UNIVERSITÀ DEGLI STUDI DI SASSARI

Ph.D COURSE IN LIFE SCIENCES AND BIOTECHNOLOGIES
CURRICULUM BIOCHEMISTRY AND MOLECULAR BIOLOGY

STUDY ON HEMOGLOBINOPATHIES IN CENTRAL PROVINCES OF VIETNAM

Chau Quynh Nguyen

Ph.D Thesis

Tutor:
Professor Monica Pirastru

Coordinator:
Professor Leonardo Sechi

Cycle XXIX (2013 - 2016)

INDEX

SUMMARY

The Vietnamese population is ethnically highly heterogeneous and the spectrum of β-thalassemia alleles is slowly defining. On the whole, six mutations of the β^0- and a mutation of the β^+-thalassemia have been identified and observed with distinct incidence in the different areas. Interaction between these mutations and the rather common Hb E leads to a variety of thalassemia syndromes, in particular to the severe forms of homozygous β-thal and Hb E-β-thal diseases. Nonetheless, epidemiological data is still insufficient and fragmented.

In this Doctoral Thesis, a screening program for hemoglobinopathies was carried out in the central provinces of Vietnam as part of the ongoing cooperation between the Universities of Hue and Sassari.

A study group of 160 subjects referred to Hue Medicine and Pharmacy College and PhuVang District Hospitals for hematological and clinical evaluation and a control group of 193 individuals were included in the study. A total of 89 samples showed abnormal hematological parameters and have undergone to Hbpathies screening. About 30% of them, having qualitative or quantitative alteration in Hb profiles, were investigated by DNA analysis.

Several mutated β-alleles were identified by nucleotide sequencing. Some of them were already described also in other countries of the Southeast Asia with different incidence. On the contrary, the β^+ promoter mutation -72 (T→A), found in a 5-year old child and in his relatives, has never been described before. *In vitro* expression studies were performed into K562 cells. The transcriptional activity of the mutated promoter is roughly half that of the wild type promoter. Although this is a mild thalassemic mutation, it is impossible to predict the severity of the phenotype and its clinical implications in the interaction with severe β-thalassemic allele. Indeed, because of the extremely high gene frequency for HbE and the widespread occurrence of β-thalassemia in Vietnam, the incidence of β^E/β^{Thal} or $\beta^{Thal}/\beta^{Thal}$ compounds is very high.

This result underline the importance of identifying and characterizing new or rare β-thalassemic alleles in carrier screening and prenatal diagnosis in order to reduce the burden of thalassemias, avoid unnecessary transfusions in TI and start early transfusions in TM.

INTRODUCTION

1. Structure and function of hemoglobin

Hemoglobin (Hb) is the iron-containing protein within the red blood cells that is responsible for the transport and exchange of oxygen. Hb is a tetrameric protein composed of four polypeptide chains, two α-globin and two β-like globin polypeptide chains encoded by individual specific genes. The α- and β-globin chains are very similar in structure. Both of them contain primarily α-helix secondary structure with no β-sheets. Each α- or β-globin chain folds into 8 α-helical segments (A-H) which, in turn, fold to form globular tertiary structures. The folded helices form a pocket that holds the heme group. The heme group is composed of a heterocyclic ring, named as porphyrin, which contains a charged iron molecule. The porphyrin ring includes four pyrrole molecules cyclically linked together by methene bridges with the iron ion bound in the center (Figure 1). The nitrogen atoms of the pyrrole molecules form coordinate covalent bonds with four of the iron's six available positions which all lie in one plane.

The iron is bound to the globular protein via the imidazole ring of the F8 histidine residue (also known as the proximal histidine) below the porphyrin ring. A sixth position can reversibly bind oxygen by a coordinate covalent bond, completing the

octahedral group of six ligands.

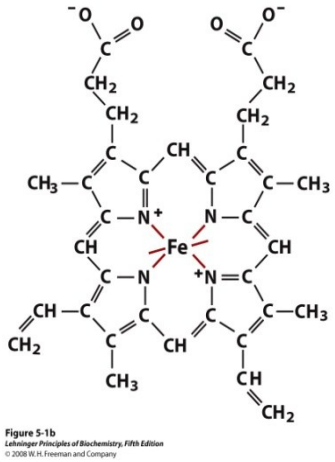


Figure 1. The heme prosthetic group. *The iron atom of heme has six coordination bonds: four in the plane of, and bonded to, the flat porphyrin ring system, and two perpendicular to it.*

The iron ion may be either in the Fe^{2+} or in the Fe^{3+} state, but ferriHb (Fe^{3+}), also called metHb, cannot bind oxygen. In binding, oxygen temporarily and reversibly oxidizes Fe^{2+} to Fe^{3+} while

oxygen temporarily turns into superoxide, thus iron must exist in the +2 oxidation state to bind oxygen. If superoxide ion associated to Fe^{3+} is protonated the Hb iron will remain oxidized and incapable to bind oxygen. In such cases, the enzyme metHb reductase will be able to eventually reactivate metHb by reducing the iron center.

In binding, oxygen molecules leading to the whole Hb molecule undergoing a conformational shift from the low oxygen affinity T-state to the high affinity R-state (Figure 2). The shift initiates when the bound oxygen changes the heme's electronic state, resulting in the Fe-porphyrin bonds together. This causes the heme to flatten out into a nearly planar molecule. When the heme flattens, the proximal histidine is pulled along, for appoximately 0.6 Å. This is only possible with an accompanying rearrangment of the attached F-helix; it is translated about 1Å along the heme plane (Voet & Voet, 2011). The change between the T and R structures is the result of a rotation of 15 degrees between the two αβ dimmers. This rotation changes the bonds between the side chains of the αβ dimers in the F helix and therefore causes the heme molecule to change positions. In the T structure, the iron ion is pulled out of the plane of the porphyrin ring and becomes less accessible for oxygen to bind to it, thus reducing its affinity to oxygen. In the R structure the iron atom is in the plane of the porphyrin ring and is accessible to bind oxygen, thus increasing its oxygen affinity.

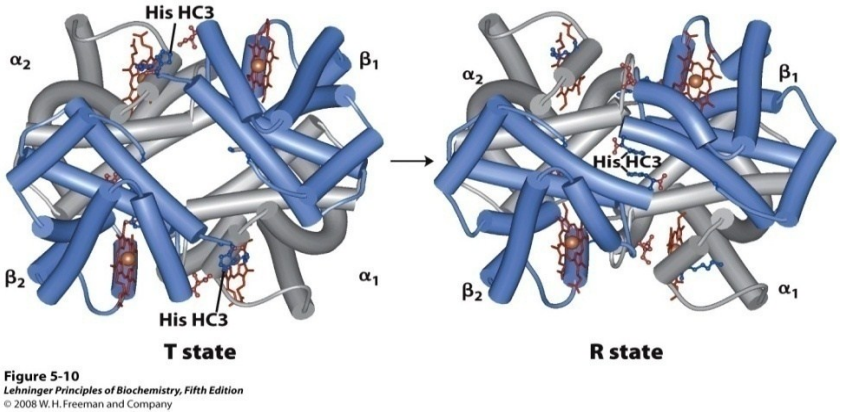


Figure 2. The T → R transition. *The transition from the T state to the R state shifts the subunit pairs substantially, affecting certain ion pairs. Most noticeably, the His HC3 residues at the carboxyl termini of the β subunits, which are involved in ion pairs in the T state, rotate in the R state toward the center of the molecule, where they are no longer in ion pairs. Another dramatic result of the T → R transition is a narrowing of the pocket between the β subunits.*

The capacity of Hb to take up oxygen molecules in the lungs and then release them in the tissues is regulated by several factors both within the Hb molecule itself and through external chemical factors. One of the most enormous regulators of the oxygen affinity of the Hb is the presence of oxygen itself. In the lungs where the oxygen levels are high, the Hb has a higher affinity for oxygen and this affinity increases disproportionately with the number of molecules it already has bound to it (Figure 3). In other words, after the oxyHb binds one molecule of oxygen, its affinity for oxygen increases until the Hb is fully saturated. In the same way, the deoxyHb has a lower affinity for oxygen and this affinity decreases disproportionately with the number of molecules it already has bound. Thus, the loss of one oxygen molecule from the deoxyHb lowers the affinity for the remaining oxygen. This regulation is known as cooperativity and is essential to the functioning of the Hb because it allows the oxyHb to carry the maximum amount of oxygen to the tissues and then allows the deoxyHb to release the maximum amount of oxygen into the tissues. Cooperativity is a function of the Hb's unique structural characteristics, and it was found that the cooperative effects of the Hb totally disappear if the Hb is split in half . Essentially, Hb is an allosteric protein that has more than one shape and can undergo conformational changes in its structure based on environment conditions.

