uses dNTPs, a reverse transcriptase and random primers, Oligo-dT or a gene-specific primer in a reverse transcriptase buffer. When the reverse transcriptase reaction has been completed, cDNA will have been produced from the original single stranded mRNA. Then standard polymerase chain reaction, or second strand reaction, is initiated (Figure 3.4).

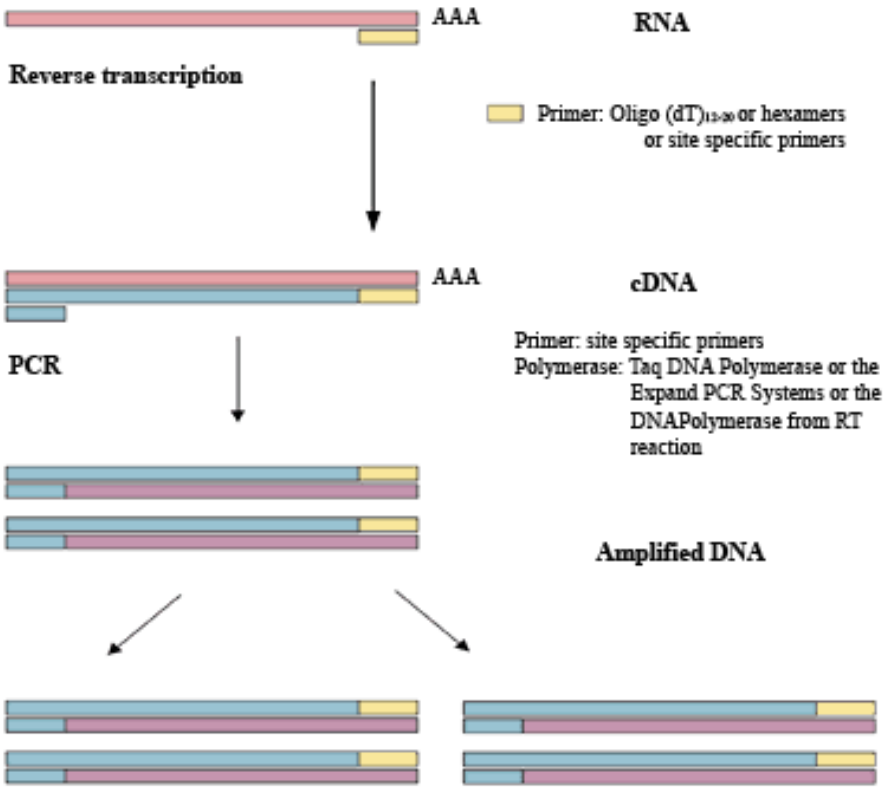


Figure 3.4 cDNA synthesis using RT-PCR (reverse transcriptase PCR). The first stand synthesis of cDNA (catalyzed by reverse transcriptase) may be primed by using either a gene-specific primer (GSP), oligo(dT), or a mixture of random hexameric oligonucleotides. Second strand synthesis (amplification cycle 1; catalyzed by a thermostable DNA polymerase) is primed with the sense primer. Amplification continues in the presence of both sense and antisense primers.

7.2 Quantification of ZHX2 gene by reverse transcriptase polymerase chain reaction (RT-PCR)

RT-PCR was performed by using the Enhanced Avian HS RT-PCR Kit (Sigma). ZHX2 specific primers were used.

7.2.1 First strand cDNA synthesis from total RNA

0.25 μg total RNA of each sample was mixed with 1 μl Anchored oligo $(dT)_{23}$ (0.5 μg/μl), 1 μl dNTP mix (10 mM). DEPC-treated water was added up to the final volume of 10 μl. Samples were mixed gently and briefly centrifuged to collect all components at the bottom of the tube. Each sample was incubated at 70°C for 10 min for denature RNA secondary structure and chilled on ice for at least 1 min. Samples were briefly centrifuged and 10 μl of the reaction mix, containing 2 μl 10X buffer for eAMV-RT, 1 μl RNAse inhibitor, 1 μl of eAMV RT, 6 μl DEPC-treated water , were added. Contents of the tubes were mixed and incubated at 50 °C for 50 min. The reaction was inactivated by incubating at 70°C for 15 min. The first strand cDNA is now ready for subsequent PCR amplification. The cDNA synthesis was verified by amplification of the GAPDH gene via PCR with control primers then they were visualized after gel electrophoresis by ethidium bromide staining.

7.2.2 PCR amplification of cDNA

Two μl of reverse transcribed cDNA was used for PCR amplification with specific primer pairs are listed in Table 3.2. The reaction volume was 25 μl in 0.6 ml PCR tube which consisted of 2 μl template DNA (cDNA) from RT reaction, 0.5 μl JumpStart AccuTaq LA DNA polymerase (2.5U/μl), 2.5 μl 10X AccuTaq buffer, 2.5 μl dNTP mix (2 mM), 13.5 μl water and 2 μl oligonucleotide primers (5 pmol/μl). The reaction was mixed. The optimized PCR was conducted in an automated PTC-200 DNA Engine Thermal cycle. The following PCR profile was used: after an initial denaturation for 5 minute at 95 ºC, and then followed by 30 cycles of 45 second 95 ºC for denaturation, annealing at 57 ºC for 30 second, and primer extension at 72 ºC for 45 second. Final extension at 72 ºC for 5 minutes was performed after final cycle to promote completion of partial extension products and complete the annealing of single-stranded complementary products, and followed by chilling at 4 ºC for stopped reaction. The PCR product was applied to an 1.5% agarose gel followed by visualized

after gel electrophoresis by ethidium bromide staining.and bands of the expected size were gel-extracted and sequenced.

8. Statistical analysis

All statistical analysis was conducted by software package STATA version 9.0 and MedCalC. The differential among hematological data such as fetal hemoglobin (HbF) level measured as percentage of total hemoglobin, adjusted fetal hemoglobin and proportion of Hb F to Hb E according to ZHX2 genotype in subjects were compared by one way ANOVA. The allele frequencies or genotype frequencies were analyzed by gene counting method and their distribution between the study groups were compared by Chi–square (χ^2) analysis. The Chi–square analysis was carried out to test deviation of genotype frequencies from those predicted by the Hardy–Weinberg equilibrium hypothesis. Moreover, the Chi–square analysis was also determined the influence of ZHX2 polymorphism in the study. To assess the association among ZHX2 and each factor (i.e. percentage of Hb F), Chi-square test were applied. Crude (unadjusted or exposure) odds ratio and 95% confidence interval (CI) were estimated. To determine the association among ZHX2 and *Xmn*I polymorphisms on percentage of fetal hemoglobin and on severity of thalassemia, a condition logistic model or multiply logistic regression analysis were fitted. Adjusted odds ratio and 95% CI obtained from the model were reported. The nominal level of statistical significance for all analyses was <0.05.

CHAPTER IV
RESULTS

1. Clinical characteristics of the study population

The study populations were selected from 3 groups of Thai-Chinese β°-thalassemia/Hb E patients, including 150 mild, 150 moderate, and 150 severe β°-thalassemia/Hb E patients with normal α-globin genes. Thereby, the cohorts used in this study consisted of 450 patients; 219 males and 231 females. The characteristics of β°-thalassemia/Hb E patients were divided into subgroups according to clinical severity shown in Table 4.1.

The subjects of this study under age group 2-64 years old, 17.6 ± 12.4 (mean ± SD). Their physical characteristics were: height 138.5 ± 22.4 cm. and body mass 33.9 ± 14.1 kg. Total hemoglobin levels ranged between 3.4 and 12.1 g/dL and hematocrit levels varied from 12.6 – 41.0 %. The subjects'age, height, weight, baseline hemoglobin level and percent Hematocrit are statistically significant different among the severity. Furthermore, Hemoglobin F (g/dL), percentage of Hemoglobin F and propotion of Hb F to Hb E are assorted between 0.19 -6.56, 6.85 - 66.56, 0.07 - 1.99 respectively. And all of them showed significant difference among the severity group as well ($p<0.0001$).

Table 4.1 Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients according to severity (Data expressed as mean ± SD)

Parameter	mild (n = 150)	moderate (n = 150)	severe (n = 150)	p-value*
Age (years)	18.23±11.82	19.55±14.69	15.04±9.73	0.0050
Sex Male/Female	69/81	76/74	74/76	na
Weight (kg)	38.01±15.36	34.59±14.32	29.01±10.79	<0.0001
Height (cm)	143.00±23.41	139.67±22.61	132.79±20.04	0.0003
Baseline Hb level (g/dL)	7.87±1.15	6.85±1.14	5.99±1.13	<0.0001
Hematocrit (%)	25.15±3.42	23.22±3.95	21.48±3.77	<0.0001
Hemoglobin F (g/dL)	3.18±1.13	1.98±0.83	1.28±0.68	<0.0001
Hemoglobin F (%)	39.36±10.84	28.69±11.85	20.69±11.25	<0.0001
Propotion of Hb F to Hb E	0.76±0.34	0.58±0.25	0.44±0.23	<0.0001

* = *Using oneway ANOVA test*

na = not applicable

2. Optimization Condition for PCR and RT-PCR

The PCR condition were optimized on normal DNA template sample by varying the annealing temperature from 55 ºC to 60 ºC and $MgCl_2$ concentration from 1.5 mM to 2.5 mM. After getting the optimized PCR condition (Table 4.2), the G779A region at exon 3 of ZHX2 gene was amplified using genomic DNA of 450 β°-Thalassemia/Hb E patients as a template by PCR method.

The RT-PCR condition were optimized on normal cDNA template sample by varying the annealing temperature from 55 ºC to 60 ºC, $MgCl_2$ concentration from 1.5 mM to 2.5 mM and primer concentration from 0.04 μM to 0.2 μM. The optimized RT-PCR condition of ZHX2 and GAPDH are shown in Table 4.2. Then the cDNA of 1 mild and 1 severe case of β°-Thalassemia/Hb E erythroid culture cells was amplified as a template.

Table 4.2 The optimized PCR and RT-PCR conditions *****

Gene	Primer		Annealing	$MgCl_2$
	Primer sequences (5'—3')	(μM)	Temp (°C)	(mM)
PCR				
ZHX2				
Forward	– AGCTCTGCGAAGAGGACTTG –	0.2	58	1.5
Reverese	– CGGAGTCTGATTCAGCCAGT –	0.2		
RT-PCR				
ZHX2				
Forward	– AGCTCTGCGAAGAGGACTTG –	0.2	57	1.5
Reverese	– ACTTCCCTGTCTAGGCCTG –	0.2		
GAPDH				
Forward	– TCATCCCTGAGCTGAACGGG –	0.2	57	1.5
Reverese	– TTACTCCTTGGAGGCCATGTG –	0.2		

***** Each PCR reaction has 25 μL total volume with 1 X buffer, 0.2 mM dNTPs, 1 U Taq polymerase.

3. Analysis of G779A polymorphism of ZHX2 gene

In order to analyze the G779A base substitution at exon 3 of the ZHX2 gene, the gDNA was amplified by PCR with oligonucleotide primers (Figure 4.1). The 276 bp amplified-DNA product which contained a cutting site for the *Bgl*I restriction enzyme was separated in 1.5% agarose gel (Figure 4.2). The restriction enzyme *Bgl*I cut 276 bp fragment into 149 bp and 127 bp that was defined as the G allele. In contrast, the A allele variant had no site cut by *Bgl*I. The fragment after digestion would remain 276 bp. After restriction enzyme digested PCR-amplified DNA, the DNA fragments were analyzed by 3% agarose gel electrophoresis (Figure 4.3).

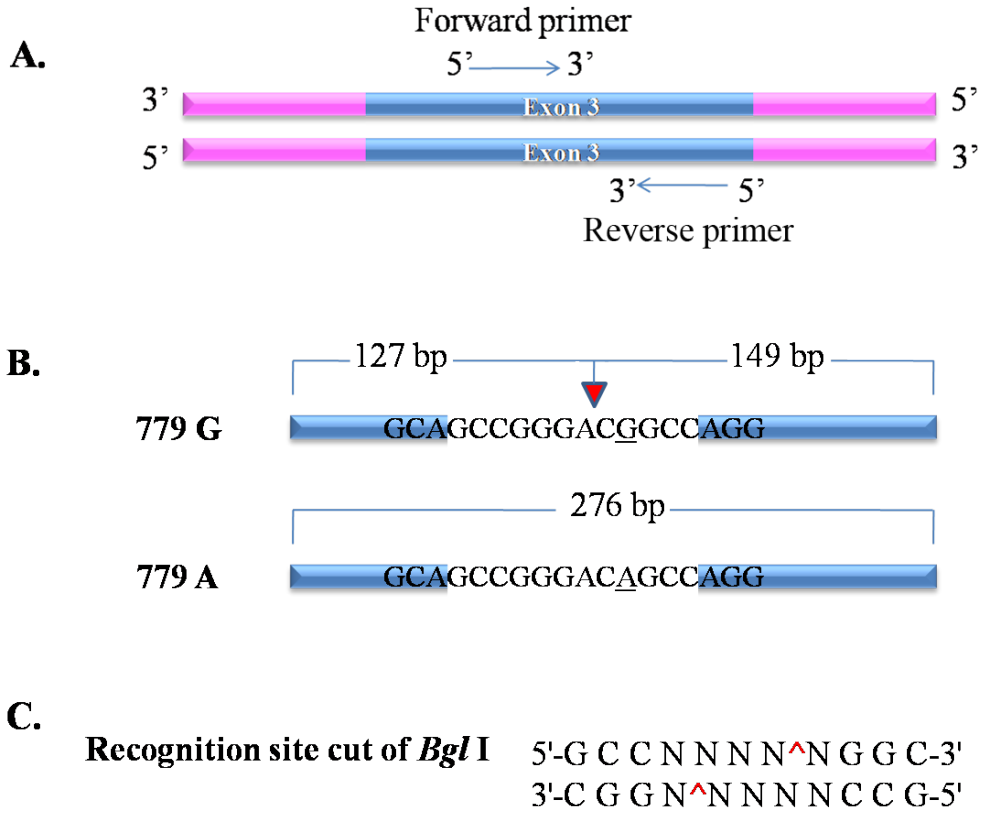


Figure 4.1 (A). PCR amplification of ZHX2 gene used oligonucleotide primers. (B). The amplicon which had one restriction site cut for restriction enzyme *Bgl*I would be digested into 2 fragments of 127 and 149 bp in the present of G at nucleotide 779, whereas the amplicon in the presence of A at nucleotide 779 remain undigested (276 bp). (C). 11-bp sequence and the cut site by *Bgl*I are displayed.

