MOLECULAR STUDY OF TRANSFUSION DEPENDENT THALASSEMIA PATIENTS WHO ATTENDED PEDIATRIC DAY CARE HOSPITAL UNIVERSITI SAINS MALAYSIA, KELANTAN FROM JANUARY 2006 TO APRIL 2006.

by

DR. ROZITAH BINTI RAZMAN

Dissertation Submitted In
Partial Fulfillment Of The
Requirements For The Degree Of
Master Of Medicine

(Paediatrics)

UNIVERSITI SAINS MALAYSIA
2008

ACKNOWLEDGEMENTS

First of all I thank the Almighty god, the most merciful for giving me strength and health to finish this study.

I would like to express my sincere thanks to all supervisors : Assoc. Prof Dr Zilfalil Alwi, (Director of Human Genome Centre, HUSM), Dr Ariffin Nasir and Dr Nursarwany, (Pediatric Lecturer, HUSM) for their help, advice and guidance, correcting this paper and encouraging me to complete this dissertation.

Thank you to all lecturers in Pediatric Department HUSM for their teaching and support throughout my years in HUSM.

I also would like to thank staffs of Human Genome Centre, HUSM especially En. Nizam Zahary, Cik Shafawati and Puan Noratifah who have helped me doing the test and teaching me while in the lab and in many other ways.

My appreciation also goes to my family, all specialists in Pediatric Department, Hospital Raja Perempuan Zainab II, Kota Bharu and all my colleagues who have helped me in various ways.

LIST OF TABLE

LIST OF FIGURES

ABREVIATIONS

β	Beta
DNA	Deoxyribonucleic Acid
PCR	Polymerase Chain Reaction
LCR	Locus Control Region
HS	Hypersensitive Site
HBB gene	Human Hemoglobin Beta gene
dNTP	deoxynucleoside triphosphate
bp	base pair
ARMS	Amplification Refractory Mutation System
RFLP	Restriction Fragment Length Polymorphism.

ABSTRACT

Introduction: Thalassemia has emerged as one of the most common public health problems in Malaysia, particularly among Malaysian Chinese and Malays. This study aims to determine the spectrum of Thalassaemia gene mutations found in transfusion dependent Thalassaemia patients who attended Pediatric Daycare Unit, Hospital Universiti Sains Malaysia, Kelantan, Malaysia. The findings are important for establishing the prenatal diagnosis in our Human Genome Centre.

Methods: This is a cross sectional study in which 38 transfusion dependent Beta Thalassaemia patients were screened for six different mutations previously shown to be prevalent in the Malaysian population. Sampel collection was started in January, 2006 till April, 2006. DNA was extracted from leucocytes collected from the peripheral vein, amplified by PCR and digested by six restriction enzymes for detection of mutations. The mutation were correlated with the clinical severity based on the following clinical parameters : age at presentation, pre-transfusion hemoglobin level, mean volume of blood transfusion per kilo body weight per year, spleen size, spleenectomy and growth failure were recorded in these patients to determine the severity of each group of thalassaemia type depicted by the mutation. For the statistical analysis, Kruskal-Wallis test and univariate analysis were used.

Results: Five of the six known Beta-globin gene defects occurring in the Malaysian population were detected, namely, IVS-1 nt5 (G>C), IVS-1 nt1 (G>T), Codon 26 (G>A), Codon 41-42 (4 bp del) and Codon 19 (A>G). The mutation which was not observed in this study was in Codon 15 (G>A). The two most common mutations observed were Codon 26 (G>A) (54.3%) and IVS-1 nt5 (G>C) (20%). Three patients did not show any

of the six mutations. There were no significance different in age at presentation (p=0.23), pre-transfusion hemoglobin level (p=0.2), volume of blood transfusion given to the patient (p=0.42) and also spleen size (p=0.59) between groups of type of Thalassaemia.

Conclusion: Our results show that the majority of Kelantan Beta Thalassaemia patients have similar beta-globin gene defects as the rest of the Malaysian population. However, mutations in the three patients were not identified. The findings complement the existing data on the Beta Thalassaemia gene mutation in Malaysia.