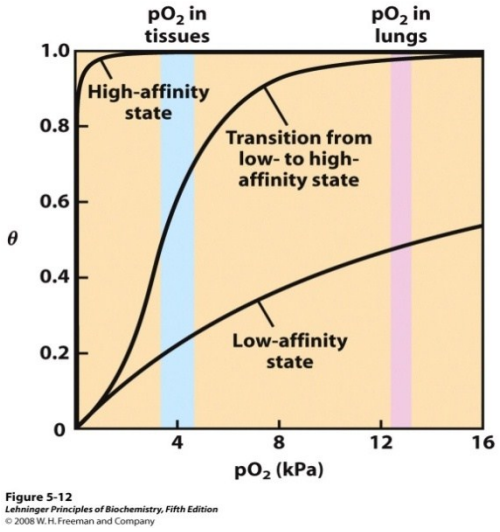


Figure 3. The sigmoid (cooperative) binding curve of human Hb*. The sigmoid binding curve can be analyzed as a hybrid curve reflecting the transition from a low-affinity to a high-affinity state. Because of its cooperative binding, as manifested by a sigmoid binding curve, Hb is more sensitive to the small differences in O_2 concentration between the tissues and the lungs, allowing it to bind oxygen in the lungs (where pO_2 is high) and release it in the tissues (where pO_2 is low).*

The oxygen affinity of Hb can also be regulated by external chemical factors including pH, carbon dioxide, and BPG (2,3-bisphosphoglycerate) (Figures 4,5). In general any chemical agents that strengthen the bonds between the α subunits and prevent the rotation to the R structure decrease the oxygen affinity of the Hb.

When CO_2 is released into the blood from the tissues it acidifies the blood by increasing the concentration of hydrogen ions. This lowering in pH causes the oxygen affinity of the Hb to decrease, which is known as the Bohr effect. The molecular basis behind the Bohr effect is that the T structure of Hb binds hydrogen more readily than the R structure, so under a condition of low pH (high hydrogen ion concentration) the T structure, which has a decreased oxygen affinity, dominates.

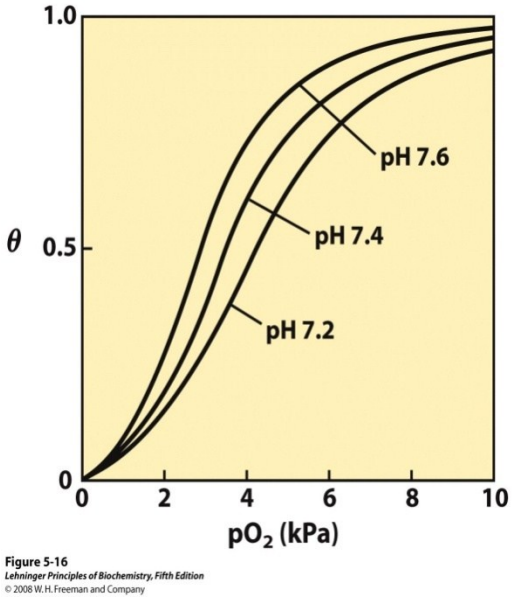


Figure 4. Effect of pH on oxygen binding to Hb. *The pH of blood is 7.6 in the lungs and 7.2 in the tissues.*

CO_2 has a similar effect on the Hb, but instead of binding to the heme molecule like oxygen, CO_2 binds to the N-terminus of the α-globin molecule. The CO_2 binds better to the globin in the T structure, so the release of oxygen in the tissues by the T structure of Hb facilitates the uptake of CO_2. Then in the lungs, the uptake of oxygen causes the Hb to change to the R structure, which causes the release of the CO_2 into the lungs, because CO_2 does not bind as well to the R structure.

Finally, BPG is an allosteric effector that changes the oxygen affinity of Hb by binding to the Hb itself. BPG can bind to the T structure of Hb, because of the change in structural conformation which allows it to fit, but cannot bind to Hb in the R structure. Therefore, the presence of BPG lowers the oxygen affinity by keeping the Hb in the T structure longer.

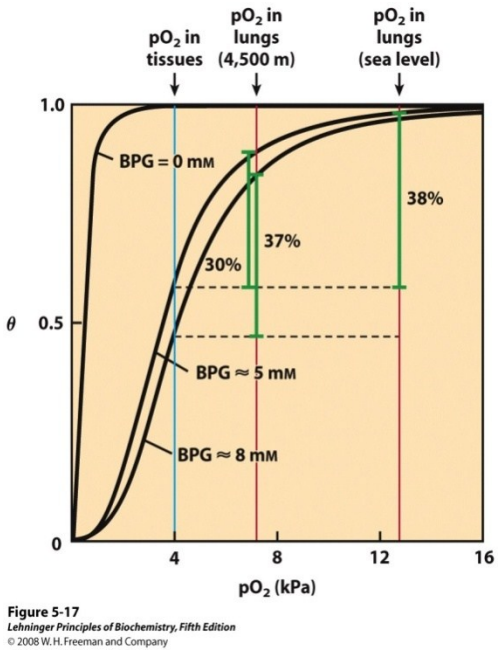


Figure 5-17
Lehninger Principles of Biochemistry, Fifth Edition

Figure 5. Effect of BPG on oxygen binding to Hb.

The allosteric regulation of deoxyHb-mediated nitrite reduction can yield bioactive NO (Nitric oxide). Moreover, Hb cross-linking and conformational locking can dramatically impact on nitrite-reduction kinetics, NO-formation and bioactivity thereby resulting in Hbs that can more effectively couple nitrite-reduction to stimulation of NO-bioactivity. Nitrite reduction kinetics measured by deoxyHb consumption do not necessarily reflect the yield of NO. In this regard, the determination of NO gas production and assessment of NO-dependent signaling provides a more accurate indicator of the ability of different Hbs to produce bioactive NO from nitrite reduction. This observation might be of particular relevance when assessing the ability of different HBOCs (Hb-based oxygen carrier) to function as nitrite reductase under physiological condition (Cantu-Medellin et al., 2011).

2. The human globin genes

Hematopoiesis during mammalian development is characterized by the progressive appearance of distinct populations of cells at stage-specific sites within the embryo. In concert with the progression from primitive to definitive hematopoiesis, the developing erythroid system expresses stage-specific forms of Hb in a process known as Hb switching.

In humans, six different Hb types are produced: Hb Gower I ($\zeta_2\varepsilon_2$), Gower II ($\alpha_2\varepsilon_2$), and Portland ($\zeta_2\gamma_2$) are found in the embryo; fetal Hb (HbF; $\alpha_2\gamma_2$) is present mainly in the fetus, but also in the embryo and adult, whereas HbA ($\alpha_2\beta_2$) and HbA_2 ($\alpha_2\delta_2$) are seen in adults. Also, all Hbs undergo post–translational modification forming minor Hb such as HbF_1 (acetylated form of HbF) or HbA_{1C} (non–enzymatic glicated form of HbA).