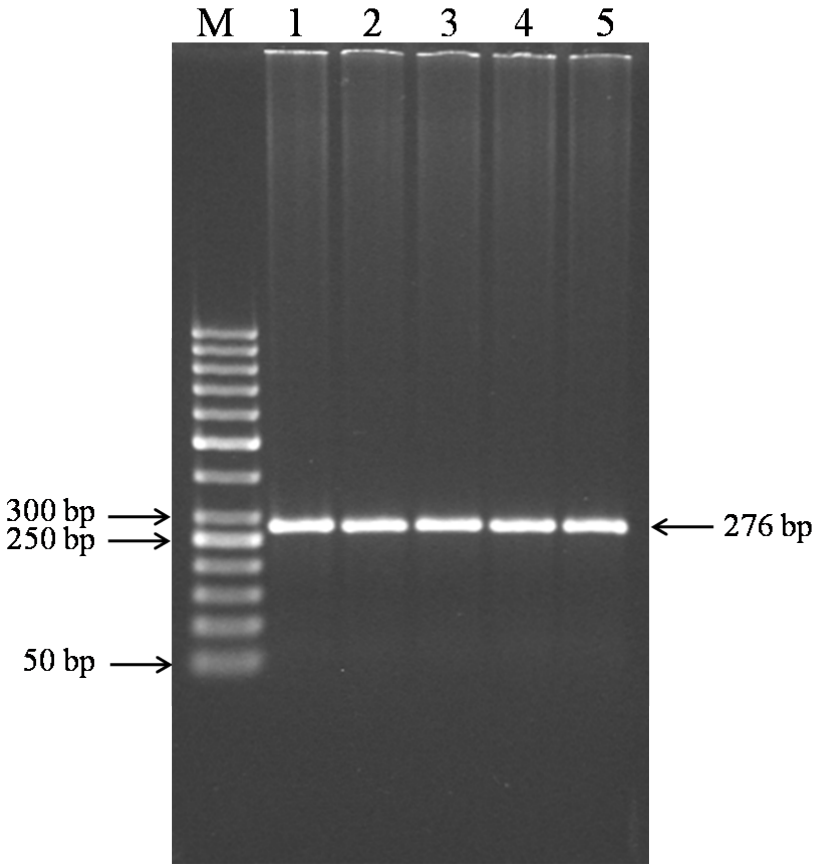


Figure 4.2 The PCR-amplified product of ZHX2 gene in 1.5 % agarose gel electrophoresis. Lane M was 50 bp ladders. Lanes 1-5 were ZHX2-PCR product (276 bp).

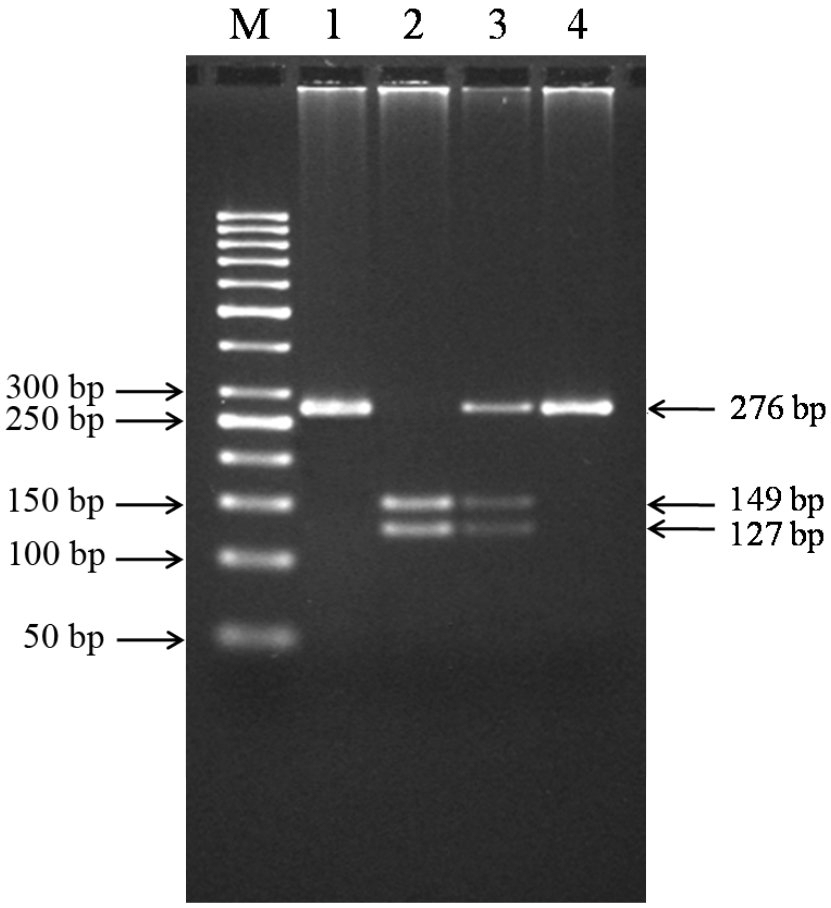


Figure 4.3 Analysis of the G779A polymorphism of ZHX2. PCR product of ZHX2 were digested with *Bgl*I and analyzed by 3 % agarose gel electrophoresis. Lane M was 50 bp ladders. Lane 1 was uncut ZHX2 PCR product. Lane 2 was homozygous GG (149 bp and 127 bp), lane 3 was heterozygous AG (276 bp, 149 bp and 127 bp), and lane 4 was homozygous AA (276 bp).

4. Study on the genotype frequencies and allele frequencies of G779A polymorphism of ZHX2 gene

4.1 Hardy-Weinberg equilibrium

Genotype distributions and allele frequencies of ZHX2 G779A polymorphism among 3 groups were shown in Table 4.3. All genotypes of ZHX2 G779A polymorphism were conform with Hardy-Weinberg equilibrium using (p=0.05) binomial equation as followed: $(p+q)^2 = p^2 + 2pq + q^2$.

4.2 Genotype frequencies

Genotype frequencies of ZHX2 G779A polymorphism were analyzed in 150 mild, 150 moderate and 150 severe cases with β°-thalassemia/Hb E disease. Table 4.4 showed the distribution of genotype frequencies. The frequencies of GG, AG and AA genotypes in mild cases were 38.0%, 50.7% and 11.3%, respectively, moderate cases: 36.7%, 50.7% and 12.7%, respectively and severe cases: 40.0%, 50.0% and 10.0% respectively.

Chi-squares analysis indicated that there was no statistically significant difference in G779A genotype distributions between mild and moderate cases (p=0.9050). In the same way, homozygous GG genotype among mild and severe cases, moderate and severe cases as shown in this table were also not significantly different (p=0.8129, 0.6348).

4.3 Allele frequencies

Allele frequencies of G and A of ZHX2 G779A polymorphism among 3 groups were calculated and displayed in Table 4.4. Estimated allele frequencies of G allele and A allele were not significant difference between mild and moderate cases (χ^2=0.064, p=0.8001). The frequency of G allele in mild cases was 63.3% which was not significant higher from 65.0% in severe cases. Consequently, this result refers to no significantly difference of G and A alleles were observed between mild and severe cases (χ^2 = 0.116, p = 0.7334). The result in moderate and severe cases was similarly not significant difference (χ^2 = 0.460; p = 0.4975).

Table 4.3 The genotype frequencies distribution of ZHX2 G779A polymorphism in Hardy-Weinberg equilibrium

ZHX2 G779A polymorphism				
Genotype	mild (n = 150)	moderate (n = 150)	severe (n = 150)	All
GG	57	55	60	172
AG	76	76	75	227
AA	17	19	15	51
χ^2	1.240	0.850	1.467	3.447
p-value	0.2655	0.3566	0.2258	0.0634
Allele G	0.63	0.62	0.65	0.63
95% CI	0.5774-0.6858	0.5639-0.6731	0.5944-0.7018	0.6024-0.6652
Allele A	0.37	0.38	0.35	0.37
95% CI	0.3142-0.4226	0.3269-0.4361	0.2982-0.4056	0.3348-0.3976

Table 4.4 Distribution of genotype frequencies and allele frequencies of G779A polymorphism of ZHX2 gene in mild, moderate and severe cases with β°-thalassemia/Hb E patients

Genotype	mild (n = 150)	moderate (n = 150)	severe (n = 150)
GG	57 (38.0%)	55 (36.7%)	60 (40.0%)
AG	76 (50.7%)	76 (50.7%)	75 (50.0%)
AA	17 (11.3%)	19 (12.7%)	15 (10.0%)
Allele G	190 (63.3%)	186 (62.0%)	195 (65.0%)
Allele A	110 (36.7%)	114 (38.0%)	105 (35.0%)

Mild cases VS Moderate cases:

GG : $\chi^2 = 0.014$; p = 0.9050; df = 1 *Allele frequency: $\chi^2 = 0.064$; p = 0.8001; df = 1*

Mild cases VS Severe cases:

GG : $\chi^2 = 0.056$; p = 0.8129; df = 1 *Allele frequency: $\chi^2 = 0.116$; p = 0.7334; df =1*

Moderate cases VS Severe cases:

GG : $\chi^2 = 0.226$; p = 0.6348; df = 1 *Allele frequency: $\chi^2 = 0.460$; p = 0.4975; df = 1*

5. Clinical characteristics of the study population according to ZHX2 gene (G779A) genotypes

The anthropometric characteristics of β°-thalassemia/Hb E patients were divided into subgroups according to ZHX2 gene (G779A) genotypes shown in Table 4.5.

The subjects' age and their physical characteristic; height, weight are not statistically significant different among the three genotypes of ZHX2 gene (G779A). Furthermore, the hematological data; baseline hemoglobin level, percent Hematocrit, Hemoglobin F (g/dL), percentage of Hemoglobin F and propotion of Hb F to Hb E also showed no significant difference among the three genotypes of ZHX2 (G779A).

Table 4.5 Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients according to ZHX2 gene (G779A) genotypes (Data expressed as mean ± SD)

Parameter	AA (n = 51)	AG (n = 227)	GG (n = 172)	p-value*
Age (years)	16.82±12.15	17.82±12.55	17.58±12.35	0.8923
Sex Male/Female	24/27	118/109	89/83	na
Weight (kg)	33.36±15.33	34.19±14.05	33.68±13.86	0.9199
Height (cm)	137.20±24.05	138.84±22.07	138.56±22.46	0.9110
Baseline Hb level (g/dL)	7.094±1.19	6.84±1.43	6.91±1.35	0.5053
Hematocrit (%)	24.33±3.43	23.12±5.71	23.53±4.19	0.2731
Hemoglobin F (g/dL)	2.40±1.19	2.14±1.21	2.08±1.17	0.3179
Hemoglobin F (%)	37.24±12.25	35.15±11.58	34.34±11.37	0.2951
Propotion of Hb F to Hb E	0.66±0.35	0.59±0.30	0.57±0.31	0.2423

* = *Using oneway ANOVA test*

na = not applicable

6. Assessment the risk of ZHX2 G779A polymorphism genotype and allele frequencies between mild and severe cases with β°-thalassemia/Hb E

To study the association of ZHX2 polymorphism on mild and severe cases, the logistic regression was used to evaluate for the risk contribution of ZHX2 G779A polymorphism on mild compared with severe cases. If the odds ratio is more than 1, there will be a positive correlation between the ZHX2 G779A genotypes or allele frequencies with severe case. In contrast, when the odds ratio is less than 1, there is a negative correlation. Moreover, 95% confidence interval (95% CI) values were considered as the risk factor indicator if the values of 95% CI equal and more than 1.

The odds ratio of ZHX2 G779A polymorphism genotype and allele frequencies between mild subjects and severe subjects in this study were summarized in Table 4.6. The odds ratio of AG and AA genotypes were 0.9375 (95% CI: 0.5784-1.5195, p=0.7934) and 0.8382 (95% CI: 0.3830-1.8346, p=0.7085), respectively in severe compared with mild cases. The results showed that having A or G allele of ZHX2 G779A polymorphism did not confer risk of being mild or severe among β°-thalassemia/Hb E patients.

Table 4.6 The odds ratio of ZHX2 G779A polymorphism genotype and allele frequencies between mild and severe cases with β°-thalassemia/Hb E

Genotype	Mild (n=150)	Severe (n=150)	Odds ratio	95% CI	p-value
GG	57 (38.0%)	60 (40.0%)	1.0000	-	-
AG	76 (50.7%)	75 (50.0%)	0.9375	0.5784-1.5195	0.7934
AA	17 (11.3%)	15 (10.0%)	0.8382	0.3830-1.8346	0.6588
Allele G	190 (63.3%)	195 (65.0%)	0.9301	0.6661-1.2986	0.6704
Allele A	110 (36.7%)	105 (35.0%)			

7. Assessment the risk of ZHX2 G779A polymorphism genotype and allele frequencies between high and low percentage of Hb F with β°-thalassemia/Hb E

G779A genotype and allele frequencies of ZHX2 gene were also used to evaluate the odds ratio for the risk assessment on high percentage of Hb F subjects compared with low percentage of Hb F subjects. The odds ratio of G779A polymorphism genotype distributions and allele frequencies between high and low percentage of Hb F with β°-thalassemia/Hb E patients in this study were summarized in Table 4.7. In this study, the results showed that the odds ratio (OR) of AG genotype, using the genotype of major common allele as a reference, in low percentage of Hb F compared with high percentage of Hb F was 0.9126, 95% CI: 0.6139-1.3568 (p=0.6514) and AA genotype was 0.6077, 95% CI: 0.3199-1.1543 (p=0.1281). This is due to no different lower GG frequency in patients with high percentage of Hb F. Furthermore, frequency of G allele demonstrated were no association with risk of low percentage of Hb F compared with high percentage of Hb F (OR = 0.8364, 95% CI : 0.6368-1.0985, p=1.990), as shown in Table 4.7. These results indicated that GG genotype and G allele were not associated with the risk of having low percentage of Hb F and may not exert a protective effect on high percentage of Hb F.

For the separation of % Hb F threshold, % HbF of 29 has been evaluated as a threshold level for diminished Hb F level. Because from the histogram of percentage of Hb F demonstrated the percentage of Hb F seperated into 2 groups and there are the cutoff point at 29% (Figure 4.4). The high Hb F group shown the different from low Hb F group that is the spleen size < 18 cm, almost no 3rd degree of retardation and they tend to have fewer requirements for blood transfusion.

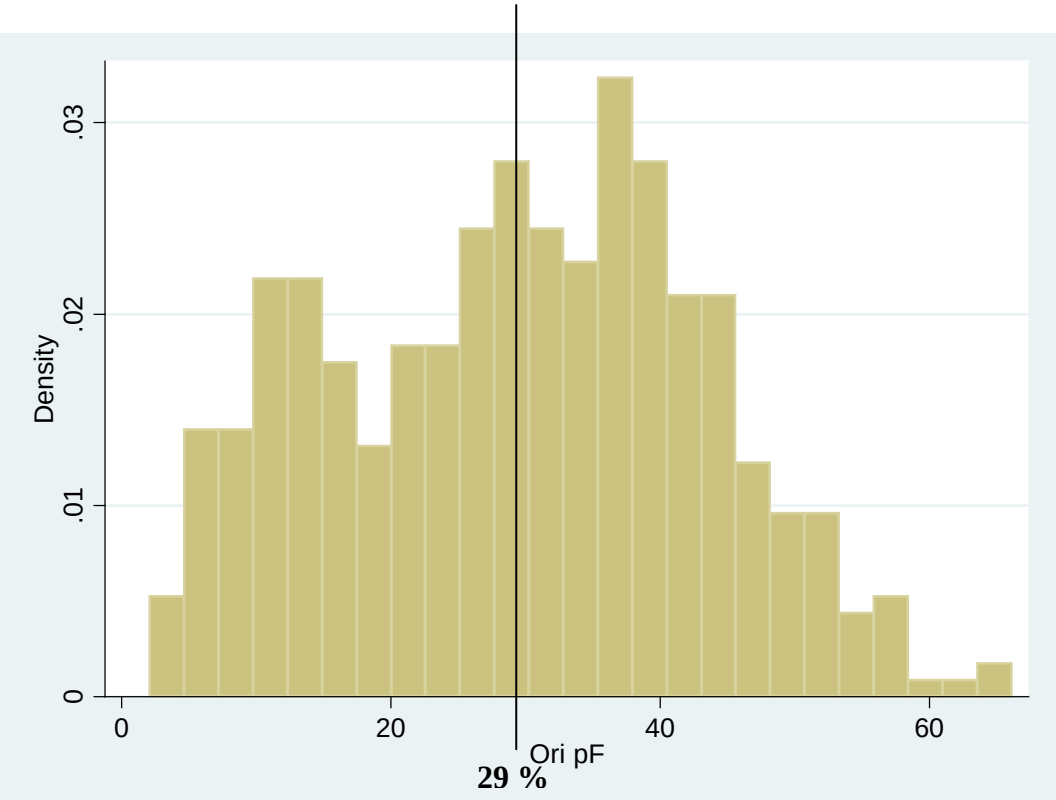


Figure 4.4 Histogram of percentage of Hb F

Table 4.7 The odds ratio of ZHX2 G779A genotype and allele frequencies between high and low percentage of Hb F with β°-thalassemia/Hb E

Genotype	High Percentage of Hb F (n=239)	Low Percentage of Hb F (n=211)	Odds ratio	95% CI	p-value
GG	87 (36.4%)	85 (40.3%)	1.0000	-	-
AG	120 (50.2%)	107 (50.7%)	0.9126	0.6139-1.3568	0.6514
AA	32 (13.4%)	19 (9.0%)	0.6077	0.3199-1.1543	0.1281
Allele G	294 (61.5%)	277 (65.6%)	0.8364	0.6368-1.0985	0.1990
Allele A	184 (38.5%)	145 (34.4%)			

8. The influence of ZHX2 polymorphism on the percentage of Hb F

Perincheri, S. et al (39) demonstrated that ZHX2 gene was related to the regulation of gene which expressed in fetal state. The one of gene that expressed is γ-gene which involve in Hb F level. The results of this study, we found that among the 450 cases of β°-thalassemia/Hb E subjects showed no significant differences in high and low percentage of Hb F among ZHX2 gene (G779A) genotypes as shown in Table 4.8.