Keywords: Beta thalassemia, mutation, β globin gene, PCR-RFLP.

ABSTRAK

Pengenalan : Thalassemia kini merupakan salah satu masalah kesihatan yang utama di Malaysia terutama dikalangan masyarakat Melayu dan kaum Cina. Tujuan kajian ini adalah untuk melihat spectrum mutasi gen Thalassemia dikalangan pesakit Thalassemia yang memerlukan transfusi darah yang kerap di Pusat Rawatan Harian Kanak-Kanak, Hospital Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia. Penemuan ini penting untuk membantu dalam membuat diagnosis prenatal di Pusat Genom Manusia.

Metodologi: Ini merupakan kajian hirisan lintang yang mana 38 orang pesakit Thalassemia yang memerlukan transfusi darah yang kerap, telah disaring untuk mengesan enam mutasi berlainan yang telah dikenal pasti sebelum ini amat prevalen di dalam populasi di Malaysia. Pengumpulan sample telah dimulakan pada bulan Januari,2006 sehingga April, 2006. DNA di ekstrak daripada sel darah putih pesakit yang dikumpulkan daripada vena pesakit dan kemudian dilakukan kaedah ujian PCR ke atas darah tersebut. Enam jenis enzim digunakan untuk memproses hasil PCR tadi dan seterusnya analisa untuk mengenal mutasi dilakukan. Korelasi diantara mutasi dengan keterukan penyakit berdasarkan parameter klinikal seperti umur pada kali pertama pesakit datang ke klinik, paras hemoglobin sebelum transfusi darah, saiz limpa, pembuangan limpa dan pertumbuhan yang terencat turut direkodkan. Kruskall-Wallis test dan univariate analisis digunakan untuk analisis statistik

Keputusan : Lima daripada enam gen Beta-globin telah dikesan iaitu, IVS1-nt 5(G>C), IVS1- nt1 (G>T), Codon 26(G>A), Codon 41-42(4 bp del) dan Codon 19(G>A). Mutasi

yang tidak dijumpai dalam kajian ini ialah Codon 15 (G>A). Dua mutasi yang paling banyak dijumpai ialah Codon 26(G>A) (54.3%) dan IVS1-nt5 (G>C) (20%). Tiga pesakit tidak mempunyai semua mutasi yang disebutkan diatas. Dalam kajian parameter klinikal didapati tidak ada perbezaan yang bermakna bagi umur ketika pesakit mula-mula datang ke hospital (p=0.23), paras hemoglobin sebelum transfusi (p=0.2), jumlah darah yang telah diberi kepada pesakit (p=0.42) dan saiz limpa (p=0.59) di dalam setiap kumpulan jenis Thalassemia.

Kesimpulan : Keputusan kajian kami menunjukan majoriti pesakit Beta thalassaemia di Kelantan mempunyai mutasi gen yang sama dengan populasi penduduk Malaysia yang lain. Walau bagaimanapun, mutasi untuk 3 pesakit tidak dapat ditentukan. Penemuan ini telah berjaya menambah data yang telah sedia ada berkenaan mutasi gen Beta thalassemia di Malaysia..

Kata kunci : Thalassemia, mutasi, gen Beta-globin, PCR-RFLP.

1. INTRODUCTION

Thalassaemias are a heterogenous group of genetic disorders in which the production of normal hemoglobin is partly or completely suppressed because of defective synthesis of one or more globin chain.

Beta thalassaemia major is the most severe form of beta-thalassaemia and results from the inheritance of the homozygous state of the phenotype β^{o}/β^{o}.

It has been estimated that there are probably as many as 100,000 living patients with homozygous beta-thalassaemia in the world (Wheatheral, 1995). In addition to Mediterranean countries, β-globin gene disorders are frequently found in Asia and the Far East.

About 4.5% of the people in Malaysia are heterozygous carriers for beta-thalassaemia and carrier couples are at risk of producing a child with beta-thalassaemia major where incidence annually is 2.1/1,000 live births (George E, 1998).