Embryonic Hb are expressed in primitive erythroblasts developing in the yolk sac during the first several weeks after conception. The first major Hb switching event occurs as ζ- and ε-globin expression ceases and α- and γ-globin synthesis begins, leading to production of HbF (Figure 6). These events are coincident with the transition of the site of erythropoiesis from the yolk sac to the fetal liver. The second switch involves the perinatal decline of HbF synthesis coupled with the increased synthesis of the adult form HbA. After 2 years of age, HbF is present as a minor component of total Hb in only a few percentage of mature red blood cells in healthy persons (Steinberg et al., 2009).

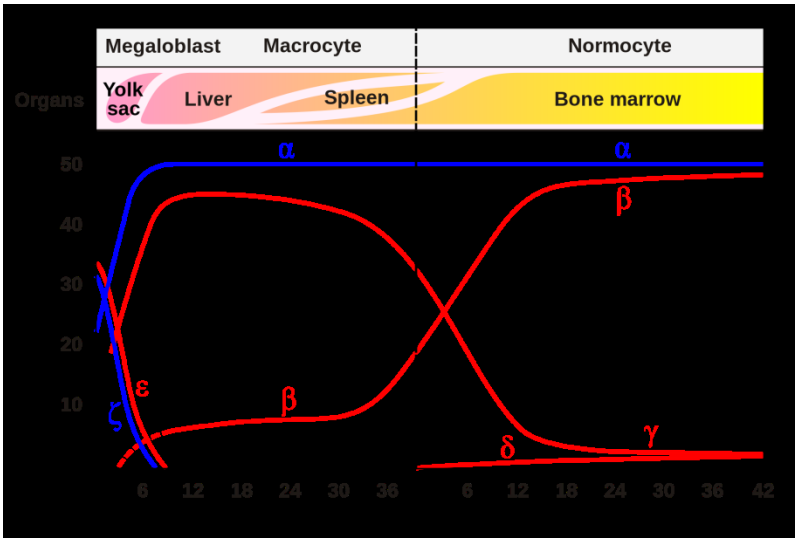


Figure 6. Switch of the human globin genes and location of erythropoiesis.

The genes encoding the different globin chains of Hb are members of an ancient gene family and they share a common structural organization.

Before the precise knowledge of globin gene organization, it was clear that there must exist at least one gene for each of the different globin chain (ζ, α, ε, γ, β, δ). Further studies based on HbF heterogeneity and globin variants showed that α and γ genes are duplicated and structurally different: the two α genes encode the same protein, while γ genes are different only for one amino acid (136 Gly/Ala: $^{G}\gamma$ and $^{A}\gamma$, respectively) (Steinberg et al., 2009).

The coding region of each globin gene in humans and other vertebrates is interrupted at two positions by stretches of noncoding DNA called intervening sequences (IVSs) or introns (Tilghman et al., 1978). In the β-like globin genes, the introns interrupt the sequence between codons 30 and 31 and between codons 104 and 105; in the α-globin gene family, the IVSs interfere the coding sequence between codons 31 and 32 and between codons 99 and 100 (Figure 7). Despite the precise codon position numbers at which the interruption occurs differ between the α- and β-like globin genes, the introns occur at precisely the same position in the aligned primary sequence of the α- and β-globin chains. The first intervening sequence (IVS-1) is shorter than the second intervening sequence (IVS-2) in both α - and β -globin genes, but IVS-2 of the human β -globin gene is much larger than that of the α -globin gene (Steinberg et al.,2009).

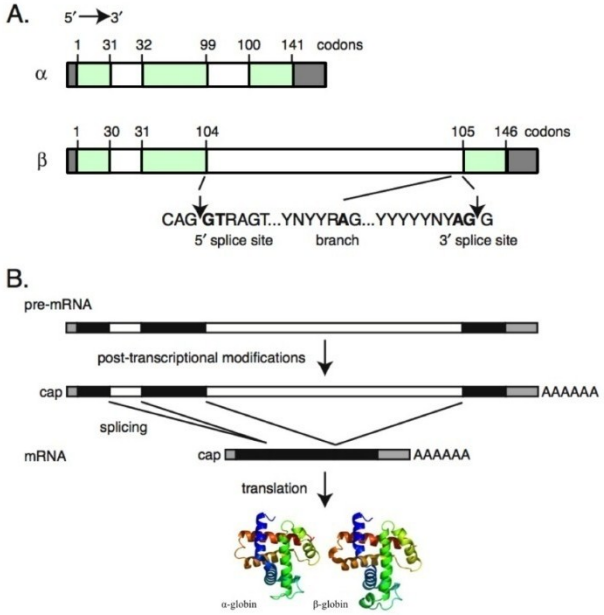
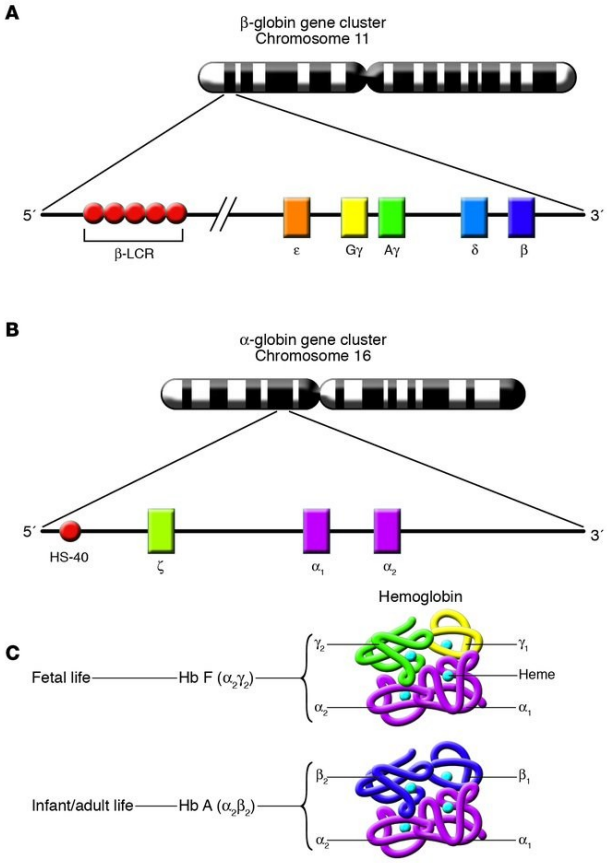


Figure 7. *A)* ***General structure of globin genes. B) The pathway for maturation and expression of mRNA of the globin genes.***

The α–like and β–like genes map on different chromosomes: the so called "β–globin or HBB cluster" maps on chromosome 11 (11p15.5), while "α–globin or HBA



cluster" locates on chromosome 16 (16p13.3), 150 kb from the telomere of its short arm (Figure 8).

Figure 8. Basic organization of human β- (A) and α-globin (B) gene clusters and composition of Hb produced (C).

The conceptually simple pattern of globin gene expression during development is the result of a complex series of regulatory events. Numerous epigenetic and transcriptional regulators are necessary for switching to occur, in what remains an incompletely understood process, despite several decades of research.

The synthesis of the α–like and β–like globin chains are balanced throughout human development. Imbalanced globin chain synthesis caused by the reduction of one of the globin chains results in thalassemias.

To facilitate their coordinately regulated transcription, the α- and β-globin loci are structurally similar with a series of evolutionarily conserved regulatory elements that direct tissue-specific transcription of far-downstream globin genes. Furthermore, the genes are arranged in the order of their developmental expression with primitive (embryonic) globin genes followed by definitive (fetal/adult) globin genes. To account for their similarity in genomic structure, the α and β loci arose by duplication of an ancestral locus at an early stage of vertebrate evolution ~500 million years ago and have been constrained by purifying selection (Steinberg et al., 2009).