Table 4.8 The influence of ZHX2 polymorphism and percentage of Hb F in β°-thalassemia/Hb E patients

Genotype	GG (n=172)	AG (n=227)	AA (n=51)	p-value*
High Percentage of Hb F	40.63±7.82	40.35±7.47	41.15±8.28	0.7772
Low Percentage of Hb F	17.95±7.56	18.33±7.14	19.03±7.11	0.8329

* = *Using oneway ANOVA test*

9. Study on the distribution of ZHX2 G779A polymorphism genotype and allele frequencies between genders in Thai population

Next, we study if whether there is sex preference among G779A genotypes. The results was shown in Table 4.9 indicated the distribution of ZHX2 G779A polymorphism genotypes between genders among three groups of subjects and showed no significant difference between male and female.

Table 4.9 Distribution of ZHX2 G779A polymorphism genotype and allele frequencies between genders in Thai population (n = 450)

Genotype frequencies			
Genotype	Male	Female	Statistical value
GG AG AA	83 (37.9%) 109 (49.8%) 27 (12.3%)	89 (38.5%) 118 (51.1%) 24 (10.4%)	$\chi^2 = 0.4229$ p = 0.809, df = 2
Total	219 (100%)	231 (100%)	
Allele frequencies			
Allele	Male	Female	Statistical value
G A	275 (62.8%) 163 (37.2%)	296 (64.1%) 166 (35.9%)	$\chi^2 = 0.1598$ p = 0.689, df = 1

10. Study on the distribution of ZHX2 genotypes according to *Xmn*I genotypes in mild and severe cases with β°-thalassemia/Hb E subjects and high percentage of Hb F and low percentage of Hb F subjects of β°-thalassemia/Hb E patients

From the Tables 4.10 and 4.11, genotype distributions and allele frequencies for combination of both ZHX2 and *Xmn*I polymorphisms were significant different in severity and individuals with percentage of Hb F. The result from the Table 4.10 showed that there was significant association between ZHX2-GG together with *Xmn*I-CC and severe phenotype (OR=5.6667, 95%CI=2.1032-15.2678, p=0.000281). Similar result was observed with comparison of % Hb F groups in β°-thalassemia/Hb E (OR=5.7750, 95%CI=2.6264-12.6980, p<0.0001) (Table 4.11). There was significant association between ZHX2-GG and *Xmn*I-CC and low percentage of Hb F. These results suggested that these two polymorphisms have synergistic effect on risks of being severe and having low %Hb F among β°-thalassemia/Hb E patients.

Table 4.10 Distribution of ZHX2 genotypes according to *Xmn*I genotypes in mild cases and severe cases with β°-thalassemia/Hb E subjects

	ZHX2 - GG		ZHX2 - AG+AA	
	severe	mild	severe	mild
*Xmn*I - CC	24	6	32	11
*Xmn*I - TT+TC	36	51	58	82
Odds ratio	5.6667		4.1129	
95% CI	2.1032-15.2678		1.9174-8.8220	
χ^2	11.817 (df=1)		13.035 (df=1)	
p-value	0.0006 (p=0.000281 by Fisher 's exact test)		0.0003	

Table 4.11 Distribution of ZHX2 genotypes according to *Xmn*I genotypes in high percentage of Hb F and low percentage of Hb F subjects of β°-thalassemia/Hb E patients

	ZHX2- GG		ZHX2- AG+AA	
	low percentage of Hb F	high percentage of Hb F	low percentage of Hb F	high percentage of Hb F
*Xmn*I - CC	36	10	43	15
*Xmn*I - TT+TC	48	77	83	137
Odds ratio	5.7750		4.7317	
95% CI	2.6264-12.6980		2.4754-9.0448	
χ^2	19.812 (df=1)		23.108 (df=1)	
p-value	<0.0001		<0.0001	

11. The comparison of ZHX2 allele frequencies in different population

The allele frequencies of G779A polymorphism from Thai population in this study were compared with other racial population as shown in Table 4.12. Chi-square analysis was used to test the statistically significant differences in G779A allele frequencies among the different populations. The result showed differences in allelic frequencies between Thai and various populations. Caucasian population showed highly diverse allele frequencies comparing with that of Thai population. The G allele frequencies showed comparable extent among Asian populations.

In conclusion, these comparisons of G779A polymorphism of ZHX2 gene vary among different racial or ethnic groups. A possible explanation for the discrepancy in the ZHX2 polymorphisms among population could be due to random genetic drift or a selective mechanism.

Table 4.12 Comparison of G779A polymorphism in different populations.

Allele frequencies information in other populations							
Population Type	Population Name	Sample Size	frequencies chart	Allele frequencies		p-value	Submitter
				G	A		
	Thailand	450		**0.63**	**0.37**	**-**	**Present study**
NORTH AMERICA	AFD_EUR_PANEL	48		0.98	0.02	<0.0001	PERLEGEN
NORTH AMERICA	AFD_AFR_PANEL	44		1.00	0.00	<0.0001	PERLEGEN
NORTH AMERICA	AFD_CHN_PANEL	48		0.65	0.35	0.9133	PERLEGEN
EUROPE	HapMap-CEU	116		0.99	0.01	<0.0001	CSHL-HAPMAP
EAST ASIA	HapMap-HCB	88		0.61	0.39	0.5617	CSHL-HAPMAP
EAST ASIA	HapMap-JPT	84		0.66	0.34	0.5733	CSHL-HAPMAP
WEST AFRICA	HapMap-YRI	116		1.00	0.00	<0.0001	CSHL-HAPMAP
NORTH AMERICA	AGI_ASP population	78		0.97	0.03	<0.0001	APPLERA_GI
EUROPE	HapMap-CEU	120		0.98	0.02	<0.0001	CSHL-HAPMAP
EAST ASIA	HapMap-HCB	90		0.61	0.39	0.6118	CSHL-HAPMAP
EAST ASIA	HapMap-JPT	90		0.66	0.34	0.5519	CSHL-HAPMAP
WEST AFRICA	HapMap-YRI	120		0.98	0.02	<0.0001	CSHL-HAPMAP

http://www.genecards.org/cgi-bin/snps/snp_link.pl?rs_number=3802264&file=/home/genecards/versions/2.39build76/cards_usr/entries/ZH/card_ZHX2.txt;&kind=AlleleFreqData;&chrom=8

12. Detection of mRNA expression of ZHX2 by reverse transcriptase-polymerase chain reaction (RT-PCR)

The mRNA expression patterns of differentially expressed genes in cultured erythroid cells was investigated by reverse transcriptase-polymerase chain reaction (RT-PCR) using exon specific primers. Glyceraldehydes-3-phosphate dehydrogenase (GAPDH) was used as the endogenous control gene and the levels of expression were normalized to GAPDH expression. The expected RT-PCR product size of GAPDH and ZHX2 were 351 and 318 bp respectively. The differentiation of erythroid cells was observed in the erythroid cell culture from 1 mild and 1 severe case of β°-thalassemia/Hb E subjects. Preproerythroblasts, proerythroblasts, basophilic erythroblasts were identified as the major populations on days 7, 9, and 12, respectively. The expression levels showed differences on each day of culture (see figure 4.5, 4.6). ZHX2 expression in erythroid precursor cells is estimated to be relatively low, while expression in ovary, prostate, spleen, skeletal muscle and pancreas tissue is high (34). The human ZHX2 in mild case also shows similar expression patterns as observed in severe cases with β°-thalassemia/Hb E. These studies suggest that ZHX2 mRNA is transcribed in erythroid precursor cells, which confirms that the ZHX2 protein may involve in transcription in erythroid precursor cells and not only in the liver (39).

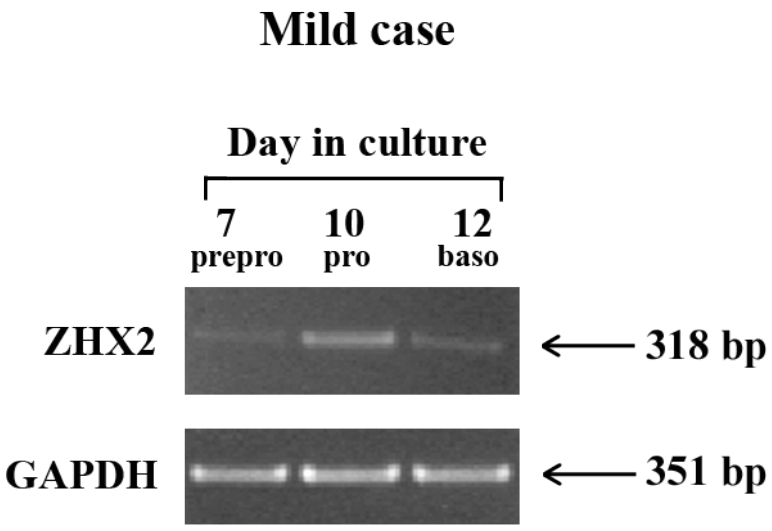


Figure 4.5 RT-PCR analysis of ZHX2 mRNA expression of mild case in agarose gel. The 351 bp and 318 bp bands corresponded to GAPDH and ZHX2. Lane 1, 2, 3 were ZHX2 and GAPDH expression at day 7, 10, 12 of culture, respectively; prepro indicates preproerythroblasts; pro, proerythroblasts; baso, basophilic erythroblasts.

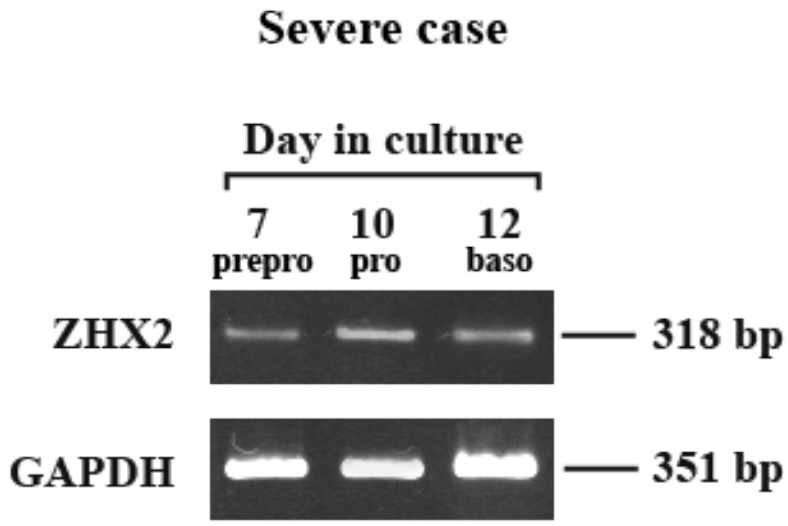


Figure 4.6 RT-PCR analysis of ZHX2 mRNA expression of severe case in agarose gel. The 351 bp and 318 bp bands corresponded to GAPDH and ZHX2. Lane 1, 2, 3 were ZHX2 and GAPDH expression at day 7, 10, 12 of culture, respectively; prepro indicates preproerythroblasts; pro, proerythroblasts; baso, basophilic erythroblasts.

CHAPTER V
DISCUSSION

β-thalassemia is an autosomal recessive genetic disease with many genes involved. It is a heterogeneous disorder caused by variations in the inactivation mechanism of the β-globin genes. Homozygous β-thalassemia and β°-thalassemia/hemoglobin (Hb) E are accounting for a majority of β-thalassemic syndromes in Southeast Asia. Despite seemingly genotypes, the patients, especially those with β°-thalassemia/Hb E disease, have a remarkable variability in anemia, growth development, hepatosplenomegaly, and transfusion requirements. The genetic factors may differ in each race or ethnic group and variations in genes may alter these parameters. Thus, it is very interesting to investigate the associations between these gene and β°-thalassemia/hemoglobin (Hb) E in each ethnic group for therapy and prevention. Despite remarkable successes in the treatment of β°-thalassemia/ hemoglobin (Hb) E in the past decades, it is still the leading cause of death and premature disability in developed and developing countries. Possible factors that influence the severity of anemia in thalassemia may be inherited or noninherited. The inherited factors include type of β-thalassemia mutation, coinheritance of α-thalassemia, and factors that stimulate Hb F production. Up to present time, it is still speculative whether genotypes could be predictive of phenotype.

Many genes have been proposed as candidates for a genetic cause of Hb F level. In recent years, SNPs have been used as genetic markers for the study of complex traits. Co-inheritance of α-thalassemia and homozygosity for *Xmn*I site polymorphism modify phenotype (148-150). Winichagoon *et al.*(151) found that mild phenotype may be seen even in the absence of detectable α-thalassemia and *Xmn*I +/+. A genetic interaction between the *Xmn*I-$^{G}\gamma$ site and a locus on chromosome 8q was reported to influence adult F-cell levels (28).

In our study, there was significant association between ZHX2 - GG and *Xmn*I - CC and low percentage of Hb F. These results suggested that these two polymorphisms have synergistic effect on severity and percentage of Hb F among β°-thalassemia/Hb E patients.

Here we provide evidence for the first time that G allele of G779A polymorphism of ZHX2 is associated with lessened Hb F levels among β°-thalassemia/Hb E patients with *Xmn*I-$^{G}\gamma^{-/-}$ and *Xmn*I-$^{G}\gamma^{+/-}$ polymorphisms. In this study, we revealed a no significant association between ZHX2 and percentage of Hb F in a β°-thalassemia/Hb E of Thai population. So these results indicate that ZHX2 variants, by modulating HbF levels, act as an important devious ameliorating factor of the β-thalassemia phenotype and it is likely they could help ameliorate other hemoglobin disorders. We expect our findings will help to characterize the molecular mechanisms of fetal globin regulation and could eventually contribute to the development of new therapeutic approaches for β-thalassemia.

CHAPTER VI
CONCLUSION

β°-thalassemia/Hb E is a monogenic disease however its phenotype results from several genetic factors. A variety in the clinical manifestation of β-thalassemic diseases may occur from the nature of β-globin gene mutations, interaction with α-thalassemia gene interaction, or differences in the amount of hemoglobin (Hb) F production. However, the biological mechanism of human globin gene regulation underlying developmental determination during embryogenesis is not yet fully defined. Studied polymorphisms may be involved in a modulation of a relative risk of being severe affected by other factors. Simultaneous studies of several genetic variants will allow the determination of genetic susceptibility to β°-thalassemia/Hb E and evaluation of severity. Tiago Gomes de Andrade et al.(37) proposed a test for candidate gene for globin regulation in erythroid cells. This study was conducted to determine differentially expressed transcripts in reticulocytes from a normal and a HPFH-2 subject, viz., these experiments observed the downregulation of ZHX2, a transcriptional repressor, in two HPFH-2 subjects which have a delayed switch from fetal to adult hemoglobin, resulting in high levels of HbF in the adult stage, without clinical manifestations. This demonstration can assess the relationship between the genotype and Hb F level of the disease leading to phenotypic severity of patients with β-thalassemia. In our study, we can conclude that:

1. Analyzing genotypes and allele frequencies of ZHX2 polymorphism at position 779, G substituted by A in Thai β°-thalassemia/Hb E patients, has shown remarkable similarity of with other Asian populations and exhibited no association with degree severity of β°-thalassemia/Hb E patients.