This disease is a public health problem in Malaysia and the clinical picture varies from one patient to another. The management of patients with transfusion dependent thalassaemia constitutes a heavy burden for health authorities. Less than 20% of patients receive adequate iron chelation therapy and the majority are destined to die in the second or third decade of life from complications of multiple organ failure secondary to iron

overload. Only 50% of thalassaemia patients in United Kingdom able to adhere fully to iron chelation therapy. Survival to age 30 years old only between 10-40% (Bernadette M, 2000).

It is likely, that appropriate and extensive screening, accurate detection and counseling of at risk couples, in connection with antenatal diagnosis, is a promising method for the reduction of thalassaemia related mortality and morbidity.

Worldwide, in the last ten years, the spectrum of molecular defects and the clinical severity of the mutations that cause thalassaemia have been identified. The implementation of successful and effective prenatal diagnosis programs for the hemoglobin disorders in any country depends on the diversity of the ethnic groups as over 250 molecular defects has been confirmed in the β–globin gene complex (Weatherall DJ, 2001). The large heterogeneity of β–mutation in a population often lead to β–thalassaemia major patients being compound heterozygous for the disorder; for example, in a study of 50 β–thalassaemia carriers in Singapore, 64% of affected couples carried a different β–mutation each (Tan JA, 1994).

It is fortunate that despite great allelic heterogeneity of the β–globin locus, β–thalassaemia in any one ethnic group or population is caused by few common mutations together with a variable number of rare mutation (Angastiniotis M, 1995). This has allowed effective and rapid prenatal diagnosis programs for β–thalassaemia to be

established in each ethnic group or population. Each ethnic group has 4-5 common mutation that form more than 95% of the mutations seen (Clark BE, 2004). A number of DNA technique incorporating polymerase chain reaction (PCR) are currently available for DNA characterization.

Molecular characterization of the β–globin mutations has been previously studied in Malay and Chinese in Malaysia. The studies confirmed the presence of nine β–globin defects : IVS 1- nt1 (G>T), IVS nt5 (G>C),Codon 26(GAG→AAG), Codon 28(A>G), Codon 15(G>A), Codon 17 (A>T), Codon 19(A>G), Codon 41/42 (-TCTT) and IVS 2-654(C>T). The most common beta-thalassaemia mutations are IVS1 nt5 (G>C) and Codon 41-42 (-TCTT) in the Malays and Chinese respectively (George E, 1998).

Malaysia has yet to establish a nationally endorsed program for population screening for thalassaemia. Currently, carriers could be screened from the red cell indices generated by automated blood counter (a mean cell volume <80 fl, and mean cell hemoglobin, Mean Corpuscular Hemoglobin less than 27 pg may be indicative of thalassaemia). Accurate quantitation of Hb A2 can be done by automated chromatography (Hb A2 levels by HPLC more than 4.0% indicative of beta-thalassaemia) and automated hemoglobin electrophoresis for abnormal hemoglobins. However, the β–mutations could not be characterized using the above methods.

1.1 GENETIC BASIS

1.1.1 Hemoglobin Type

Transport of oxygen from the lungs to the tissue is mediated by a highly specialized molecule, hemoglobin, which is contained within the circulating red cells. Each red cell contains approximately 300 million molecules of hemoglobins, weighing about 30 picograms per cell.

Each molecule of hemoglobin is formed by 2 pairs of identical sub-units, the globin chains, named with Greek alphabet

i) Alpha globin cluster (ζ and α globin chain)

ii) Beta globin cluster (ε, γ, β and δ globin chain).

The globin chains pairing form the following 4 major types of hemoglobin:

i) The Embryonic hemoglobin

This is detected from 3rd to 10th week of gestation and represent: ζ2ε2, α2ε2, ζ2γ2.

ii) The Fetal hemoglobin (HbF)

This is the predominant oxygen carrier during pregnancy and is a α2γ2 molecule.

iii) Adult hemoglobin (HbA)

This replaces HbF shortly after birth. It comprise of $\alpha 2\beta 2$ globin chains.

iv) Minor Adult component hemoglobin (HbA2) – $\alpha 2\delta 2$ chains.

Under normal conditions, the red cells of the adult human contain approximately 98% HbA, traces (<1%) of HbF and <3.8% of HbA2.