In the globin gene clusters, there was the presence of additional gene-like structures with sequence homology and an exon–intron structure similar to the actively expressed globin genes. These DNA segments have been called pseudogenes. One, called ψβ1, is in the HBB cluster between the γ- and δ- globin genes. At least two (and possibly four) are in the HBA cluster. The two clear examples are ψζ1 and ψα1, located between the active ζ-globin and α-globin genes. All three (ψβ1, ψζ1 and ψα1) are characterized by the presence of one or more mutations that render them incapable of encoding a functional globin chain. This inability to encode a functional globin polypeptide does not necessarily render the pseudogenes inactive for transcription. The pseudogene ψβ1 is transcribed and spliced, as shown by several spliced expressed sequence tags, whereas no evidence has been provided that ψα1 is transcribed. These pseudogenes appear to have arisen by gene duplication events within the globin gene clusters followed by mutation and inactivation of the duplicated gene and subsequent accumulation of additional mutations through loss of selective pressure. Two other α-like globin genes have been identified and characterized in the HBA cluster, but their roles, if any, in encoding globin polypeptides are still uncertain (Steinberg et al., 2009).

The expression pattern of globin genes depends on local cis-acting sequences at the promoter sequences. The globin gene promoters share a number of characteristic sequences at specific distances from the transcriptional start site (TSS). These sequences include TATA, CCAAT and CACCC which are situated at 30, 70-78, and 80- 140 bp upstream of the TSS, respectively. The TATA and CCAAT sequences are found in many eukaryotic promoters while the CACCC sequence exists predominantly in erythroid specific genes. Importance of these sequences for normal globin gene expression has been demonstrated by the presence of natural mutations at these sites that lead to down regulation of the globin genes causing thalassemia (par. 4.2.). Other regulatory sequences that control the globin genes include the enhancer and silencer elements

that interact with protein complexes to influence promoter activity and transcription of target genes. Remote cis-acting sequences, such as the locus control region (LCR) of the HBB cluster and the major regulatory element (MRE) of the HBA cluster are also crucial for the expression of the globin genes. The LCR spans a 16kb region which starts 5kb upstream of β-globin gene. The function was identified as essential for high level expression of these genes (Tuan et al., 1985; Forrester et al., 1986; Grosveld et al., 1987; Bender et al., 2000). Molecular experiments such as DNAse Hypersensitivity assays have highlighted the active parts of the β-globin locus. Genes that have an 'open' chromatin structure (euchromatin) are transcriptionally active, whereas transcriptionally silent genes have a 'closed' chromatin structure (heterochromatin). Open chromatin is hypersensitive to nuclease digestion while closed chromatin is resistant to DNase I digestion. Thus, transcriptionally active sites can be identified by their hypersensitivity to DNase I. In 1980s Tuan et al. and Groudine et al. discovered the 5 hypersensitive sites (HS1-5) that are located upstream of the β-globin gene. These sites are phylogenetically conserved sequences and they play a crucial role in the regulation of the β-globin gene locus (Forrester et al., 1986, Stamatoyannopoulos et al., 2005; Stamatoyannopoulos et al., 2009). There are 2 more hypersensitive sites that are positioned further upstream but only HS1-5 were shown to be erythroid specific (Stamatoyannopoulos et al., 2005). HS1-4 are hypersensitive to DNase I in erythroid cells, whereas HS5 is ubiquitous in many non-erythroid cells. The hypersensitive site cores are 200-300 bp and contain binding sites for ubiquitous and erythroid specific transcription factors (Hardison et al., 1997).

3. HBB cluster regulation

The β-LCR has important functions, such as enhancement of transcription, remodeling the chromosome structure through histone modifications, acting as an insulator to regulate enhancer promoter interactions and enabling the correct positioning of the β-globin locus in the nucleus (Stamatoyannopoulos et al., 2005; Bender et al., 2006). Various studies supported the importance of the LCR for globin regulation; as evidenced by results that deletions of the β-LCR in transgenic mice leads to reduction of globin gene expression by 100-fold. Another study showed that deletion of individual HS sites in the β-LCR disrupts the activation of globin genes proving that an intact LCR is necessary for activation of the globin genes (Milot et al., 1996; Bender et

al., 2000). At the transcriptional level, specific sequences in the promoters compete with other promoters to interact with the LCR through binding of transcription factors. Binding of these transcription factors is regulated through remodeling of the chromatin by various histone modifications that enables activation of different parts of the globin locus during development. Acetylation and methylation are the histone modification observed at the β-globin locus. Acetylation mainly harbors an open state of chromatin thus acts as an activator. On the other hand methylation may act both as an activator and repressor depending on the site that is methylated.

Several models have been proposed to explain how the β-LCR functions to activate the downstream globin genes and of these, the looping model is the most widely accepted (Figure 9). This model suggests that the HS sites of the β-LCR fold into a holocomplex, bringing the LCR closer to the appropriate gene and thereby the transcription factors bound to the LCR are delivered to the gene locus to interact with the basal transcriptional apparatus and activate globin gene expression (Stamatoyannopoulos et al., 2005).

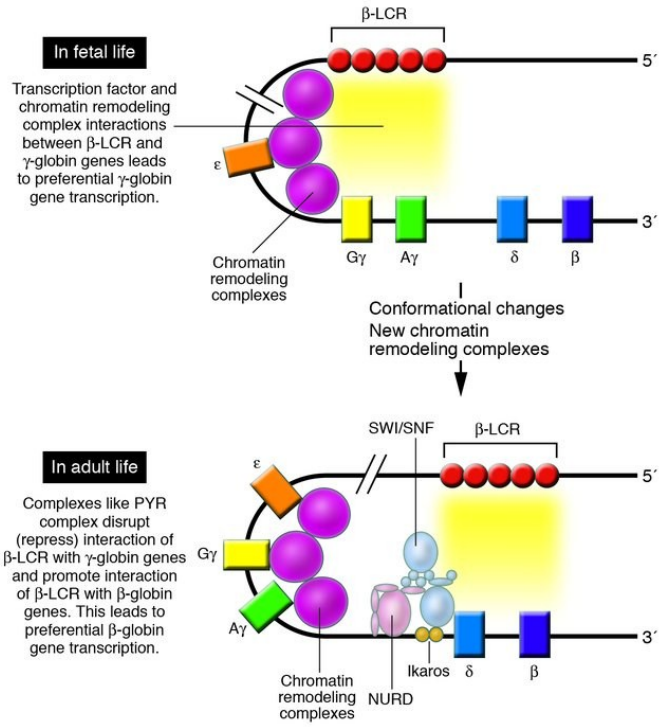


Figure 9. ***Model of transcription complex recruitment to the β-globin gene.*** *The LCR and TS interact to form a chromatin hub. In the active chromatin hub, the expressed genes interact with the HS sites. The genes come in close proximity to the LCR holocomplex by as yet unknown mechanisms that may involve local remodeling of chromatin structure at the active promoters. Transcription complexes are then transferred from the LCR to high affinity binding sites at the globin gene promoters.*

Studies have demonstrated the interaction of the LCR with downstream globin gene promoters via the looping model. Carter et al. (2002) used an RNA FISH method called RNA TRAP to tag and recover chromatin near the murine β-globin gene which recovered HS1, HS2 and HS3 together with the β-globin gene. This demonstrated that the HS sites and β-globin gene were in close proximity.

A further study used chromatin conformation capture to measure the proximity of LCR sequences and active globin genes in erythroid cells. Clustering of active regulatory elements at these sequences was referred to as an active chromatin hub (ACH) (Tolhuis et al., 2002).