2. This study provides no evidence in supporting the association of G779A polymorphism of ZHX2 gene with Hb F level and severity of β°-thalassemia/Hb E

3. However, this preliminary study of ZHX2 gene expression during erythroid development does support its fetal-expressed gene suppression.

4. ZHX2 is indeed expressed at low level in early developmental stage of human erythroid culture cells.

5. It is possible that other genetic loci exist and provide larger effect on Hb F level, consequently are more important as risk factors for β°-thalassemia/Hb E.

REFERENCES

1. Lichtman MA, Kioos TJ, Kaushansky K, Beutler E, Seligsohn U, JT. P. Williams Hematology. 7th ed. USA: McGeaw-Hill companies, Inc.; 2006.
2. Eleftheriou A. About Thalassemia Nicosia: Thalassaemia International Federation; 2007.
3. Korf B. Human Genetics and Genomics. 3rd ed. USA: Blackwell Publishing; 2006.
4. Barbara JB. Haemoglobinopathy Diagnosis. Blackwell Publishing; 2006.
5. Winichagoon P, Higgs DR, Goodbourn SE, Clegg JB, Weatherall DJ, Wasi P. The molecular basis of alpha-thalassaemia in Thailand. EMBO J 1984;3:1813-8.
6. Weatherall DJ. Fortnightly review: The thalassaemias. BMJ 1997;314:1675.
7. Weatherall DJ, Clegg JB. The Thalassemia Syndrome. 4th ed. Oxford : Blackwell Science; 2001.
8. Thein SL, Winichagoon P, Hesketh C, Best S, Fucharoen S, Wasi P, et al. The molecular basis of beta-thalassemia in Thailand: application to prenatal diagnosis. Am J Hum Genet 1990;47:369-75.
9. Fucharoen S, Fucharoen G, Sriroongrueng W, Laosombat V, Jetsrisuparb A, Prasatkaew S, et al. Molecular basis of beta-thalassemia in Thailand: analysis of beta-thalassemia mutations using the polymerase chain reaction. Hum Genet 1989;84:41-6.
10. Weatherall DJ, Clegg JB. Inherited haemoglobin disorders: an increasing global health problem. Bull World Health Organ 2001;79:704-12.
11. Chakraborty D, Bhattacharyya M. Antioxidant defense status of red blood cells of patients with beta-thalassemia and Ebeta-thalassemia. Clin Chim Acta 2001;305:123-9.

12. Winichagoon P, Fucharoen S, Wilairat P, Chihara K, Fukumaki Y, Wasi P. Identification of five rare mutations including a novel frameshift mutation causing beta zero-thalassemia in Thai patients with beta zero-thalassemia/hemoglobin E disease. Biochim Biophys Acta 1992;1139:280-6.
13. Angastiniotis M, Modell B. Global epidemiology of hemoglobin disorders. Ann N Y Acad Sci 1998;850:251-69.
14. Fucharoen S, Winichagoon P. Hemoglobinopathies in Southeast Asia: molecular biology and clinical medicine. Hemoglobin 1997;21:299-319.
15. Fucharoen S, Winichagoon P, Pootrakul P. Variable severity of Southeast Asian β^0-thalassemia/Hb E disease. Birth Defects Orig Artic Ser 1998;23:241-8.
16. Winichagoon P, Thonglairoum V, Siriboon W. Severity differences in β-thalassemia/haemoglobin E syndromes: implication of genetic factors. Br J Haem 1993;83:633-9.
17. Cohen AR, Galanello R, Pennell DJ, Cunningham MJ, Vichinsky E. Thalassemia. Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. 2004:14-34.
18. Fucharoen S, Winichagoon P, Pootrakul P, Piankijagum A, Wasi P. Variable severity of Southeast Asian beta^0-thalassemia/Hb E disease. Birth Defects Orig Artic Ser 1987;23:241-8.
19. Fucharoen G, Fucharoen S, Sanchaisuriya K, Sae-Ung N, Suyasunanond U, Sriwilai P, et al. Frequency Distribution and Haplotypic Heterogeneity of beta(E)-Globin Gene among Eight Minority Groups of Northeast Thailand. Hum Hered 2002;53:18-22.
20. Suwanmanee T, Sierakowska H, Fucharoen S, Kole R. Repair of a splicing defect in erythroid cells form patients with β-thalassemia/HbE disorder. Mol Ther 2002;6:718-26.
21. Agarwal S, Gulati R, Singh K. Hemoglobin E-beta thalassemia in Uttar Pradesh. Indian Pediatr 1997;34:287-92.
22. Forget BG. Molecular basis of hereditary persistence of fetal hemoglobin. Ann N Y Acad Sci 1998;850:38-44.

23. Garner C, Tatu T, Reittie JE, Littlewood T, Darley J, Cervino S, et al. Genetic influences on F cells and other hematologic variables: a twin heritability study. Blood 2000;95:342-6.
24. Panigrahi I, Agarwal S, Gupta T, Singhal P, Pradhan M. Hemoglobin E-beta Thalassemia: Factors Affecting Phenotype. Indian Pediatr 2004;42:357-62.
25. Craig J, Rochette J, Fisher C, Weatherall DJ, Marc S, Lathrop G, et al. Dissecting the loci controlling fetal haemoglobin production on chromosomes 11p and 6q by the regressive approach. Nat Genet 1996;12:58–64.
26. Close J, Game L, Clark B, Bergounioux J, Gerovassili A, Thein SL. Genome annotation of a 1.5 Mb region of human chromosome 6q23 encompassing a quantitative trait locus for fetal hemoglobin expression in adults. BMC Genomics 2004;5:33-46.
27. Garner C, Silver N, Best S, Menzel S, Martin C, Spector TD, et al. Quantitative trait locus on chromosome 8q influences the switch from fetal to adult hemoglobin. Blood 2004;104:2184-6.
28. Garner C, Tatu T, Best S, Creary L, Thein S. Evidence of Genetic Interaction between the β-Globin Complex and Chromosome 8q in the Expression of Fetal Hemoglobin. Am J Hum Genet 2002;70:793-9.
29. Dover GJ, Smith KD, Chang YC, Purvis S, Mays A, Meyers DA, et al. Fetal hemoglobin levels in sickle cell disease and normal individuals are partially controlled by an X-linked gene located at Xp22.2. Blood 1992;80:816-24.
30. Garner C, Mitchell J, Hatzis T, Reittie J, Farrall M, Thein SL. Haplotype Mapping of a Major Quantitative-Trait Locus for Fetal Hemoglobin Production, on Chromosome 6q23. Am J Hum Genet 1998;62:1468-74.
31. Sampietro M, Thein SL, Contreras M, Pazmany L. Variation of HbF and F-cell number with the G-gamma Xmn I (C-T) polymorphism in normal individuals [letter]. Blood 1992;79:832-3.
32. Gilman JG, Huisman TH. DNA sequence variation associated with elevated fetal G gamma globin production. Blood 1985;66:783-7.

33. Groudine M, Kohwi-Shigematsu T, Gelinas R, Stamatoyannopoulos G, Papayannopoulou T. Human fetal to adult hemoglobin switching: changes in chromatin structure of the beta-globin gene locus. Proc Natl Acad Sci U S A 1983;80:7551-5.
34. Kawata H, Yamada K, Shou Z, Mizutani T, Yazawa T, Yoshino M, et al. Zinc-fingers and homeoboxes (ZHX) 2, a novel member of the ZHX family, functions as a transcriptional repressor. Biochem J 2003;373(Pt 3):747-57.
35. Kawata H, Yamada K, Shou Z, Mizutani T, Miyamoto K. The mouse zinc-fingers and homeoboxes (ZHX) family; ZHX2 forms a heterodimer with ZHX3. Gene 2003;323:133-40.
36. Mattaj IW, Englmeier L. Nucleocytoplasmic transport: The Soluble Phase. 1998. p. 265-306.
37. de Andrade TG, Peterson KR, Cunha AF, Moreira LS, Fattori A, Saad ST, et al. Identification of novel candidate genes for globin regulation in erythroid cells containing large deletions of the human beta-globin gene cluster. Blood Cells Mol Dis 2006;37:82-90.
38. Spear BT, Jin L, Ramasamy S, Dobierzewska A. Transcriptional control in the mammalian liver: liver development, perinatal repression, and zonal gene regulation. Cell Mol Life Sci 2006;63:2922-38.
39. Perincheri S, Dingle RW, Peterson ML, Spear BT. Hereditary persistence of alpha-fetoprotein and H19 expression in liver of BALB/cJ mice is due to a retrovirus insertion in the Zhx2 gene. Proc Natl Acad Sci U S A 2005 ;102:396-401.
40. Neildez-Nguyen TMA, Wajcman H, Marden MC, Bensidhoum M, Moncollin V, Giarratana M-C, et al. Human erythroid cells produced ex vivo at large scale differentiate into red blood cells in vivo. Nat Biotech 2002;20:467-72.
41. Steinberg MH. Disorders of hemoglobin. New York: Cambridge University Press; 2001.
42. Amoyal I, Fibach E. Hemoglobin Switch in the Newborn: A Flow Cytometry Analysis. Neonatology 2007;91:61-8.

43. Mader SS. Inquiry into Life. 10th ed. USA: McGraw-Hill Science / Engineering/Math; 2002.
44. Marengo-Rowe AJ. The thalassemias and related disorders. Proc (Bayl Univ Med Cent) 2007;20:27-31.
45. Peterson KR. Hemoglobin switching: new insights. Curr Opin Hematol 2003 ;10:123-9.
46. Olivieri NF. The beta-thalassemias. N Engl J Med 1999;341:99-109.
47. Stamatoyannopoulos G, Nienhuis AW, Majerus PW, Varmus H. The Molecular Basis of Blood Disease. 2nd ed. Philadephia: Elsevier; 1994.
48. Wood WG, Stamatoyannopoulos G, Lim G, Nute PE. F-cells in the adult: normal values and levels in individuals with hereditary and acquired elevations of Hb F. Blood 1975;46:671-82.
49. Weatherall DJ. Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. Nat Rev Genet 2001;2:245-55.
50. Moschonas N, de Boer E, Grosveld FG, Dahl HH, Wright S, Shewmaker CK, et al. Structure and expression of a cloned beta o thalassaemic globin gene. Nucleic Acids Res 1981;9:4391–401.
51. Eaton WA. The relationship between coding sequences and function in haemoglobin. Nature 1980;284:183-5.
52. Matthews CM, Trotman CNA. Ancient and Recent Intron Stability in the Artemia Hemoglobin Gene. J Mol Evol 1998;47:763-71.
53. Jarman AP, Wood WG, Sharpe JA, Gourdon G, Ayyub H, Higgs DR. Characterization of the major regulatory element upstream of the human alpha-globin gene cluster. Mol Cell Biol 1991;11:4679-89.
54. Tuan D, Solomon W, Li Q, London IM. The "beta-like-globin" gene domain in human erythroid cells. Proc Natl Acad Sci U S A 1985;82:6384-8.
55. Grosveld F, van Assendelft GB, Greaves DR, Kollias G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell 1987;51:975-85.
56. Higgs DR, Wood WG, Jarman AP, Sharpe J, Lida J, Pretorius IM, et al. A major positive regulatory region located far upstream of the human alpha-globin gene locus. Genes Dev 1990;4:1588-601.

57. Vakoc CR, Letting DL, Gheldof N, Sawado T, Bender MA, Groudine M, et al. Proximity among Distant Regulatory Elements at the [beta]-Globin Locus Requires GATA-1 and FOG-1. Mol Cell 2005;17:453-62.
58. Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W. The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet 2003;35:190-4.
59. Gribnau J, de Boer E, Trimborn T, Wijgerde M, Milot E, Grosveld F, et al. Chromatin interaction mechanism of transcriptional control in vivo. EMBO J 1998;17:6020-7.
60. Bank A. Regulation of human fetal hemoglobin: new players, new complexities. Blood 2006;107:435-43.
61. Patrinos GP, de Krom M, de Boer E, Langeveld A, Imam AMA, Strouboulis J, et al. Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. Genes Dev 2004;18:1495-509.
62. Ebb D, Tang DC, Drew L, Chin K, Berg PE, Rodgers GP. Identification of Upstream Regulatory Elements That Repress Expression of Adult Beta-like Globin Genes in a Primitive Erythroid Environment. Blood Cells Mol Dis 1998;24:356-69.
63. Trudel M, Costantini F. A 3' enhancer contributes to the stage-specific expression of the human beta-globin gene. Genes Dev 1987;1:954-61.
64. Langdon SD, Kaufman RE. Gamma-Globin Gene Promoter Elements Required for Interaction With Globin Enhancers. Blood 1998;91:309-18.
65. Ristaldi MS, Casula S, Porcu S, Marongiu MF, Pirastu M, Cao A. Activation of the [delta]-Globin Gene by the [beta]-Globin Gene CACCC Motif. Blood Cells Mol Dis 1999;25:193-209.
66. Charnay P, Mellon P, Maniatis T. Linker scanning mutagenesis of the 5'-flanking region of the mouse beta-major-globin gene: sequence requirements for transcription in erythroid and nonerythroid cells. Mol Cell Biol 1985;5:1498-511.
67. Stamatoyannopoulos G, Blau CA, Nakamoto B, Josephson B, Li Q, Liakopoulou E, et al. Fetal hemoglobin induction by acetate, a product of butyrate catabolism. Blood 1994;84:3198-204.

68. Choi OR, Engel JD. Developmental regulation of beta-globin gene switching. Cell 1988;55:17-26.
69. Bulger M, Schubeler D, Bender MA, Hamilton J, Farrell CM, Hardison RC, et al. A Complex Chromatin Landscape Revealed by Patterns of Nuclease Sensitivity and Histone Modification within the Mouse {beta}-Globin Locus. Mol Cell Biol 2003;23:5234-44.
70. Tolhuis B, Palstra R-J, Splinter E, Grosveld F, de Laat W. Looping and Interaction between Hypersensitive Sites in the Active [beta]-globin Locus. Mol Cell 2002;10:1453-65.
71. Bulger M, Bender MA, van Doorninck JH, Wertman B, Farrell CM, Felsenfeld G, et al. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters. Proc Natl Acad Sci U S A 2000;97:14560-5.
72. Grosveld FG, Stamatoyannopoulos G. Hemoglobin Switching. 3rd ed. Philadelphia: W.B. Saunders Publishing Co.; 2001.
73. Tsai SF. Cloning of cDNA for the major DNAbinding protein of the erythroid lineage through expression in mammalian cells. Nature 1989;339:446-51.
74. Evans T, Felsenfeld G. The erythroid-specific transcription factor eryfl: A new finger protein. Cell 1989;58:877-85.
75. Orkin SH. GATA-binding transcription factors in hematopoietic cells. Blood 1992;80:575-81.
76. Strauss EC, Orkin SH. In vivo Protein-DNA Interactions of Hypersensitive Site 3 of the Human {beta}-Globin Locus Control Region. Proc Natl Acad Sci U S A 1992;89:5809-13.
77. Cao A, Moi P. Regulation of the globin genes. Pediatr Res 2002;51:415-21.
78. Blobel GA, Simon MC, Orkin SH. Rescue of GATA-1-deficient embryonic stem cells by heterologous GATA - binding proteins. Mol Cell Biol 1995;15:626-33.
79. Weiss MJ, Keller G, Orkin SH. Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. Genes Dev 1994;8(10):1184-97.