1.1.2 β Globin Gene – Structure, Function And Expression

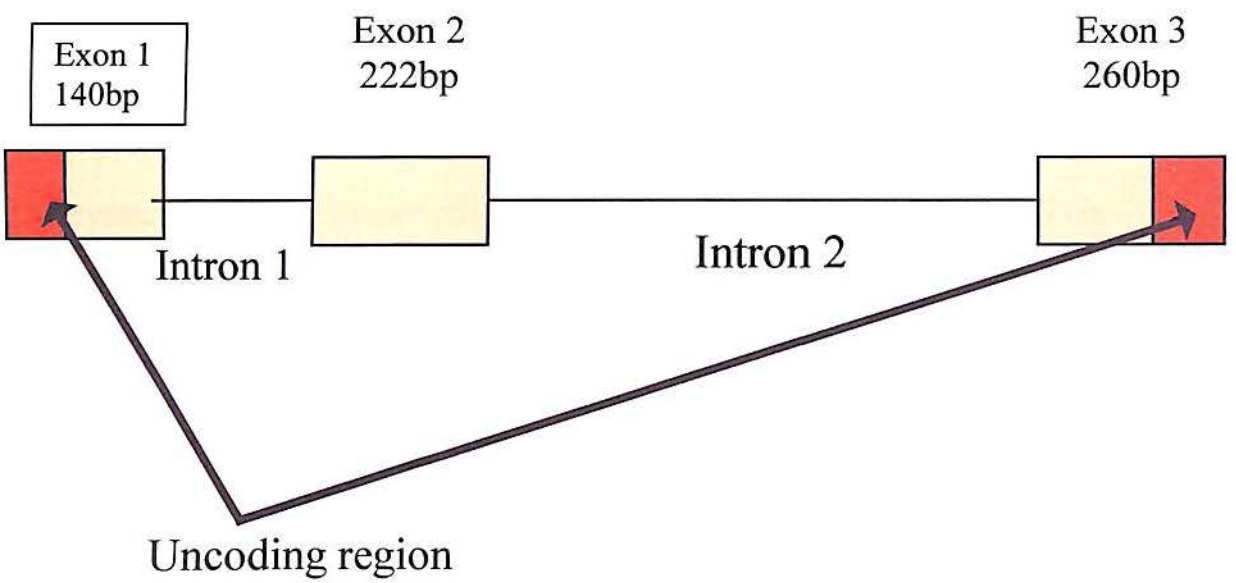

Figure 1.1 : β globin gene (HBB gene) structure (Arthur B, 2005)

β globin is encoded by a structural gene found in a cluster with the other β like genes on the short arm of chromosome 11. The cluster contains five functional genes, 5' - ε - Gγ - Aγ - ψβ - δ - β - 3' which are arrange in the order of their developmental expression. Upstream of the entire β globin complex is the locus control region (LCR), which consisted of five DNase 1 hypersensitive (HS) sites (designated HS 1-5) distributed between 6 and 20 kb 5' of the ε gene (Stamatoyannopoulus G, 2001).

The LCR plays a critical role in β globin gene expression by maintaining an open chromatin state and acting as a powerful enhancer of globin gene transcription; in its absence, the level of β globin gene expression is low (Li Q, 1999).

Four of the LCR sites (HS 1- 4) are erythroid-specific, encompassing binding sequences for erythroid-restricted transcription factors (GATA-1 and NF-E2), while HS5 is ubiquitous (Lodish HF, 1972).There is one other hypersensitive site approximately 20 kb 3' to the β gene. The two extreme HS sites flanking the β complex have been suggested to mark the boundaries of the β globin gene domain.

The general structure of the β globin gene is typical of the other globin loci. The genomic sequence, which codes for 146 amino acids, spans 1600 bp; the transcribed region is contain in three exons separated by two introns or intervening sequences (IVSs).

Exon 2 encodes the residues involved in haem binding and αβ dimer formation, while exons 1 and 3 encode for the non-haem-binding regions of the β globin chain (Anderson KP, 1993).