Transcription factors play an important part in gene regulation at the HBB locus. The HS sites all share a core sequence ~250 nucleotides long, which is studded with motifs for transcription factors. The HS sites contain binding sites for erythroid specific transcription factors NF-E2, GATA-1 USF, Sp1 and KLF1 (Harju et al., 2005). NF-E2 is an erythroid-specific DNA-binding protein that recognizes motifs (GCTGA(G/C)TCA) in the 5'HS2 of the β-LCR (Talbot & Grosveld, 1991; Stamatoyannopoulos et al., 1995). Here, the NF-E2 binding sites are important for transcriptional activation and formation of HS sites in the LCR (Forsberg et al*.,* 2000; Ney et al*.,* 1990). Experimental evidence also shows NF-E2 binding sites are important for chromatin remodelling activity and necessary for ε-globin gene expression and formation of HS2 (Gong et al*.,* 1996). GATA-1 is also an erythroid specific transcription factor required for globin gene switching and erythroid cell maturation. It is a member of the GATA zinc finger family of transcription factors which bind to nucleic acid consensus sequence (T/A)GATA(A/G) (Evans & Felsenfeld, 1989; Tsai et al., 1989). GATA-1 binds to sites in the globin gene promoters and HS site cores of HS1-5 which contain the GATA-1 recognition sequence (Orkin, 1992). GATA-1 can also act as an activator when bound to the γ-globin gene promoter or HS1-5 (Jane et al., 1993; Stamatoyannopoulos et al., 1995). GATA-1 is also an activator of the ε-globin gene but acts as a repressor of ε-globin when bound to the ε-globin gene silencer with transcription factor YY1 (Li et al., 1998). GATA-1 is known to interact with itself and other transcription factors like SP1 and KLF1 (Merika & Orkin, 1995). Krüppel-like factor 1, KLF1 (previously known as erythroid Krüppel-like factor, EKLF) was identified by subtractive hybridization and found to be an erythroid cell-specific transcription factor, homologous to the Krüppel family of transcription factors which have roles in cell proliferation, differentiation and survival (Miller & Bieker,

1993). KLF1 contains three zinc fingers at the C-terminus which bind to a CACCC sequence. CACCC sequences are repeated in erythroid enhancers and promoters, including the β-globin gene promoter (Miller & Bieker, 1993). This sequence is noted as a site of point mutations that give rise to β-thalassemia. It is reported by Feng et al. (1994) that KLF1 is unable to transactivate in the presence of these point mutations due to a decrease in binding affinity for these target sites.

The KLF1 protein binds to the β-globin promoter with 8 fold higher affinity than to the γ-globin promoter (Donze et al., 1995). These studies showed KLF1 as a stage-specific, β-globin-specific transcription factor and proposed that it is most likely an important factor in the fetal to adult Hb switch. Further studies in mice revealed that *KLF1* -/- mice die around day 14-15 of gestation due to anemia caused by the failure to express β-globin. Moreover, the embryos showed features of β-globin deficiency as found in β-thalassemia. *KLF1*-/- embryos appeared normal during embryonic yolk sac stage of hematopoiesis, but became fatally anemic during early fetal life, at the precise time of the switch from embryonic to fetal-liver erythropoiesis (Nuez et al., 1995, Perkins et al., 1995).

KLF1 binding also enhances the interaction between the β-globin promoter and the LCR and so the interaction between the γ-globin promoter and LCR decreases (Miller & Bieker 1993, Donze et al., 1995). The human β-globin locus was studied in *KLF1*-/-/human β-locus transgenic mice (Wijgerde et al., 1996). The ε and γ globin genes were expressed normally in *KLF1-/-* fetuses, with a complete lack of β-globin expression. In *KLF1*+/- /β-globin transgenic mice, there was a shift in the γ to β ratio caused by an increase in the number of actively transcribed γ genes and a decrease in transcribed β genes. The authors proposed that the reduction in KLF1 in the *KLF1*+/- mice reduced the time the LCR was in complex with the β-globin gene. Because KLF1 does not directly bind γ-globin, an increase in γ-globin is due to the LCR being less occupied with β-globin, therefore less β-globin activation and more interaction is formed between the LCR and γ-globin (Wijgerde et al., 1996). This was accompanied by changes in chromatin structure at the β-globin promoter, and HS3 of the β-LCR. Following this work, KLF1 was found to be required for the activity of 5'HS3 of the β-globin LCR by binding directly to the core fragment within HS3. Increasing levels of KLF1 lead to changes in the balance from γ to β-globin gene expression, which results in an earlier switch of the globin genes and the amount of KLF1 also influences the rate of the switching process (Tewari et al., 1998). These studies, collectively

demonstrate the importance of KLF1 in the γ to β globin switch and LCR-β-globin gene interactions.

4. Disorders of hemoglobin

To date, more than 1600 disorders of Hb synthesis and of the structure have been identified. These are collected, and continuously update, in the Globin Gene Server (HbVar, http://globin.cse.psu.edu/).

Hb disorders can be broadly classified into two general categories: qualitative Hemoglobinopathies (Hbpathies) in which there is a structural defect in one of the globin subunits, and quantitative Hbpathies characterized by the absence or reduced output of one or more of the globin chains. These are called the thalassemia syndromes.

4.1. Qualitative disorders of globin structure

The qualitative Hbpthies class is mainly composed of missense mutations that cause single amino acid substitutions in the globin protein, resulting in an abnormal, or "variant" Hb tetramer. Less commonly, Hb variants are associated with deletions or multiple amino acid substitutions (Thom et al., 2013).

The majority of human Hb mutants were discovered as an incidental finding, unassociated with any hematologic or clinical phenotype since most commonly the amino acid change is innocuous, perturbing neither the structure nor function of the Hb molecule. Occasionally, alteration of a single amino acid dramatically perturbs the behavior of the Hb and produces a disease state. The most common and medically important Hb variants include HbS [β6(A3)Glu→Val], HbC [β6(A3)Glu→Lys] and HbE [β26(B8) Glu→Lys]. In addition to these prevalent mutant proteins, there are also >1000 other known naturally occurring Hb variants, which are rare individually but common collectively (Forget & Bunn, 2013).

Sickle Cell Disorders

Sickle Hb (HbS) results from an amino acid substitution at the sixth residue of the β-globin subunit: β6 Glu→Val. Approximately 8% of African Americans are heterozygous for this Hb variant, a condition called sickle cell trait or HbAS. In equatorial Africa, where malaria is endemic, the prevalence of HbAS is much higher and can reach over 30% in some populations because of survival advantage of HbAS heterozygotes from complications of falciparum malaria.

RBCs of persons with HbAS typically have 40% HbS and 56%-58% HbA. Individuals with HbAS are typically asymptomatic; severe hypoxia is required for them to experience manifestations of sickle cell disease, called sickling. The basis of sickling in patients homozygous for the disorder, called sickle cell anemia or HbSS, is polymerization of deoxy-HbS resulting in the formation of multistranded fibers that create a gel and change the shape of RBCs from biconcave discs to elongated crescents. The polymerization/sickling reaction is reversible following reoxygenetion of the Hb. Thus, an RBC can undergo repeated cycles of sickling and unsickling.

There are two major pathophysiological consequences of sickling: repeated cycles of sickling damage the red blood cell membrane leading to abnormalities of permeability and cellular dehydration, eventually causing premature destruction of RBCs and a chronic hemolytic anemia. Moreover, sickled RBCs are rigid, increase blood viscosity and obstruct capillary flow, causing tissue hypoxia and, if prolonged, cell death, tissue necrosis/infarction, and progressive organ damage (Forget & Bunn, 2013).