80. Pevny L, Simon MC, Robertson E, Klein WH, Tsai S-F, D'Agati V, et al. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. Nature 1991;349:257-60.
81. Drissen R, Palstra RJ, Gillemans N, Splinter E, Grosveld F, Philipsen S, et al. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. Genes Dev 2004;18:2485-90.
82. Miller IJ, Bieker JJ. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. Mol Cell Biol 1993;13:2776-86.
83. Shyu YC, Lee TL, Wen SC, Chen H, Hsiao WY, Chen X, et al. Subcellular transport of EKLF and switch-on of murine adult beta maj globin gene transcription. Mol Cell Biol 2007;27:2309-23.
84. Wani MA, Means RT, Jr., Lingrel JB. Loss of LKLF function results in embryonic lethality in mice. Transgenic Res 1998;7:229-38.
85. Andrews NC, Erdjument-Bromage H, Davidson MB, Tempst P, Orkin SH. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. Nature 1993;362:722-8.
86. Martin F, van Deursen JM, Shivdasani RA, Jackson CW, Troutman AG, Ney PA. Erythroid Maturation and Globin Gene Expression in Mice With Combined Deficiency of NF-E2 and Nrf-2. Blood 1998;91:3459-66.
87. Shivdasani RA, Orkin SH. Erythropoiesis and Globin Gene Expression in Mice Lacking the Transcription Factor NF-E2. Proc Natl Acad Sci 1995;92:8690-4.
88. Yang Z, Engel JD. Biochemical characterization of the developmental stage- and tissue- specific erythroid transcription factor, NF-E4. J Biol Chem 1994;269:10079-87.
89. Zhou W, Zhao Q, Sutton R, Cumming H, Wang X, Cerruti L, et al. The role of P22 NF-E4 in human globin gene switching. J Biol Chem 2004:M402191200.

90. Zhao Q, Cumming H, Cerruti L, Cunningham JM, Jane SM. Site-specific Acetylation of the Fetal Globin Activator NF-E4 Prevents Its Ubiquitination and Regulates Its Interaction with the Histone Deacetylase, HDAC1. J Biol Chem 2004;279:41477-86.
91. Weatherall DJ. Single gene disorders or complex traits: lessons from the thalassemias and other monogenic diseases. BMJ 2000;321:1117-20.
92. Bentley D. The human genome project-an overview. Med Res Rev 2000;20:189-96.
93. Barnes MR, Gray IC. Bioinformatics for Geneticists. Wiley, English. 2003.
94. Brookes AJ. The essence of SNPs. Gene 1999;234:177.
95. Weatherall DJ, Clegg JB. The Thalassemia Syndrome, 4th ed. Blackwell Science, Oxford. 2001.
96. Chakravarti A. Single nucleotide polymorphism to a future of genetic medicine. Nature 2001;409:822-3.
97. Riva A, Kohana IS. SNPper: retrieval and analysis of human SNPs. Bioinformatics 2002;18:1681-5.
98. Strittmatter WJ, Roses AD. Apolipoprotein E and Alzheimer's disease. Annu Rev Neurosci 1996;19:53-77.
99. Sripichai O, Fucharoen S. Genetic polymorphisms and implications for human diseases. J Med Assoc Thai 2007;90:394-8.
100. International HapMap Consortium. The International HapMap Project. Nature 2003;426:789-96.
101. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. Genet Med 2002;4:45-61.
102. Tuzmen S, Schechter AN. Genetic diseases of hemoglobin: diagnostic methods for elucidating beta-thalassemia mutations. Blood Rev. 2001;15:19-29.
103. Weatherall DJ. ABC of clinical haematology. The hereditary anaemias. BMJ 1997;314:492-6.
104. Barbara J B. Haemoglobinopathy Diagnosis. Blackwell Publishing; 2006.
105. Anand R, Boehm CD, Kazazian HH, Jr., Vanin EF. Molecular characterization of a beta zero-thalassemia resulting from a 1.4 kilobase deletion. Blood 1988;72:636-41.

106. Thein SL, Al-Hakim I, Hoffbrand AV. Thalassaemia intermedia: a new molecular basis. Br J Haematol 1984;56:333-7.
107. Aksoy M. Thalassaemia intermedia: a genetic study in 11 patients. J Med Genet 1970;7:47-51.
108. Colosimo A, Guida V, De Luca A, Cappabianca MP, Bianco I, Palka G, et al. Reliability of DHPLC in mutational screening of beta-globin (HBB) alleles. Hum Mutat 2002;19:287-95.
109. Ambekar SS, Phadke MA, Balpande DN, Mokashi GD, Khedkar VA, Bankar MP, et al. The Prevalence and Heterogeneity of Beta Thalassemia Mutations in The Western Maharashtra Population: A Hospital Based Study. IJHG 2001;1:219-23.
110. Leoncini-Franscini L. Book review. Ann Oncol 1999;10:1270.
111. Weatherall DJ. Pathophysiology of thalassaemia. Baillieres Clin Haematol 1998;11:127-46.
112. Higgs DR, Thein SL, Woods WG. The molecular pathology of the thalassaemias. In: Weatherall DJ, Clegg B, eds. The thalassaemia syndromes. 4th ed. Oxford: Blackwell Science ; 2001.p.133-91.
113. Fucharoen S, Winichagoon P. Hemoglobinopathies in Southeast Asia. Hemoglobin 1987;11:65-88.
114. Clegg JB, Weatherall DJ, Milner PF. Haemoglobin Constant Spring[mdash]A Chain Termination Mutant. Nature 1971;234:337-40.
115. D. M. Hunt DRHPWJBCDJW. Haemoglobin Constant Spring has an unstable alpha-chain messenger RNA. Br J Haematol 1982;51:405-13.
116. Fucharoen S, Winichagoon P. Hemoglobinopathies in Southeast Asia: Molecular Biology and Clinical Medicine. Hemoglobin 1997;21:299-319.
117. Fucharoen S, Ketvichit P, Pootrakul P. Clinical manifestration of β-thalassemia/hemoglobin E disease. J Pediatr Hemato Oncol. 2000;22:552-7.
118. Weatherall DJ, Clegg JB. Inherited haemoglobin disorders: an increasing global health problem. Bull World Health Organ 2001;79:704-12.

119. Swee Lay T. Genetic insights into the clinical diversity of β thalassaemia. Br J Haematol 2004;124:264-74.

120. Kudo S, Harigae H, Watanabe N, Takasawa N, Kimura J, Kameoka J-i, et al. Increased HbF levels in dyserythropoiesis. Clinica Chimica Acta 2000;291:83-7.

121. Merghoub T, Perichon B, Maier-Redelsperger M, Dibenedetto SP, Samperi P, Ducrocq R, et al. Dissection of the association status of two polymorphisms in the β-globin gene cluster with variations in F-cell number in non-anemic individuals. Am J Hematol 1997;56:239-43.

122. Jiang J, Best S, Menzel S, Silver N, Lai MI, Surdulescu GL, et al. cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood 2006;108:1077-83.

123. Menzel S, Jiang J, Silver N, Gallagher J, Cunningham J, Surdulescu G, et al. The HBS1L-MYB intergenic region on chromosome 6q23.3 influences erythrocyte, platelet, and monocyte counts in humans. Blood 2007;110:3624-6.

124. Pandit RA, Svasti S, Sripichai O, Munkongdee T, Triwitayakorn K, Winichagoon P, et al. Association of SNP in exon 1 of HBS1L with hemoglobin F level in beta(0)-thalassemia/hemoglobin E. Int J Hematol 2008.

125. Miyoshi K, Kaneto Y, Kawai H, Ohchi H, Niki S, Hasegawa K, et al. X-linked dominant control of F-cells in normal adult life: characterization of the Swiss type as hereditary persistence of fetal hemoglobin regulated dominantly by gene(s) on X chromosome. Blood 1988;72:1854-60.

126. Chang YC, Smith KD, Moore RD, Serjeant GR, Dover GJ. An analysis of fetal hemoglobin variation in sickle cell disease: the relative contributions of the X-linked factor, beta-globin haplotypes, alpha-globin gene number, gender, and age. Blood 1995;85:1111-7.

127. Belayew A, Tilghman SM. Genetic analysis of α-fetoprotein synthesis in mice. Mol Cell Biol 1982;2:1427-35.

128. Morford LA, Davis C, Jin L, Dobierzewska A, Peterson ML, Spear BT. The oncofetal gene glypican 3 is regulated in the postnatal liver by zinc fingers and homeoboxes 2 and in the regenerating liver by alpha-fetoprotein regulator 2. Hepatology 2007;46:1541-7.
129. Gehring WJ, Affolter M, Burglin T. Homeodomain proteins. Annu Rev Biochem 1994;63:487-526.
130. Yamada K, Printz RL, Osawa H, Granner DK. Human ZHX1: cloning, chromosomal location, and interaction with transcription factor NF-Y. Biochem Biophys Res Commun 1999 ;261:614-21.
131. Mantovani R. The molecular biology of the CCAAT-binding factor NF-Y. Gene 1999;239:15-27.
132. Yamada K, Kawata H, Shou Z, Hirano S, Mizutani T, Yazawa T, et al. Analysis of zinc-fingers and homeoboxes (ZHX)-1-interacting proteins: molecular cloning and characterization of a member of the ZHX family, ZHX3. Biochem J 2003;373(Pt 1):167-78.
133. Hirano S, Yamada K, Kawata H, Shou Z, Mizutani T, Yazawa T, et al. Rat zinc-fingers and homeoboxes 1 (ZHX1), a nuclear factor-YA-interacting nuclear protein, forms a homodimer. Gene 2002;290:107-14.
134. Hill RE, Jones PF, Rees AR, Sime CM, Justice MJ, Copeland NG, et al. A new family of mouse homeo box-containing genes: molecular structure, chromosomal location, and developmental expression of Hox-7.1. Genes Dev 1989;3:26-37.
135. Barthelemy I, Carramolino L, Gutierrez J, Barbero JL, Marquez G, Zaballos A. zhx-1: a novel mouse homeodomain protein containing two zinc-fingers and five homeodomains. Biochem Biophys Res Commun 1996;224:870-6.
136. Huang M, Qian F, Hu Y, Ang C, Li Z, Wen Z. Chromatin-remodelling factor BRG1 selectively activates a subset of interferon-alpha-inducible genes. Nat Cell Biol 2002;4:774-81.

137. Sripichai O, Whitacre J, Munkongdee T, Kumkhaek C, Makarasara W, Winichagoon P, et al. Genetic Analysis of Candidate Modifier Polymorphisms in Hb E-{beta}0-Thalassemia Patients. Ann N Y Acad Sci. 2005;1054:433-8.

138. Fibach E, Rachmilewitz EA. The two-step liquid culture: a novel procedure for studying maturation of human normal and pathological erythroid precursors. Stem Cells 1993;11 Suppl 1:36-41.

139. Aoyagi K. PCR. In: Gerstein AS, editor. Molecular biology Problem Solver. New York: Wiley-Liss, Inc.; 2001. p. 291-329.

140. Weaver RF. Molecular biology. 2nd ed. New York: McGraw-Hill Companies,Inc.; 2003.

141. Rapley R. Molecular biology and basic techniques. In: Wilson K, Walker J, editors. Principles and Techniques of Practical Biochemistry. 5th ed. Cambridge University Press; 2000.

142. Rolfs A, Schuller I, Finckh U, Weber-Rolfs. I. PCR. In: Clinical Diagnostics and Research. Heidelberg: Springer-Verlag; 1992.

143. Nicholl DST. An Introduction to Genetic Engineering. 2nd ed. Cambridge, UK: Cambridge University Press; 2002.

144. Sawhney SK, Singh R. Introductory Practical Biochemistry. 2nd ed. UK: Alpha Science International, Ltd.; 2005.

145. Turner PC, McLennan AG, Bates AD, White MRH. Instant Notes in Molecular Biology. 2nd ed. UK: Bios Scientific Publishers Ltd.; 2000.

146. O'Connell J. RT-PCR Protocols. In: Walker JM, editor. Totowa, New Jersey: Humana Press Inc.; 2002.

147. Santos CFd, Sakai VT, Machado MAdAM, Schippers DN, Greene AS. Reverse transcription and polymerase chain reaction: principles and applications in dentistry. J Appl Oral Sci 2004;12:1-11.

148. Fucharoen S, Winichagoon P. Clinical and hematologic aspects of hemoglobin E beta-thalassemia. Curr Opin Hematol 2000;7:106-12.

149. Nadkarni A, Ghosh K, Gorakshakar A, Colah R, Mohanty D. Variable clinical severity of Hb E beta-thalassemia among Indians. J Assoc Physicians India 1999;47:966-8.

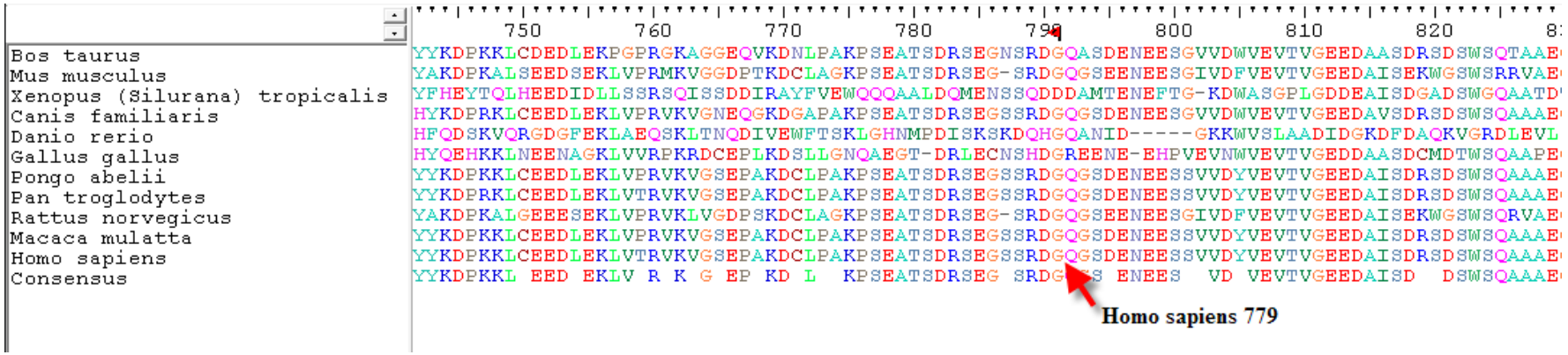
150. Other significant hemoglobinopathies. In: Bain BJ, editor. Hemoglobinopathy diagnosis. 1st ed. London: Blackwell Science; 2001. p. 154-86.

151. Winichagoon P, Fucharoen S, Chen P, Wasi P. Genetic factors affecting clinical severity in beta-thalassemia syndromes. J Pediatr Hematol Oncol 2000;22:573-80.