Many of the amino acids involved in globin subunit interactions required for the Bohr effect, and 2,3-diphosphoglycerate binding, are found in exon 3. Conserved sequences important for β globin gene expression are found in the 5' promoter region, at the exon-intron junctions, and in the 3' untranslated region (3' –UTR) at the end of the mRNA sequences (Grosveld F, 1998)

The β globin gene promoter includes three positive *cis*-acting elements : a TATA box (position -28 to -31) , a CCAAT box (position -2 to -76) and duplicated CACCC motifs (proximal at positions -86 to -90, and distal at position -101 to -105). While the CCAAT and TATA elements are found in many eukaryotic promoters, the CACCC sequence is found predominantly in erythroid cell-specific promoters. Binding of the erythroid Kruppel like factor (EKLF) to the CACCC motif appears to be crucial for normal adult β globin expression (Jane SM, 1992)

In addition to these motifs, the region upstream of the β globin promoter contains two binding motifs for the erythroid transcription factor GATA-1. The importance of these various 5'-flanking sequences for normal gene expression is underscored by β thalassaemia arising from point mutations in these sequences specifically in and around

the TATA box and the CACCC motifs in the -80 to -100 region. An enhancer is also found in intron 2 and 3' of the globin gene, 600-900 bp downstream of the poly (A) site (Li Q, 1999).

The 5'-UTR occupies a region of 50 nucleotides between the CAP site, the start of transcription, and the initiation (ATG) codon. There are two prominently conserved sequences in the 5'-UTR of the various globin genes (both α and β). One is the CTTCTG hexanucleotide found 8-13 nucleotides downstream from the CAP site, i.e. at positions +8 to +13 (Hanscombe, 1991).

The second conserved sequence is CACCATG, in which the last three nucleotides form the initiation codon (ATG). Again, the importance of these sequence in the regulation of the β gene expression is exemplified by several mutations in the 5'-UTR causing β thalassaemia. The 3'-UTR constitutes the region of 132 nucleotides between the termination codon (TAA) and the poly (A) tail with one conserved sequence, AATAAA, located 20 nucleotides upstream of the poly (A) tail (Myers RM, 1986). Several mutations affecting the AATAAA sequence and other sequences in the 3'-UTR causing β thalassaemia (Myers RM, 1986)

The developmental regulation of the globin gene reflect their sequential activation in a 5' to 3' direction. Transcription of the ε gene in the embryonic yolk gene switches after

the sixth week of gestation to the transcription of the two γ genes in the fetal liver, and then around the prenatal period, to that of the δ (minor adult) and β (major adult) genes.

At six month of life, Hb F comprises <5% of the total hemoglobin and continues to fall reaching the adult level of <1% at 2 years of age. It is at this stage that mutations affecting the β gene become clinically apparent. The 'switch' from fetal (γ) to adult (β) haemoglobin production is not complete, and small amounts of γ expression persist in adult life. All adults have residual amounts of fetal haemoglobin (Hb F, $\alpha 2\gamma 2$), present in a subset of erytrocytes called F cells which also contain adult ($\alpha 2\beta 2$) haemoglobin. These levels of Hb F and F cells in adults vary considerably, and are largely genetically controlled (Thien SL, 2005).

The developmental expression of the individual globin genes relies on two mechanisms, gene silencing and gene competition, governed by direct physical interactions between the globin promoters and the β-LCR , which are dependent on the transcription environment in embryonic, fetal and adult cells. While the ε and γ globin genes are autonomously silenced at the appropriate developmental stage, expression of the adult β globin gene depends on the lack of competition from the γ gene for the LCR sequences (Grosveld F, 1999).

This is supported by the concomitant down-regulation of the *cis* β gene when the γ gene is persistently up-regulated by mutations in the promoter region in hereditary

persistence of fetal haemoglobin (Hanscombe O, 1991). In addition, mutations that affect the β promoter, which remove competition for the β–LCR, tend to be associated with variable increase in the γ and δ gene expression (Jane SM, 1995).

Beta-thalassaemia occurs when there is a quantitative reduction of beta globin chains, caused by mutations of human hemoglobin beta gene (HBB gene). The gene is located on the short (p) arm of chromosome 11 at position 15.5. The size of HBB gene is 1605 base pair (Thien SL, 2000)

Most of the mutations involved a change in a single DNA building block (nucleotide) within or near the HBB gene. Other mutations insert or delete one to several nucleotides in the HBB gene. These genetic defects lead to a variable reduction in beta globin output ranging from minimal deficit (β^+ thalassaemia allele) or complete absence (β^o thalassaemia).