The amount and type of other non-S Hb variants in the RBC can influence the extent or rate of sickling, and thus clinical severity. HbSC disease is associated with significant clinical manifestations. HbSC patients are compound heterozygotes for β^S/β^C alleles. HbC is caused by a mutation in the sixth position of the β-globin chain, where glutamic acid is substituted by lysine. Two independent factors conspire to make HbSC a disease. The presence of HbC in the RBC greatly enhances potassium efflux and cell dehydration, which increases corpuscular Hb concentration (Bunn et al., 1982) and promotes polymerization. Sickling in SC patients is further enhanced by the fact that they have ~50% HbS, whereas HbAS individuals have ~40% HbS. This difference is because α-globin subunits bind more readily to negatively charged β^A-subunits than to positively charged β^C-subunits (Forget & Bunn, 2013).

Unstable Hb variants

A substantial minority of Hb mutants have substitutions that modify the solubility of the molecule in the red cell. Mutations that alter any step in globin processing, including subunit folding, heme interaction, dimerization or tetramerization, can destabilize Hb. Five general mechanisms can occur: amino acid substitutions within the heme pocket, disruption of secondary

structure, substitution in the hydrophobic interior of the subunit, amino acid deletions and elongation of the subunit (Bunn & Forget, 1986).

The intraerythrocytic precipitated material derived from the unstable abnormal Hb is detectable by a supravital stain as dark globular aggregates called Heinz bodies. These intracellular inclusions reduce the life span of the red cell and generate a hemolytic process of varied severity called congenital Heinz body hemolytic anemia. When a red cell hemolysate of an affected individual is heated to 50°C or treated with 17% isopropanol, a precipitate usually develops. The Hb electrophoresis often reveals an abnormal banding pattern. Definitive diagnosis requires analysis of either globin structure or DNA sequence (Forget & Bunn, 2013).

Hb variants with altered oxygen affinity

Over 25 Hb variants have been encountered in individuals with erythrocytosis. Amino acid substitutions cause an increase in oxygen affinity usually because stabilize the R state relative to the T state and/or inhibit responses to environmental allosteric regulators that stimulate O_2 release, including H^+ (Bohr effect) or 2,3-BPG. Because T to R state transitions are mediated largely through α1β2 interactions, high affinity variants frequently result from substitutions that alter this interface.

About half of the high-affinity variants can be detected by Hb electrophoresis. Definitive diagnosis is established by demonstration of a "shift to the left" in the oxyHb binding curve.
Low O_2 affinity Hb variants typically present with cyanosis. In general, these variants are caused by globin amino acid substitutions that tip the quaternary equilibrium of Hb tetramers from the R to the T state. This does not inhibit Hb-O2 release in tissue capillaries, but rather, interferes with Hb-O2 uptake if the P50 has increased to ≥50 mm Hg. Paradoxically, low O_2 affinity Hb variants can be associated with mild anemia thought to be caused by increased O_2 tissue delivery with reduced erythropoietic drive. In addition, some low O_2 affinity mutants are unstable and therefore associated with not only cyanosis but also Heinz body hemolytic anemia (Thom et al., 2013).

MetHb ("M-Type") variants

Hb iron must be in its reduced (Fe^{2+}, ferrous) state to bind O_2. Moreover, oxidized (Fe^{3+}, ferric, met) Hb is intrinsically unstable with a tendency to release heme. Hb reduction is maintained

through intrinsic features of the Hb protein and extrinsic antioxidant pathways within RBCs. Exposure to oxidant drugs or toxins, genetic alterations in erythroid metHb reductase enzyme systems, or globin chain variants can predispose to methemoglobinemia. These disorders present as “pseudocyanosis,”, despite adequate arterial oxygenation.

M-type Hb variants are predisposed to spontaneous oxidation. Globin variants associated with metHb formation are typically caused by amino acid substitutions within the heme pocket. For example, several different M-Hbs occur when the α- or β-globin proximal or distal histidine residues are replaced. Comparative studies of the M-Hbs have contributed greatly to understanding the biochemical properties of the heme iron, including its interactions with various ligands and nearby amino acids such as the proximal and distal histidines (Thom et al., 2013).

Globin Chain Elongation Mutants

Antitermination and frameshift mutations that add irrelevant amino acids to the carboxyl terminus of globin proteins produce interesting variants that can damage erythrocytes. The most clinically significant example is Hb Constant Spring (HbCS: α142Stop→Gln), caused by an antitermination mutation at the α2 stop codon (Clegg et al., 1971; Efremov et al., 1971; Milner et al., 1971; Clegg & Weatherall, 1974). This elongates the protein by 31 amino acids, generating an unstable protein that is relatively underrepresented in hemolysates. In addition, HbCS mRNA is rapidly degraded in developing erythrocytes, owing to ribosomal entry into the 3'UTR, causing displacement of RNA-bound stabilizing proteins with a resultant thalassemia syndrome (Thom et al., 2013). HbCS contributes to α-thalassemia syndromes, particularly when combined with two α-globin deletional alleles (–/αCSα), which produces a severe form of HbH disease (par. 4.4.).

4.2. The thalassemia syndromes

The thalassemias are a heterogenous group of inherited disorders in which there is a quantitative defect in the production of one of the globin subunits, either total absence or marked reduction. Although there is a dearth of the affected Hb subunit, with most cases the few subunits synthesized are structurally normal.

The thalassemia syndromes are most prevalent in the Mediterranean region, the Middle East, the Indian subcontinent and South-East Asia, representing a serious health problem in certain areas where gene frequencies reach 3-10% of the population. The severity of β -thalassemia is directly linked to the degree of imbalance in the production of α- and β-globin chains (Weatherall &

Clegg, 2001; Orkin et al., 1998; Thein & Wood, 2009). They can be classified at several levels. First, there is a clinical classification, which simply describes the degree of severity. Second, the thalassemias can be defined by the particular globin chain that is synthesized at a reduced rate. Finally, there is a molecular classification relating to the specific mutation that is responsible for defective globin chain synthesis.

Under normal circumstances, the rate of synthesis of α-globin must be more or less matched by the total synthesis of β-, δ- and γ-globin chains. The myriad manifestations of this complex group of disorders result from the imbalanced synthesis of α-like and non-α-like globin chains and from the accumulation of unpaired counterpart.

Unpaired globin chains are unstable: they form aggregates and precipitate within the cell, causing decreased deformability, membrane damage and selective removal of the damaged cell. Ineffective erytropoiesis and shortened red cell survival will lead to chronic anemia.

4.3. The β-thalassemias

In the β-thalassaemias there is absent or decreased production of β-globin. Unlike α-thalassemia, in which deletions in the α-globin gene cluster account for most of the mutations (par 4.4.1), the vast majority of β-thalassemias are caused by mutations involving one (or a limited number of nucleotides) within the β gene or its immediate flanking regions (Giardine et al., 2011). Rare deletion forms of β-thalassemia have also been described. One of these deletions is caused by "unequal" crossing over between the linked and partially homologous δ- and β-globin genes, resulting in the formation of a fusion δβ-globin gene, the Lepore gene, that has a low level of expression. Large deletions involving part or all of the β-globin gene cluster are responsible for the δβ-thalassemias and the εγδβ-thalassemias.

The β-thalassemias can be subclassified into those in which there is total absence of normal β-globin subunit synthesis, the β^0-thalassemias, and those in which some structurally normal β-globin subunits are synthesized, but in markedly decreased amounts, the β^+-thalassemias. The molecular basis of the β-thalassemias are very heterogeneous, with almost 300 different mutations having been described (Thein, 2013). Despite the marked heterogeneity in the molecular basis of the β-thalassemias, the clinical phenotype of these disorders is relatively homogeneous because of their common pathophysiology: deficiency of HbA tetramers and excess accumulation of free α-subunits incapable of forming Hb tetramers (Nienhuis & Nathan, 2012).