APPENDIX

APPENDIX A
MULTIPLE SEQUENCE ALIGNMENT

Appendix A: Multiple sequence alignment



APPENDIX B

CLINICAL DATA AND BIOCHEMICAL CHARACTERISTICS OF β°-THALASSEMIA/Hb E PATIENTS

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
1	002N	mild	female	19	168	55	8.5	25.4	3.49	41.10	0.76	TC	AG
2	006M	mild	male	22	157	45	8.2	24.7	2.68	36.78	0.75	CC	AG
3	010N	mild	female	20	155	48	8.0	27.2	3.46	43.20	0.85	TC	AG
4	018N	mild	female	26	155	51	8.6	26.3	0.87	10.10	0.13	TC	AG
5	019N	mild	female	10	131	25	7.2	23.0	3.40	47.17	0.91	TT	AA
6	020N	mild	male	6	112	16	7.7	21.7	3.64	47.30	1.06	TC	GG
7	030N	mild	female	14	148	38	8.4	25.7	4.19	49.90	1.16	TC	GG
8	043N	mild	female	27	154	41	5.7	21.5	2.11	37.00	0.64	TC	GG
9	047N	mild	male	4	105	18	7.8	29.5	3.12	39.98	1.03	TC	GG
10	048N	mild	male	7	113	19	8.7	27.1	3.78	43.50	0.85	TT	AG
11	059N	mild	male	12	136	28	6.9	23.0	1.92	27.84	0.48	TC	GG
12	067R	mild	female	28	154	47	6.1	18.9	2.92	47.80	0.95	TC	GG
13	068R	mild	female	28	165	51	5.6	17.9	2.35	41.93	0.76	TC	AG
14	070N	mild	female	23	167	58	9.5	28.7	4.29	45.20	0.92	TC	AG
15	073R	mild	female	22	168	69	7.4	24.9	2.45	33.13	0.52	TC	AA
16	1001S	mild	male	19	177	69	8.1	24.3	1.93	23.77	0.32	TC	GG
17	1005S	mild	female	32	155	40	4.2	14.3	0.75	17.79	0.23	TC	AG
18	1007S	mild	female	52	158	45	6.4	19.6	2.25	35.14	0.64	TC	AG
19	1011S	mild	male	26	171	62	6.8	21.6	1.91	28.02	0.40	TC	AG
20	1015S	mild	female	27	158	54	5.2	23.7	0.87	16.77	0.54	TC	AG
21	1016S	mild	female	36	151	46	9.3	27.7	3.84	41.26	0.73	TC	AG
22	1020R	mild	female	16	158	47	6.7	21.2	3.70	55.28	1.25	TC	AG
23	102M	mild	male	32	170	70	9.3	16.0	3.36	36.10	0.63	TC	AG
24	1033S	mild	female	27	172	58	7.7	24.0	2.77	35.92	0.58	TC	AA
25	1039S	mild	male	40	169	65	12.1	36.6	6.57	54.27	1.21	TC	AG
26	1040S	mild	male	23	170	50	7.9	25.3	3.16	40.06	0.69	TC	GG
27	1041S	mild	female	54	160	41	8.2	24.9	3.80	46.31	0.89	TT	AG

Appendix B:Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
28	1044S	mild	female	32	155	45	7.3	23.0	3.88	53.17	1.17	TC	GG
29	1047M	mild	female	41	158	49	8.5	26.9	3.14	36.91	0.61	CC	AG
30	1049N	mild	male	32	170	55	7.0	23.4	2.35	33.61	0.52	TC	GG
31	104N	mild	female	2	84	10	7.2	24.0	3.37	46.87	0.92	TC	AA
32	106N	mild	male	11	135	29	8.0	19.1	3.65	45.60	0.88	TC	AG
33	107N	mild	female	12	151	40	8.7	25.1	3.44	39.50	0.70	TC	GG
34	118M	mild	male	33	170	55	7.8	23.9	2.64	33.85	0.53	TC	AG
35	120R	mild	female	29	165	46	6.8	23.0	0.87	12.87	0.16	CC	AG
36	124M	mild	male	22	158	45	7.3	22.7	3.60	43.68	0.72	TC	GG
37	125M	mild	female	11	125	22	7.8	25.3	3.22	41.34	0.73	TC	GG
38	135N	mild	female	8	120	19	7.8	24.3	2.98	38.16	0.65	TC	AA
39	139R	mild	female	31	150	42	5.1	18.5	2.25	44.19	0.82	TC	GG
40	145R	mild	male	7	121	24	8.0	26.1	2.99	37.32	0.61	TC	AG
41	146R	mild	male	4	100	20	7.0	25.2	3.00	42.81	0.83	TC	AG
42	148R	mild	female	10	122	22	7.9	26.8	3.54	44.80	0.84	TC	AG
43	149R	mild	female	7	119	21	7.2	24.5	2.50	34.71	0.56	TC	GG
44	151M	mild	female	45	145	43	6.2	20.9	1.67	26.94	0.38	TC	AG
45	161C	mild	male	8	133	29	8.3	25.3	3.01	36.24	0.58	TT	AG
46	168C	mild	female	9	122	20	7.9	25.0	4.14	52.36	1.15	TC	AG
47	188CR	mild	female	5	103	17	6.2	21.6	2.96	47.78	0.94	TC	GG
48	194CR	mild	female	6	104	16	8.5	29.5	3.14	36.97	0.60	TC	AG
49	204CR	mild	female	14	147	38	8.3	27.9	4.73	57.04	1.37	TC	AG
50	206CR	mild	female	8	130	26	8.9	29.8	3.58	40.22	0.69	TT	GG
51	214CR	mild	female	2	82	12	8.0	26.8	3.34	41.69	0.74	CC	AG
52	220CR	mild	female	12	139	39	7.4	26.2	1.23	16.61	0.21	TC	AA
53	222CR	mild	female	10	130	29	7.4	25.9	3.95	53.39	1.18	TC	AG
54	225CR	mild	male	14	137	26	7.9	27.2	3.36	42.51	0.76	TC	AG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
55	231CR	mild	female	7	120	21	8.5	29.2	3.60	42.35	0.79	TC	GG
56	244CR	mild	female	17	161	60	7.5	26.5	1.82	24.26	0.33	CC	AG
57	254CR	mild	male	10	127	29	6.7	22.2	2.38	35.56	0.57	TT	GG
58	257CR	mild	male	10	132	26	8.1	27.5	3.66	45.22	0.85	TC	AG
59	259CR	mild	female	12	144	33	9.4	30.9	2.59	27.59	0.39	TT	GG
60	270CR	mild	male	12	140	25	8.6	28.4	4.28	49.80	1.02	TC	AA
61	278C	mild	male	2	88	10	8.6	27.0	4.97	57.81	1.51	TC	AA
62	286C	mild	female	13	152	37	9.3	31.0	3.98	42.84	0.79	TC	AG
63	306C	mild	female	11	138	26	6.8	22.0	2.15	31.62	0.50	CC	GG
64	342C	mild	male	12	152	56	9.0	33.0	2.56	28.39	0.41	TC	GG
65	348R	mild	female	34	163	51	7.5	25.5	2.28	30.41	0.47	TC	GG
66	355R	mild	female	13	152	45	8.4	26.9	3.82	45.51	0.86	TC	AG
67	356R	mild	male	16	165	52	8.8	29.0	3.28	37.26	0.61	TC	AA
68	357R	mild	male	4	101	14	9.7	32.1	6.14	63.30	1.78	TC	GG
69	359M	mild	female	40	160	47	8.2	28.6	2.08	25.31	0.35	TC	GG
70	360R	mild	female	13	153	36	7.9	25.0	4.51	57.08	1.38	TC	AG
71	372NR	mild	female	13	145	37	8.0	26.7	3.06	38.22	0.65	TC	GG
72	377NR	mild	male	6	116	20	6.7	24.1	2.44	36.35	0.59	TC	AG
73	380NR	mild	female	8	130	25	6.7	23.3	1.50	22.44	0.30	CC	AG
74	393NR	mild	female	8	123	22	8.0	26.7	3.71	46.35	0.90	TC	GG
75	399NR	mild	female	14	151	34	7.4	26.6	2.25	30.42	0.46	TC	AG
76	408NR	mild	male	30	160	53	8.6	28.9	3.85	44.80	0.84	TC	AG
77	411NR	mild	female	15	152	53	9.0	30.9	3.38	37.55	0.63	TC	AG
78	420NR	mild	female	34	142	44	6.6	23.2	2.73	41.35	0.73	TC	AG
79	428NR	mild	female	33	141	51	6.5	22.5	3.25	50.00	1.08	TC	GG
80	429NR	mild	female	46	155	45	6.7	23.6	2.02	30.18	0.45	TC	GG
81	435NR	mild	female	30	153	48	6.4	21.9	2.94	45.90	0.88	TC	AG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
82	440R	mild	female	42	158	45	6.3	19.0	2.64	41.98	0.75	CC	GG
83	442R	mild	male	24	170	56	8.1	27.1	3.24	40.06	0.69	TC	AG
84	448R	mild	female	15	152	52	8.2	26.4	4.18	50.94	1.07	TC	AG
85	450R	mild	female	6	109	17	7.8	24.0	2.65	34.00	1.08	TT	AG
86	458R	mild	male	3	97	16	7.8	24.3	3.61	46.26	0.89	TC	GG
87	461R	mild	female	14	152	38	7.8	25.0	3.33	42.71	0.77	TC	GG
88	462R	mild	female	20	155	48	10.7	33.3	5.78	54.00	1.21	TC	AG
89	478N	mild	male	22	169	55	7.8	26.5	1.47	18.83	0.24	CC	GG
90	486C	mild	female	5	105	14	6.3	22.0	2.31	36.69	0.60	TC	AG
91	492R	mild	male	14	163	42	7.5	24.4	2.71	36.10	0.58	TC	GG
92	493R	mild	female	16	144	33	7.3	23.5	2.46	33.68	0.53	TC	AG
93	496R	mild	female	12	158	40	7.4	22.5	3.50	47.28	0.93	TC	AG
94	498R	mild	female	17	160	47	7.2	23.3	1.98	27.45	0.39	CC	GG
95	501R	mild	female	7	121	23	6.7	21.8	1.84	27.53	0.40	TC	AG
96	506CB	mild	male	4	103	16	8.1	25.3	2.22	27.45	0.39	TT	AG
97	512CB	mild	female	11	145	32	7.8	25.0	2.80	35.86	0.57	TC	AG
98	513CB	mild	male	15	159	48	8.2	26.0	1.97	23.98	0.32	TC	AG
99	525R	mild	male	13	152	66	10.6	31.3	5.92	55.88	1.31	TC	AG
100	527R	mild	female	8	151	44	7.9	25.4	3.99	50.50	1.05	TC	AG
101	528R	mild	female	23	150	50	6.3	19.6	2.57	40.83	0.71	CC	AA
102	537SB	mild	female	19	156	45	8.4	26.2	3.28	39.08	0.67	TT	AG
103	541SB	mild	female	37	159	43	7.6	24.3	2.76	36.38	0.92	TC	AA
104	543SB	mild	male	35	173	52	6.5	21.1	1.37	21.12	0.28	TC	GG
105	546SB	mild	female	29	167	53	7.6	23.2	3.29	43.23	0.78	TC	AG
106	547SB	mild	female	38	163	57	7.0	22.5	3.55	50.64	1.07	TC	GG
107	548SB	mild	male	33	159	41	9.8	30.2	6.33	64.61	1.87	TC	GG
108	551SB	mild	male	15	156	41	8.0	25.5	3.65	45.68	0.86	TC	AG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
109	566SB	mild	male	13	132	26	7.7	24.0	4.54	58.98	1.49	TC	GG
110	571SB	mild	male	35	173	54	8.5	24.8	3.92	46.10	0.88	TT	AG
111	585SB	mild	female	4	98	13	7.8	22.7	4.30	55.18	1.27	CC	GG
112	592SB	mild	female	12	140	32	8.7	26.2	3.38	38.87	0.66	TC	AG
113	596SB	mild	male	12	142	27	9.0	27.4	3.49	38.79	0.65	TT	AG
114	604SB	mild	female	32	149	48	9.1	27.1	4.42	48.62	0.98	TC	AG
115	618C	mild	female	38	160	53	9.0	28.0	4.62	51.39	1.09	TC	GG
116	619C	mild	female	7	128	27	8.6	27.0	3.24	37.73	0.63	TT	AG
117	624CB	mild	male	16	156	40	6.8	21.1	2.80	41.17	0.72	TC	GG
118	626CB	mild	male	14	144	38	8.2	25.3	3.36	41.01	0.72	TC	GG
119	631CB	mild	female	6	106	15	6.7	21.1	3.35	49.95	1.03	TC	AA
120	637CB	mild	male	16	160	43	6.0	20.9	0.84	14.05	0.17	CC	AG
121	649CB	mild	male	17	151	40	10.8	32.0	5.89	54.51	1.24	TT	AG
122	651R	mild	female	2	75	9	7.6	24.2	3.45	45.45	0.86	TC	GG
123	665LP	mild	male	12	125	25	7.4	23.2	2.80	37.88	0.63	TT	AA
124	671LP	mild	male	8	120	22	6.8	21.2	2.77	40.75	0.71	TT	GG
125	687LP	mild	female	10	135	27	7.0	22.7	1.62	23.08	0.31	CC	GG
126	699LP	mild	male	11	133	24	8.6	28.2	3.97	46.20	0.88	TC	AA
127	707LP	mild	female	9	123	24	10	30.4	5.99	59.88	1.54	TC	AG
128	719C	mild	female	6	112	24	8.9	26.0	3.24	36.45	0.60	TC	AG
129	723R	mild	male	12	139	29	7.1	23.1	3.14	44.28	0.83	TC	AA
130	726UT	mild	male	19	167	46	9.2	28.6	3.70	40.20	0.70	CC	AA
131	733UT	mild	male	7	116	18	9.1	28.3	3.67	40.31	1.03	TC	AG
132	735UT	mild	male	5	104	15	7.6	24.0	3.32	43.65	0.81	TC	AG
133	741UT	mild	male	10	130	24	8.2	25.5	4.23	51.62	1.10	TC	AG
134	765UT	mild	male	12	126	22	7.0	23.2	2.03	28.99	0.43	TC	AG
135	825NT	mild	male	13	150	34	8.0	25.1	2.69	33.64	0.52	TC	AG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
136	834C	mild	male	12	152	37	7.6	25.0	2.39	31.41	0.50	TC	GG
137	854N	mild	male	23	175	63	7.6	23.2	3.25	42.74	0.76	TC	GG
138	863N	mild	male	3	92	14	8.3	25.0	4.59	55.35	1.27	TC	GG
139	878S	mild	male	38	167	67	7.3	23.3	2.43	33.30	0.51	CC	AG
140	891S	mild	male	47	160	62	9.9	28.6	6.50	65.69	1.99	TC	AA
141	892S	mild	male	25	165	50	9.7	31.0	3.57	36.80	0.60	TC	GG
142	897S	mild	male	23	173	58	8.6	25.7	4.65	54.03	1.22	TC	GG
143	898S	mild	male	14	144	30	7.7	23.9	3.14	40.71	0.71	TC	AG
144	899S	mild	male	28	157	50	8.3	26.2	2.54	30.61	0.46	TC	GG
145	936S	mild	male	17	174	56	7.8	24.2	1.82	23.35	0.33	TC	GG
146	938N	mild	male	26	172	56	8.7	26.0	2.74	31.46	0.47	TT	GG
147	948S	mild	male	24	168	51	10	28.7	5.05	50.54	1.04	TC	GG
148	961S	mild	male	24	168	50	8.0	25.7	2.30	28.70	0.41	TC	GG
149	969R	mild	male	3	97	17	8.8	28.0	3.59	40.79	0.71	TT	GG
150	998S	mild	male	33	167	47	6.0	19.2	1.52	25.28	0.35	CC	AG
151	009N	moderate	female	22	157	53	6.2	22.9	2.74	44.15	0.81	TC	GG
152	013M	moderate	male	52	166	46	4.8	16.6	0.29	6.10	0.07	CC	AG
153	014M	moderate	female	36	156	49	6.5	22.2	2.91	44.70	0.94	TC	GG
154	022M	moderate	female	17	160	56	7.1	24.1	2.41	33.90	0.59	TC	AG
155	023M	moderate	female	32	167	65	7.4	21.6	2.25	30.38	0.44	TC	AG
156	024N	moderate	male	38	160	43	6.4	21.5	1.74	27.18	0.38	TC	AG
157	031N	moderate	female	27	159	48	6.5	21.2	1.40	21.60	0.29	TC	GG
158	032M	moderate	female	54	162	50	5.6	18.6	1.67	29.80	0.48	TC	GG
159	036M	moderate	male	40	162	75	4.9	16.7	1.48	30.20	0.46	TC	AG
160	042M	moderate	male	50	155	39	4.9	16.5	1.74	35.60	0.56	TC	GG
161	045N	moderate	male	18	175	58	7.2	25.4	2.31	32.03	0.50	TC	AA
162	050M	moderate	male	27	166	60	6.3	20.6	1.42	22.55	0.30	TC	AG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
163	052S	moderate	female	47	155	45	5.3	18.0	1.77	33.40	0.55	CC	AA
164	053S	moderate	female	35	165	49	5.9	21.6	1.92	32.50	0.52	TC	AG
165	054M	moderate	male	39	163	46	6.6	20.9	1.12	17.00	0.23	TC	GG
166	057M	moderate	male	38	162	45	7.1	25.1	1.22	17.24	0.21	CC	AG
167	058M	moderate	female	29	165	52	5.4	18.8	2.10	38.90	0.69	TC	AG
168	060M	moderate	female	54	153	41	6.0	20.4	1.97	32.80	0.50	TC	AG
169	061M	moderate	male	19	149	35	5.8	21.0	1.83	31.50	0.47	TC	AG
170	062M	moderate	male	28	166	45	7.0	23.2	1.99	28.40	0.45	CC	AA
171	063M	moderate	male	32	167	52	6.2	22.5	2.06	33.30	0.58	TC	AA
172	064M	moderate	male	38	161	43	6.8	23.6	1.01	14.90	0.18	TC	AG
173	065S	moderate	female	30	150	43	4.9	17.7	0.54	11.00	0.13	CC	GG
174	093R	moderate	female	34	160	45	6.3	21.2	2.48	39.34	0.67	TC	AG
175	095M	moderate	female	38	156	43	6.6	22.0	1.85	28.04	0.41	TC	GG
176	113M	moderate	female	14	149	37	5.9	20.0	2.27	38.41	0.65	CC	AG
177	117M	moderate	female	53	160	56	6.9	22.6	2.06	29.91	0.46	TC	AG
178	130N	moderate	male	25	168	53	6.4	22.0	1.71	26.79	0.37	TC	GG
179	133M	moderate	male	35	157	49	7.1	23.5	1.79	25.23	0.35	TC	AG
180	134M	moderate	female	23	147	45	5.3	18.2	1.43	27.01	0.38	TC	AG
181	141M	moderate	female	47	165	58	8.7	29.1	3.93	45.16	0.86	CC	GG
182	142N	moderate	male	36	174	54	5.0	18.2	0.79	15.79	0.20	TC	AA
183	143M	moderate	female	21	158	50	5.7	21.0	2.21	38.80	0.67	TC	AG
184	144M	moderate	male	32	160	50	6.3	41.0	0.83	13.23	0.16	TC	GG
185	147R	moderate	female	47	160	55	6.1	21.3	1.53	25.15	0.37	TC	GG
186	150N	moderate	female	32	153	47	7.0	23.9	2.23	31.87	0.50	TC	GG
187	156C	moderate	male	5	112	21	6.2	20.0	0.61	9.89	0.39	TC	AG
188	157C	moderate	male	4	103	19	9.6	31.1	0.59	6.17	0.18	TC	AG
189	159C	moderate	male	6	112	18	8.0	24.5	0.20	2.46	0.40	CC	GG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	HbF (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
190	169C	moderate	female	3	91	13	8.0	22.9	1.51	18.83	0.86	CC	AG
191	170C	moderate	female	6	122	22	7.5	22.9	1.10	14.70	0.56	TC	AA
192	172C	moderate	female	6	110	17	7.7	24.9	1.14	14.76	0.81	TC	GG
193	175C	moderate	female	11	143	36	7.2	24.0	2.02	28.06	0.70	TC	GG
194	177C	moderate	female	9	121	26	7.0	23.6	2.03	28.94	0.42	.	GG
195	187CR	moderate	male	11	136	28	7.0	23.4	2.79	39.88	0.72	TC	AA
196	193CR	moderate	female	9	122	21	6.9	24.9	1.92	27.78	0.39	TC	GG
197	205C	moderate	female	6	113	20	6.6	30.1	2.38	36.00	0.59	TT	GG
198	216CR	moderate	male	14	146	33	7.2	25.4	3.18	44.22	0.82	TT	AG
199	218CR	moderate	male	9	128	27	6.1	20.7	1.27	20.88	0.53	TC	AG
200	227CR	moderate	female	5	100	14	5.3	18.3	2.48	46.88	0.93	TC	GG
201	228CR	moderate	female	13	139	31	8.7	28.0	4.38	50.35	1.05	TC	AG
202	234CR	moderate	male	16	155	47	8.0	29.0	2.17	27.17	0.39	CC	AG
203	245CR	moderate	male	4	103	18	7.2	23.5	2.21	30.64	0.85	TC	AG
204	246CR	moderate	male	7	118	21	7.8	25.2	1.68	21.54	0.73	TC	AG
205	247CR	moderate	male	11	132	27	6.9	24.2	2.29	33.20	0.51	TC	AG
206	250CR	moderate	male	8	119	22	6.4	21.3	1.94	30.27	0.78	TC	AG
207	252CR	moderate	male	13	135	24	7.1	25.4	2.17	30.55	0.45	TC	AG
208	253CR	moderate	female	12	132	24	6.5	21.3	3.00	46.21	0.89	CC	GG
209	255CR	moderate	female	15	140	29	7.3	25.7	2.77	37.99	0.63	TC	GG
210	256CR	moderate	female	15	136	27	7.6	26.6	2.88	37.90	0.63	TC	GG
211	263CR	moderate	female	10	127	26	6.5	30.9	2.62	40.36	0.70	TC	AA
212	268CR	moderate	female	10	132	28	6.2	23.8	1.98	31.91	0.48	TC	AA
213	271C	moderate	male	12	150	48	9.2	29.9	1.24	13.44	0.47	TC	GG
214	274C	moderate	male	14	136	26	7.4	24.5	2.05	27.75	0.55	CC	GG
215	275C	moderate	male	13	154	36	6.9	25.1	1.62	23.43	0.33	TC	AG
216	277C	moderate	female	3	100	16	9.7	30.4	1.28	13.17	0.99	TC	AA