Typically, β thalassaemia inheritance pattern is autosomal recessive. The most severe, β^o thalasaemia is characterized by the complete absence of Hb A and results from the inheritance of two β^o thalassaemia alleles (homozygous). This normally results in the patient being transfusion dependent. The patient typically presents within 6 month of life with profound anaemia.

Techniques for mutational analysis of the beta globin gene (HBB) has been established (Mario Pirastru, 1983; Tan JA, 1993; Ng IS, 1994; Cao A,1997; Wheatherall DJ, 2001; Old JM, 2003). The identification of the underlying mutations maybe useful for prediction of the clinical phenotype in some cases as well as presymptomatic diagnosis of at risk family.

Molecular studies may reveal a large array of abnormalities which underlies the various phenotypes and may help in their identification.The strategy for the community control of Beta-thalassaemia major requires prior characterization of the spectrum of β-globin gene mutation in the population.

Mainly, there are 2 forms of genetic defects which produced beta thalassaemia :

a) Non-deletion form:

   This defect generally involved a single base substitution or small deletion or insertion near the beta globin gene. Most commonly, these mutation occur in promoter regions preceding the beta globin gene. Mutation affecting transcription can either involve the conserved DNA sequences that form the β globin promoter or the stretch of 50 nucleotides in the 5'-UTR (Little JA, 1995). Generally, they result in mild to moderate deficit of β globin output that reflects the relatively mild phenotype of these β+ thalassaemia.

   The C→ T mutation at position -101 to the β globin gene appears to cause an

extremely mild deficit of β globin, such that it is 'silence' in heterozygotes who have normal Hb A2 levels and normal red cells indices (Huisman T, 1997). Several mutations in the 5'-UTR e.g CAP + 1A-C, also have 'silent' phenotype.

Mutation that affect RNA processing can involve either of the invariant dinucleotides (GT at 5' and AG at 3') in the splice junction in which case normal splicing is completely abolished with the resulting phenotypes of β thalassaemia (Thien SL, 2000).Mutations within the consensus sequences at the splice junction reduce the efficiency of normal splicing to varying degree and produce a β+ phenotype that ranges from mild to severe (Jane SM, 1995).

Mutation within introns or exons might also affect the splicing pattern of the pre-mRNAs. For example, a cryptic splice site that contains the sequence GT GGT GAG G has been found in exon 1 of the β globin gene, spanning codons 24-27. Three mutations within this region activate this cryptic site, which acts as an alternative donor site in RNA processing (Thien SL, 2000).

The mutations in codon 26 (GAG→ AAG) that give rise to HbE Codon 26 (Gln→Lys) is one such mutation that activates this cryptic

splice site, with a reduction of the normal splicing that produces the HbE

variant. As HbE production is also quantifiably reduced, the compound heterozygous state, HbE / β thalassaemia, results in a clinical picture closely resembling homozygous β thalassaemia, ranging from severe anaemia and transfusion-dependency to thalassaemia intermedia.

Other RNA processing mutants affect the polyadenylation signal (AATAAA) and the 3'-UTR. These are generally mild β+ thalassaemia alleles. Mutation that abrogate mRNA translation either at the initiation or extension phases of globin synthesis are all associated with a βo phenotype. Approximately half of the β thalassaemia alleles are characterized by premature termination of β chain extension.They result from introduction of premature termination codons due to frameshift or nonsense mutations and nearly all terminate within exons 1 and 2. These mutations are associated with minimal steady state levels of mutant β mRNA in erythroid cells. In heterozygotes for such cases, no β chain is produced from the mutant allele and only half the normal β globin is present, resulting in a typical asymptomatic phenotype.

The 'silent' mutations are normally identified in the compound heterozygous states with a severe β thalassaemia allele, which results in thalassaemia intermedia, or in homozygotes who have typical phenotype of β thalassaemia trait. The 'silent' β thalassaemia alleles are not common, except for the -101