According to the severity of the phenotype, β-thalassemias can be subclassified in three different groups:

β-thalassemia trait or β-thalassemia minor (Tm) (heterozygotes for a defective β-gene). There is mild to moderate hypochromic microcytic (mean corpuscular Hb [MCH] 18–24 pg; normal values 25-35, mean corpuscular volume [MCV] 65–80 fl; normal values 80-102) anemia, without evidence of hemolysis. Heterozygotes are characterized by an increased proportion of HbA_2 (normal <3.2%, β–thalassemia trait 3.5%–6.0%), and HbF levels that vary from normal (<1.0%) to slightly raised (1.0%–3.0%) (Forget & Bunn, 2013).

β-thalassemia major (TM) (homozygotes or compound heterozygotes), is characterized by a severe transfusion-dependent hemolytic anemia associated with marked ineffective erythropoiesis resulting in destruction of erythroid precursor cells in the bone marrow. Clinical presentation of TM occurs between 6 and 24 months. Affected infants fail to thrive and become progressively pale. Feeding problems, diarrhea, irritability, recurrent bouts of fever, and progressive enlargement of the abdomen caused by spleen and liver enlargement may occur. If a regular transfusion program that maintains a minimum Hb concentration of 9.5 to 10.5 g/dL is initiated, growth and development tends to be normal up to 10 to 12 years. Transfused patients may develop complications related to iron overload. Complications of iron overload in children include growth retardation and failure or delay of sexual maturation. Later iron overload related complications include involvement of the heart, liver and endocrine glands. Individuals who have not been regularly transfused usually die before the second-third decade. Survival of individuals who have been regularly transfused and treated with appropriate chelation extends beyond age of 40 years (Galanello & Origa, 2010).

Most patients with TM carry two thalassemic β-globin genes. However, several cases of TM are caused by a compound heterozygosity with a single mutant β-globin gene. For example, HbE, a common Hb variant in Southeast Asia, contains an amino acid substitution that renders β chains mildly unstable *in vitro* with minimal clinical significance. However, this mutation also creates an alternate splice site in the β-globin mRNA, leading to reduced synthesis of productive transcripts with resultant thalassemia (Orkin et al., 1982). HbE carriers usually have no clinical symptoms, but HbE is particularly deleterious when coinherited with a more severe β-thalassemic allele. The frequency of HbE varies from region to region and differs among different ethnic groups. High prevalence of HbE is reported in various distinct groups in

Southeast Asia, ranging from 5% to 10% in the region overall, and as high as 50% in some groups of Cambodia and Thailand (Lithanatudom et al., 2016).

β-thalassemia intermedia (TI) is a less common clinical phenotype. Individuals with TI present later than TM, have milder anemia and by definition do not require or only occasionally require transfusion. At the severe end of the clinical spectrum, patients present between the ages of 2 and 6 years and although they are capable of surviving without regular blood transfusion, growth and development are retarded. At the other end of the spectrum are patients who are completely asymptomatic until adult life with only mild anemia.

TI patients have a milder disease because there is less severe α- to non-α-globin subunit imbalance than in typical TM patients, resulting in less accumulation of free α–subunits that cause the ineffective erythropoiesis. There are different possible causes for such a lowered α- to non-α-globin subunit imbalance, including: inheritance of milder β^{+}-thalassemia mutations with less severe than usual deficiency of β–globin subunit production; coinheritance of a form of α–thalassemia; coinheritance of determinants associated with increased production of the γ–subunit of HbF. Most patients with TI carry two mutant β-globin genes: they have a genotype typical of TM, but the phenotype is modified by one of the factors listed above. However, rare cases of TI are caused by heterozygosity for a single mutant β-globin gene associated with the production of a highly unstable β-globin subunit that causes RBC damage in a fashion similar to excess free α-subunits; this is the so-called "dominantly inherited β-thalassemia" (par. 4.3.1) (Thom et al., 2013).

The δβ-thalassemias (par. 4.3.1.) are associated with total deficiency of β-globin subunit production, but are clinically milder than typical cases of β^{0}-thalassemia, because there is an associated persistent high level of expression of the γ-subunit of HbF that decreases the degree of α-subunit excess.

The Hereditary Persistence of Fetal Hemoglobin (HPFH) syndromes (par. 4.3.1) are not, strictly speaking, a form of β-thalassemia because it is not associated with significant α- to non- α-globin subunit imbalance, but is characterized by high levels of persistent γ-globin production and is frequently considered within the spectrum of δβ-thalassemia (Forget & Bunn, 2013).

4.3.1. The molecular basis of β-thalassemia

Transcriptional mutants

Mutations affecting transcription involve the conserved DNA sequences that form the β-globin promoter (from 100 bp upstream to the TSS, including the functionally important CACCC, CCAAT, and ATAA boxes) and the stretch of 50 nucleotides in the 5' untranslated region (5'-UTR). Mutations in all of the three conserved sequence motifs in the β promoter, the two CACCC, CCAAT and ATAA boxes, have been identified in different patients with β-thalassemia. In general, the degree of diminished β-globin synthesis associated with mutations of the β-globin gene promoter within the CCAAT box is relatively minor. This finding is consistent with transcription studies of the mutant genes in tissue culture cells, which reveal only a mild-to-moderate decrease in transcriptional activity of these genes (Thein & Wood, 2009.

Of the mutations affecting transcription, the C→T change at position -101 to the β-globin gene appears to cause an extremely mild deficit of β-globin. The allele is so mild that heterozygotes are "silent" with borderline reduced/normal red cell indices (Maragoudaki et al., 1999).

Mutations of the Splice Site Junction

Sequences critical in the splicing process include the invariant dinucleotides GT at the 5' (Donor) and AG at the 3' (Acceptor) splice junctions in the introns.

Mutations that affect either of the invariant dinucleotides in the splice junction completely abolish normal splicing and produce the phenotype of β^0-thalassemia. These mutations can be base substitutions that change one or the other of invariant dinucleotides or short deletions that remove them. Genes bearing these mutations appear to transcribe normally and, although some alternative splicing occurs using "cryptic" donor or acceptor sites, the misspliced mRNA do not translate into functional β- globin. The misspliced mRNA species can sometimes be detected in small amounts in affected erythroid cells. They are nonfunctional because translation of the abnormally spliced or frameshifted mRNAs would usually stop prematurely due to the introduction of chain termination (nonsense) codons (Huang & Benz, 2001).

Mutations of Splice Site Consensus Sequence

Flanking these invariant dinucleotides are sequences that are fairly well conserved and a consensus sequence can be recognized at the exon–intron boundaries. They encompass the last

three nucleotides of the exon and the first six nucleotides of the intron for the 5' donor site; and the last 10 nucleotides of the intron and the first nucleotide of the exon for the 3' acceptor site. Mutations within the consensus sequences at the splice junctions reduce the efficiency of normal splicing to varying degrees and produce a β-thalassemia phenotype that ranges from mild to severe. For example, mutations at IVS1 position 5, G→C, T, or A considerably reduce splicing at the mutated donor site compared with normal. The mutations appear to activate the use of three "cryptic" donor sites (sequences that mimic the consensus sequence for a splice site but are not normally used), two in exon 1 and one in IVS1, which are utilized in preference to the mutated donor site. On the other hand, the substitution of C for T in the adjacent nucleotide, IVS1 position 6, only mildly affects normal RNA splicing even though it activates the same three cryptic donor sites as seen in the IVS1–5 mutants. Although the IVS1–6 T-C mutation is generally associated with milder β-thalassemia, studies have shown that in some cases apparently identical mutations can be severe; this is presumably related to the chromosomal background on which the mutations have arisen (Camaschella et al., 1995; Ho et al., 1998).