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
217	279M	moderate	female	19	151	45	7.4	24.6	1.55	20.91	0.27	CC	AA
218	281C	moderate	female	10	128	23	8.4	27.5	2.68	31.90	0.77	TC	AG
219	283C	moderate	male	6	106	18	8.2	26.4	1.00	12.20	0.81	TC	AG
220	284C	moderate	male	5	106	17	7.0	23.3	1.02	14.56	0.50	TC	GG
221	288C	moderate	female	9	123	23	9.4	29.7	0.66	6.98	0.31	TC	GG
222	289C	moderate	female	7	118	22	7.8	25.9	1.86	23.82	0.66	TC	GG
223	290C	moderate	female	7	115	20	6.7	22.2	1.50	22.36	0.67	TC	AG
224	294C	moderate	male	6	109	17	5.9	19.2	0.59	10.08	0.33	CC	AG
225	295C	moderate	female	5	114	14	8.3	27.5	0.99	11.87	0.26	TC	AG
226	296C	moderate	male	6	119	21	8.9	28.0	1.90	21.29	0.80	TC	AA
227	300C	moderate	female	6	118	22	6.8	22.0	0.97	14.27	0.53	CC	GG
228	302C	moderate	male	13	133	26	6.3	20.1	1.89	29.93	0.76	TC	AA
229	316R	moderate	female	7	104	14	6.7	24.0	2.60	38.73	0.69	TC	AG
230	317C	moderate	male	10	126	22	8.0	27.8	2.13	26.62	0.87	TC	AG
231	322C	moderate	female	9	132	28	7.8	25.0	0.54	6.91	0.29	CC	GG
232	323C	moderate	male	11	127	23	12.1	41.1	4.54	37.53	0.78	TC	GG
233	326C	moderate	female	14	162	46	6.7	24.7	2.64	39.34	0.67	TC	AA
234	330R	moderate	female	5	103	16	5.9	19.7	2.18	36.93	0.82	TC	AG
235	332R	moderate	female	14	155	40	4.8	15.6	2.10	43.77	0.84	TC	AG
236	340C	moderate	female	9	119	19	6.2	22.2	2.63	42.39	0.76	TC	AG
237	341C	moderate	female	9	117	18	5.9	21.0	2.51	42.49	0.76	TC	AG
238	344R	moderate	female	10	133	30	6.3	18.8	0.55	8.67	0.66	TC	AG
239	347R	moderate	male	12	136	28	8.1	25.1	0.82	10.18	0.37	CC	GG
240	350R	moderate	female	21	178	54	6.7	23.5	2.10	31.33	0.48	TC	AG
241	366NR	moderate	female	11	132	28	7.3	25.2	3.21	44.04	0.81	TC	GG
242	368NR	moderate	female	9	113	15	5.9	20.3	2.19	37.13	0.62	TC	AG
243	383NR	moderate	male	5	110	19	9.2	28.4	1.03	11.15	0.51	CC	GG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
244	386NR	moderate	female	11	132	25	5.6	18.8	2.11	37.76	0.64	TC	GG
245	387NR	moderate	male	2	81	10	5.9	21.0	3.13	52.99	1.18	TC	AA
246	392NR	moderate	female	11	139	33	6.8	23.3	2.69	39.57	0.68	TC	GG
247	395NR	moderate	male	6	104	18	6.9	25.7	1.93	27.95	0.40	TC	AG
248	396NR	moderate	male	6	111	18	8.1	26.0	1.87	23.08	0.62	TC	AA
249	398NR	moderate	female	15	143	37	8.1	28.6	2.97	36.61	0.60	TC	GG
250	403NR	moderate	female	10	117	20	6.1	20.8	2.29	37.49	0.62	TC	GG
251	409NR	moderate	male	59	162	43	6.2	21.8	1.36	21.87	0.49	TC	AG
252	416NR	moderate	female	31	150	39	6.7	22.7	0.98	14.67	0.21	CC	GG
253	419NR	moderate	female	64	155	45	6.3	20.6	1.82	28.83	0.63	CC	GG
254	421NR	moderate	female	32	150	40	7.4	27.4	2.04	27.57	0.39	TC	AA
255	423NR	moderate	male	33	155	42	6.5	23.4	1.62	24.93	0.37	TC	AG
256	431NR	moderate	female	17	132	30	6.1	20.5	0.88	14.47	0.21	CC	AG
257	438NR	moderate	male	44	173	52	6.5	22.1	1.36	20.98	0.32	TC	AG
258	439R	moderate	female	10	125	27	7.0	23.6	2.84	40.53	0.75	TC	AG
259	441R	moderate	female	26	158	50	6.9	24.6	2.62	37.97	0.63	TC	AG
260	452R	moderate	female	9	122	23	6.2	20.2	2.23	35.95	1.31	CC	GG
261	455N	moderate	male	19	164	45	6.4	20.7	2.54	39.69	0.69	CC	AG
262	459R	moderate	male	15	145	32	7.0	24.8	2.00	28.54	0.44	TC	AG
263	463R	moderate	male	17	173	63	6.0	21.0	2.09	34.84	0.64	TC	GG
264	466R	moderate	male	45	154	43	6.6	21.7	1.42	21.52	0.31	TC	AG
265	468C	moderate	male	11	123	25	6.9	23.5	1.89	27.33	0.55	CC	AG
266	469R	moderate	female	46	165	51	5.1	16.2	1.54	30.27	0.46	TC	AA
267	470R	moderate	female	49	162	49	6.9	22.2	2.84	41.15	0.74	CC	AG
268	485C	moderate	female	3	126	22	8.9	28.4	4.01	45.03	1.35	TC	AG
269	487C	moderate	female	42	160	52	8.2	28.3	1.46	17.80	0.34	TC	GG
270	490R	moderate	female	11	121	21	7.2	22.9	2.73	37.95	0.63	TC	GG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
271	494R	moderate	male	13	167	40	5.9	19.1	2.86	48.41	0.97	TC	AG
272	495R	moderate	male	5	107	15	7.7	24.2	2.68	34.76	1.15	TC	AG
273	497R	moderate	female	6	108	18	6.4	18.7	0.99	15.41	0.91	TC	GG
274	510CB	moderate	male	5	115	18	5.9	19.3	2.22	37.57	0.61	TC	AG
275	511CB	moderate	male	4	110	16	6.6	22.8	2.24	33.92	0.52	TC	GG
276	516CB	moderate	male	9	121	16	6.3	19.6	2.96	46.98	0.91	CC	AG
277	518CB	moderate	male	22	171	43	6.7	21.2	3.56	53.14	1.17	TC	GG
278	519CB	moderate	female	45	160	50	5.4	17.5	1.01	18.73	0.35	TC	AG
279	520CB	moderate	male	40	155	47	6.8	22.7	2.96	43.58	0.79	TC	AG
280	521CB	moderate	female	24	153	41	5.0	17.0	1.50	30.03	0.44	TC	AG
281	523CB	moderate	female	20	155	43	7.7	25.5	0.64	8.34	0.11	TC	AG
282	524CB	moderate	male	27	163	43	7.3	25.4	2.24	30.72	0.46	CC	GG
283	526R	moderate	male	14	157	43	6.8	22.3	1.38	20.27	0.27	TC	AG
284	529R	moderate	male	5	109	18	7.4	.	2.31	31.25	1.06	TC	AG
285	532N	moderate	female	27	151	39	5.0	18.2	2.07	41.47	0.73	TC	AG
286	542SB	moderate	male	36	160	47	6.4	21	2.04	31.91	0.49	TC	AG
287	561SB	moderate	male	9	115	18	8.1	25.8	4.11	50.72	1.06	TC	AG
288	564SB	moderate	male	15	159	44	7.9	26.3	3.64	46.12	0.89	TC	AA
289	569SB	moderate	male	9	125	26	9.2	29.1	1.29	14.06	0.30	CC	GG
290	574SB	moderate	male	37	170	62	5.6	18.5	1.43	25.61	0.36	CC	GG
291	584SB	moderate	male	12	138	27	7.9	25.2	2.06	26.08	0.44	TT	AG
292	586SB	moderate	male	17	156	45	6.9	21.7	3.12	45.23	0.87	TC	AG
293	597SB	moderate	male	7	116	22	7.7	24.5	1.79	23.25	0.52	CC	AG
294	600SB	moderate	male	10	126	26	8.3	25.6	1.87	22.54	0.57	TC	AG
295	617M	moderate	male	34	157	41	7.8	25.0	3.13	40.10	0.69	TC	GG
296	620C	moderate	male	10	139	28	6.7	17.4	2.80	41.72	0.74	TC	AG
297	641CB	moderate	male	22	166	58	5.8	21.4	0.93	16.09	0.20	CC	GG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
298	654C	moderate	male	17	177	50	7.2	24.4	1.10	15.28	0.21	TC	GG
299	673LP	moderate	male	11	131	24	6.8	22.3	2.80	41.13	0.72	TC	GG
300	676LP	moderate	male	3	89	13	7.1	22.4	2.82	39.78	0.68	TC	GG
301	001N	severe	male	6	112	14	7.5	24.4	3.10	41.39	1.04	TC	AA
302	003N	severe	male	15	142	32	5.0	16.1	0.95	19.06	0.49	CC	AG
303	004N	severe	female	17	166	43	4.7	15.4	1.09	23.10	0.35	CC	GG
304	005N	severe	female	8	122	24	8.9	26.3	0.32	3.60	0.22	TC	GG
305	008N	severe	male	17	134	26	6.3	22.2	1.61	25.60	0.39	TC	GG
306	015M	severe	male	22	152	45	8.1	26.1	3.97	49.00	1.00	TC	AA
307	016M	severe	female	29	150	39	6.0	20.0	2.83	47.19	0.91	TC	AG
308	017M	severe	female	41	155	55	5.8	19.4	0.45	7.68	0.14	CC	GG
309	025N	severe	female	12	126	24	3.5	12.9	0.67	19.02	0.31	CC	AG
310	028N	severe	male	12	134	30	6.4	21.2	2.14	33.47	0.69	TC	GG
311	033M	severe	female	34	140	42	6.7	22.1	1.29	19.24	0.67	TC	AG
312	035M	severe	female	41	154	39	6.0	19.9	2.58	43.00	0.84	TC	AG
313	037M	severe	female	38	153	37	6.2	22.0	1.37	22.10	0.30	TC	GG
314	040M	severe	female	33	156	53	6.2	22.8	1.63	26.30	0.39	CC	AG
315	049M	severe	male	30	152	48	6.5	22.7	2.46	37.92	0.63	TC	GG
316	069N	severe	female	8	121	22	5.3	25.4	1.99	37.60	0.64	TC	AA
317	072R	severe	male	25	150	40	4.2	14.3	1.40	33.33	0.53	CC	GG
318	074R	severe	male	29	162	45	7.5	25.0	2.61	34.81	0.66	TC	AG
319	075R	severe	female	24	160	52	7.1	23.8	2.05	28.85	0.42	TC	AG
320	076N	severe	female	6	104	17	5.7	19.3	2.54	44.51	0.82	CC	AG
321	077N	severe	male	5	111	16	5.9	16.2	1.40	23.80	0.40	TC	AG
322	082M	severe	male	44	136	34	4.9	16.8	1.99	40.64	0.71	TC	GG
323	084N	severe	female	10	120	22	4.8	16.9	0.37	7.64	0.17	CC	AA
324	085N	severe	female	11	120	26	4.2	24.1	2.36	56.12	1.37	CC	AG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
325	089R	severe	male	30	160	43	5.7	20.3	1.35	23.76	0.32	CC	AG
326	090R	severe	male	24	170	54	7.0	23.1	2.33	33.26	0.52	TC	AG
327	091R	severe	male	27	164	48	4.9	17.2	0.84	17.13	0.21	TC	GG
328	092R	severe	male	25	165	46	6.6	23.6	0.99	14.95	0.19	TC	GG
329	096M	severe	male	23	170	48	5.0	17.5	1.31	26.29	0.37	TC	AG
330	097N	severe	female	18	134	26	3.4	12.6	0.70	20.68	0.28	CC	GG
331	098T	severe	female	31	155	37	5.4	18.6	1.98	36.64	0.64	TC	AG
332	1006S	severe	male	17	157	37	7.9	23.0	1.90	24.10	0.56	TC	GG
333	100N	severe	female	13	136	26	5.5	24.0	1.66	30.12	0.82	TC	GG
334	1012S	severe	female	15	132	28	7.5	25.2	2.42	32.25	0.49	TC	AG
335	1029S	severe	male	14	162	40	11.8	35.2	0.94	7.97	0.31	TC	GG
336	1034S	severe	female	30	157	48	5.7	18.3	1.51	26.48	0.55	TC	AG
337	119M	severe	male	43	159	43	5.4	17.4	1.65	30.60	0.46	TC	AG
338	121N	severe	female	31	156	42	6.9	23.9	1.79	25.93	0.75	TC	GG
339	126R	severe	female	30	167	49	6.7	22.0	0.71	10.59	0.39	TC	AG
340	127R	severe	female	18	160	34	6.8	21.5	0.31	4.60	0.21	TC	AG
341	128N	severe	male	31	173	48	5.3	19.2	1.39	26.27	0.37	TC	GG
342	131M	severe	female	41	137	36	4.8	16.7	1.34	27.95	0.40	TC	GG
343	136S	severe	female	38	150	43	7.0	25.1	1.93	27.63	0.40	TC	AA
344	152M	severe	female	30	151	38	5.4	20.5	0.80	14.79	0.18	TC	AA
345	155C	severe	male	10	124	24	5.8	21.0	0.59	10.24	0.24	CC	GG
346	158C	severe	male	12	125	22	7.1	24.0	1.59	22.37	0.55	TC	AA
347	160C	severe	male	6	113	18	8.1	26.0	0.58	7.12	0.26	TC	GG
348	165C	severe	male	6	107	18	5.9	20.0	0.39	6.64	0.16	CC	AG
349	167M	severe	female	27	160	42	4.4	16.0	0.59	13.40	0.21	TC	AG
350	171C	severe	female	13	150	38	8.0	25.0	0.38	4.73	0.24	CC	AG
351	176C	severe	female	10	128	25	7.8	25.0	0.55	7.06	0.20	CC	GG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
352	178C	severe	male	14	136	27	8.6	28	0.40	4.64	0.18	CC	AA
353	182CR	severe	male	7	112	18	5.8	19.2	0.48	8.25	0.25	TC	GG
354	184CR	severe	male	11	134	33	7.3	24.5	0.86	11.85	0.24	TC	GG
355	185CR	severe	male	11	126	23	6.9	23.0	1.04	15.12	0.19	TC	AA
356	189CR	severe	male	11	135	27	6.9	23.8	1.57	22.75	0.37	TC	AA
357	196CR	severe	male	7	115	19	6.6	24.0	2.47	37.43	0.61	TC	AA
358	198CR	severe	male	2	75	9	9.1	28.5	1.34	14.74	0.97	TC	AG
359	201CR	severe	male	11	133	24	5.9	22.1	1.33	22.51	0.30	TC	AG
360	202CR	severe	male	8	120	22	6.0	21.8	1.40	23.38	0.42	TC	AG
361	207CR	severe	male	9	127	25	6.1	21.9	1.08	17.72	0.44	TC	AG
362	210CR	severe	male	10	128	24	5.8	30.6	2.00	34.40	0.62	CC	AA
363	211CR	severe	male	11	131	27	5.2	19.8	0.88	16.91	0.21	CC	GG
364	212CR	severe	male	6	104	18	5.2	17.7	1.89	36.37	0.80	TC	AG
365	219CR	severe	male	6	101	15	5.4	18.9	1.57	29.11	0.46	TC	AG
366	221CR	severe	female	10	114	20	5.8	20.7	0.71	12.32	0.24	TC	AG
367	223CR	severe	male	6	112	17	6.7	24.4	1.07	16.03	0.66	TC	AG
368	229CR	severe	female	12	129	24	7.7	26.0	0.91	11.83	0.21	CC	GG
369	230CR	severe	female	13	136	27	6.5	23.3	0.85	13.07	0.31	TC	GG
370	232CR	severe	female	12	128	24	4.9	18.9	1.15	23.52	0.32	TC	GG
371	237CR	severe	female	6	111	17	5.5	21.1	2.07	37.69	0.66	TC	GG
372	241CR	severe	female	5	98	15	4.2	15.7	0.32	7.69	0.20	CC	AG
373	248CR	severe	female	13	139	29	6.0	22.0	2.34	39.08	0.70	TC	GG
374	266CR	severe	male	11	128	24	5.4	20.4	0.86	15.90	0.32	CC	AG
375	272C	severe	female	11	127	27	7.6	85.0	0.80	10.59	0.47	CC	AG
376	273C	severe	female	7	119	23	5.9	21.0	2.79	47.29	0.93	TC	GG
377	276C	severe	male	14	145	31	5.9	22.0	1.78	30.09	0.45	TC	GG
378	280M	severe	female	46	156	40	5.9	21.0	0.77	13.11	0.25	TC	AG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
379	282C	severe	female	8	118	21	6.9	24.0	0.25	3.56	0.14	TC	AG
380	291C	severe	male	12	128	24	8.0	26.0	1.49	18.66	0.77	CC	GG
381	292C	severe	female	17	158	45	5.2	18.0	1.38	26.61	0.38	CC	AG
382	293C	severe	female	9	123	20	7.1	23.4	0.71	9.99	0.39	CC	AG
383	297C	severe	male	8	114	19	6.1	21.0	0.94	15.44	0.44	TC	GG
384	298C	severe	female	8	115	20	5.8	19.0	1.72	29.72	0.51	TC	GG
385	299C	severe	female	15	145	36	7.3	25.0	0.88	12.00	0.26	CC	GG
386	301C	severe	female	10	137	25	6.5	23.0	1.57	24.22	0.59	TC	GG
387	304C	severe	female	14	133	25	6.6	26.0	0.40	6.04	0.08	CC	AG
388	307C	severe	male	7	109	18	7.9	27.0	0.90	11.38	0.26	CC	GG
389	308C	severe	male	12	136	26	4.9	18.0	0.54	10.97	0.23	CC	AG
390	309C	severe	male	5	99	16	7.2	25.0	1.01	13.98	0.62	TC	AG
391	310C	severe	male	13	149	30	7.4	26.0	2.10	28.35	0.72	TC	AG
392	311C	severe	male	10	139	29	5.5	22.0	1.79	32.62	0.59	TT	AG
393	314R	severe	male	6	108	18	6.4	24.3	1.89	29.53	0.46	TC	AG
394	318C	severe	male	7	111	18	6.1	23.0	1.41	23.11	0.42	TC	AG
395	319C	severe	male	15	126	22	7.1	26.0	0.62	8.73	0.33	TC	GG
396	320C	severe	male	7	111	17	6.8	24.0	1.58	23.18	0.77	TC	AG
397	324C	severe	female	15	142	31	5.5	24.0	0.50	9.08	0.11	CC	AG
398	325C	severe	female	5	103	14	4.8	19.0	0.94	19.62	0.34	CC	GG
399	328C	severe	male	13	135	32	6.3	22.0	1.83	29.01	0.65	TC	AG
400	331R	severe	male	11	125	23	7.6	25.4	0.51	6.73	0.33	CC	GG
401	333R	severe	male	5	105	16	4.9	16.2	0.83	16.91	0.60	TC	GG
402	334R	severe	male	9	119	25	8.9	29.4	1.28	14.33	0.39	TC	GG
403	335R	severe	male	15	131	33	7.0	23.2	0.62	8.87	0.30	TC	AG
404	336R	severe	female	7	108	17	4.4	14.1	0.35	7.87	0.36	TC	AG
405	338C	severe	male	10	117	20	5.9	20.0	0.77	13.08	0.28	CC	AG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
406	345R	severe	male	9	120	24	4.6	15.8	0.74	16.09	0.39	TC	AG
407	346R	severe	male	12	141	34	5.8	20.7	0.62	10.71	0.19	CC	GG
408	349R	severe	male	16	150	42	5.8	21.2	0.73	12.55	0.18	TC	AG
409	352C	severe	male	12	135	29	11	24.9	2.06	18.73	0.38	TC	AG
410	369NR	severe	male	6	108	17	6.9	24.2	1.60	23.16	0.35	TC	AA
411	373NR	severe	male	6	109	12	5.2	22.1	1.30	25.00	0.35	TC	GG
412	378NR	severe	female	9	121	26	5.0	19.3	0.84	16.79	0.33	CC	AG
413	382NR	severe	male	7	104	16	4.8	20.1	1.57	32.77	0.51	TC	GG
414	384NR	severe	female	15	141	32	7.7	25.9	0.91	11.79	0.25	TC	AG
415	390NR	severe	male	15	145	28	5.4	21.2	1.03	19.10	0.24	TC	AG
416	391NR	severe	female	13	128	27	4.7	17.0	1.47	31.25	0.48	CC	GG
417	394NR	severe	male	7	112	16	5.4	19.1	0.66	12.21	0.19	CC	GG
418	413NR	severe	male	17	140	31	4.2	17.8	0.78	18.53	0.24	CC	AG
419	422NR	severe	male	20	150	40	5.4	20.6	1.05	19.39	0.25	CC	AG
420	434NR	severe	female	24	150	44	6.1	23.5	1.32	21.57	0.28	TC	AG
421	443R	severe	male	7	118	20	3.9	13.7	0.43	10.98	0.17	TC	AG
422	454C	severe	female	14	138	28	6.7	25.0	0.71	10.61	0.18	CC	AG
423	457R	severe	male	16	164	43	8.7	28.1	0.61	7.04	0.23	CC	GG
424	460R	severe	male	19	167	47	7.3	26.0	2.62	35.82	0.57	TC	GG
425	471R	severe	male	9	142	30	6.1	20.8	2.11	34.57	0.57	CC	GG
426	473R	severe	female	14	144	34	6.8	21.6	1.26	18.55	0.98	CC	AG
427	475R	severe	male	11	117	24	8.5	23.4	0.60	7.10	0.25	CC	GG
428	502CB	severe	female	10	127	24	5.2	20.0	1.45	27.81	0.48	TC	GG
429	504CB	severe	female	3	91	13	5.1	15.7	0.66	12.97	0.55	CC	AG
430	505CB	severe	male	13	164	41	8.9	28.2	0.85	9.55	0.37	TC	AG
431	544SB	severe	female	17	132	26	4.8	17.4	1.47	30.65	0.47	TC	AG
432	550SB	severe	female	10	131	26	4.3	19.2	0.48	11.16	0.13	CC	GG

Appendix B: Clinical data and biochemical characteristics of β°-thalassemia/Hb E patients (continued)

No.	ID	Severity	Sex	Age (years)	Height (cm)	Weight (kg)	Hb (g/dL)	HCT (%)	Hb F (g/dL)	Hb F (%)	Proportion of Hb F to Hb E	*Xmn*I	ZHX2
433	553SB	severe	female	10	125	24	5.9	19.8	1.44	24.45	0.82	TC	AG
434	563SB	severe	female	12	133	29	6.8	22.2	0.83	12.24	0.39	CC	AG
435	576SB	severe	female	5	107	16	4.7	16.5	1.41	30.07	0.47	TC	AG
436	603SB	severe	female	33	157	43	5.1	18.6	0.79	15.48	0.22	CC	AG
437	605SB	severe	female	6	106	16	7.2	22.5	1.00	13.83	0.55	TC	GG
438	607SB	severe	female	6	106	18	7.0	22.8	1.16	16.56	0.39	CC	GG
439	608SB	severe	female	3	85	11	7.1	22.8	2.44	34.36	0.68	TC	AG
440	612SB	severe	female	8	116	21	6.9	21.9	0.91	13.24	0.52	CC	GG
441	659M	severe	female	24	150	49	5.6	20.1	0.67	11.94	0.17	TC	AA
442	678LP	severe	female	9	128	23	4.4	16.4	1.27	28.79	0.48	CC	AG
443	679LP	severe	female	13	133	23	5.6	19.9	1.18	21.12	0.28	TC	AG
444	686LP	severe	female	11	125	23	5.3	19.3	1.19	22.44	0.30	CC	GG
445	695LP	severe	female	12	130	22	5.1	17.9	1.72	33.63	0.53	TC	GG
446	702LP	severe	female	13	153	39	5.8	20.2	2.08	35.78	0.62	TC	GG
447	713LP	severe	female	7	112	18	4.0	14.8	0.68	16.93	0.28	CC	AG
448	722R	severe	female	16	155	38	4.7	15.5	1.32	28.07	0.81	CC	AG
449	756UT	severe	female	26	159	41	5.7	19.9	1.32	23.11	0.31	CC	GG
450	850S	severe	female	19	163	43	4.9	19.8	1.27	26.89	0.39	CC	AA

BIOGRAPHY

NAME	Miss Chiranya Sangprasert
DATE OF BIRTH	13 September, 1981
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	Silpakorn University, 2000-2003: Bachelor of Science (Chemistry) Mahidol University, 2005-2008: Master of Science (Biochemistry)
RESEARCH GRANTS	Partial Thesis Writing Grant supported From Faculty of Graduate Studies, Mahidol University
HOME ADDRESS	233 Doembang RD., T. Mahachai, Muag Samuthsakorn, 74000 Thailand E-mail: Chiranya_s@hotmail.com