Mutations that Create New Alternative Splice Sites in Introns

A third category of splicing mutation is due to base substitutions in introns that generate new splicing signals, which are preferentially used instead of the normal splice sites. The associated phenotype may be either β^+ or β^0-thalassemia, depending on the site and nature of the mutation. The splicing mutation at position 110 of IVS1 is one of the most common form of β-thalassemia in the Mediterranean population. The mutation is a substitution of G to A that creates an acceptor AG in a favorable consensus sequence environment, 19 bp 5' to the normal acceptor AG of IVS1. In vitro expression studies have shown that this newly created alternative splice site is preferentially used in 80%–90% of the transcripts, whereas the normal splice site is used in only 10%–20% of the transcripts thus giving the phenotype of β^+-thalassemia (Busslinger et al., 1981; Fukumaki et al., 1982).

Another β-thalassemic gene with a T→G substitution in position 116 of IVS1, leads to a newly created 3' acceptor site. In this case, the normal acceptor sequence, although intact, is not used, and little or no normal β mRNA is produced resulting in a β^0-thalassemia phenotype (Metherall et al., 1986).

Mutations that Create Alternative Splice Sites in Exons

Four mutations have been identified in exon 1 that are associated with activation of cryptic or alternative splice sites (Thein & Wood, 2009; Weatherall & Clegg, 2001; Thein, 1998). Three of these mutations modify the cryptic splice site spanning codons 24–27 in exon 1 so that it more closely resembles the consensus splice sequence AAGGTGAGT and activates it. The codon 24 GGT-GGA mutation is translationally silent (Hattori et al., 1988), whereas codon 26 GAG-AAG and codon 27 GCC-TCC result in the β^{E} and $\beta^{Knossos}$ variants, respectively (Orkin et al., 1982, Orkin et al., 1984). The mutation in codons 26 and 27 lead to a minor use of the alternative pathway so that there is a reasonable level of normally spliced products that result in the mild β^{+}-thalassemic phenotype of the β^{E} and $\beta^{Knossos}$ alleles, respectively. The reduction in normal splicing is the molecular basis for the mild β^{+}-thalassemic phenotype of these variants, including the β^{E} allele. The β^{E} allele is particularly prevalent in southeast Asia where it can reach up to a frequency of 75% in northeast Thailand. Its interaction with β-thalassemia accounts for a large proportion of the TM in southeast Asia (Weatherall & Clegg, 2001).

Mutations Causing Abnormal Posttranscriptional Modification

An increasing number of mutations in the 5'-UTR have been characterized since the original CAP +1 A→C allele (Wong et al., 1987). The defects include single base substitutions and minor deletions distributed along the stretch of 50 nucleotides. As in the −101 C→T mutation, heterozygotes for this class of mutations are silent; the extremely mild phenotype is exemplified in a homozygote for the +1 A→C mutation who has the hematological values of α-thalassemia carrier. It is not known whether the CAP mutation causes β thalassemia by decreasing β-globin gene transcription or by decreasing the efficiency of capping (posttranscriptional addition of $m^{7}G$) and mRNA translation. In vivo and in vitro studies show that the +33 C→G mutation leads to a reduction of β mRNA that is 33% of the output from a normal β gene, milder than the mutations involving the promoter elements. Compound heterozygotes for these transcriptional mutations and the more severe β thalassemia alleles tend to have a milder disease (Ho et al., 1996).

Mutations affecting the consensus AATAAA sequence at the 3'UTR include seven base substitutions at different locations; two short deletions of 2 and 5 bp each, and one deletion of the total AATAAA sequence. These mutations markedly decrease the efficiency of the cleavage-

polyadenylation process and only about 10% of the mRNA is properly modified. Therefore, the associated phenotype is that of β^+-thalassemia of moderate severity. The remainder of the transcripts extend far beyond the normal polyadenylation site and are probably cleaved and polyadenylated after the next AATAAA consensus sequences, which occur about 0.9–3 kb downstream (Orkin et al., 1985).

Mutations affecting other sites in the 3' UTR, a C→G substitution at nucleotide 6, and a 13 bp deletion at nucleotides 90 downstream from the termination codon, also result in β^+-thalassemia (Rund et al., 1992; Hamid & Akbari, 2011).

Mutants resulting in premature termination of translation

Approximately half the β-thalassemia alleles result from the introduction of premature termination codons, either because of direct mutations creating a stop codon or a change in the reading frame by insertion or deletion of a single to a few nucleotides. These frameshifts lead to premature termination further downstream when the next nonsense codon is reached (Thein & Wood, 2009). One of the first nonsense mutations to be characterized and extensively studied was the mutation at $\beta^0$39 CAG→TAG (Humphries et al., 1984; Takeshita et al., 1984; Huang & Benz, 2001). This mutation is the second most common cause of β-thalassemia in the Mediterranean population and accounts for most of the cases of β-thalassemia in Sardinia (95%) (Masala et al., 1988; Rosatelli et al., 1987; Rosatelli et al., 1992). In homozygous β^0(39)-thalassemia TM is the most frequent phenotype, although rare TI can be found (10%) (Galanello et al., 1989). HbF, when present, is the only useful Hb tetramer synthesized and this is mostly the principal cause of the intermediate phenotype.

Dominantly inherited β-thalassemia

The syndrome comprises a distinct set of structural mutations affecting the HBB gene that are associated with typical hematological features of β-thalassemia (i.e., increased HbA_2 levels and imbalanced α-/β-globin-chain biosynthesis in heterozygotes). In these variants, the mutated β–globins are so unstable that they undergo very rapid post-synthetic degradation. The hyperunstable β-chains act in a dominant negative way, causing a disease phenotype even when present in the heterozygous state, and hence have been referred to as "dominantly inherited β-thalassemia" (Steinberg et al., 2009). Unlike the common recessive forms of β-thalassemia, which are prevalent in malaria endemic regions, dominantly inherited β-thalassemia has been

described in dispersed geographical regions. The clinical manifestations are related both to functional deficiency of β globin and to precipitation of the β-chain variants with the concomitant excess α chains overload the proteolytic intracellular mechanisms increasing ineffective erythropoiesis. Indeed, the large intraerythroblastic inclusions, that are so characteristic of this form of β-thalassemia, have subsequently been shown to be composed of both α- and β-globin chains (Ho et al., 1997). In contrast, the inclusion bodies in homozygous β-thalassemia consisted only of precipitated α-globin. The molecular mechanisms underlying the instability include: substitution of the critical amino acids in the hydrophobic heme pocket displacing heme leading to aggregation of the globin variant; disruption of secondary structure because of replacement of critical amino acids; substitution or deletion of amino acids involved in αβ dimer formation; and elongation of subunits by a hydrophobic tail. The molecular defects include missense mutations, deletions or insertion of intact codons, nonsense mutations causing premature termination codons in exon 3. Frameshifts may also result in aberrant splicing producing elongated or truncated β-globin-chain variants with abnormal carboxy-terminal ends.

Deletional forms of β-thalassemia

The β-thalassemias are rarely caused by major gene deletions: a group includes deletions that are restricted to the β-globin gene and a second group contains larger deletions involving part or all of the β-globin gene cluster.

Deletions Restricted to the β-globin gene

These deletions, ranging from 105 bp to approximately 67 kb in size, involve only the β-globin gene and its flanking DNA without affecting any of the other neighboring β-like globin genes. The phenotype associated with these deletions is that of β^0-thalassemia.

These conditions are associated with unusually high levels of HbA_2 and variable increases of HbF in heterozygotes (Thein & Wood, 2009).

It has been proposed that deletion of the β promoter removes competition for the upstream β-LCR and limiting transcription factors, allowing greater interaction of the LCR with the cis δ and γ genes, thus enhancing their expression. Although the increases in HbF are variable and modest in heterozygotes for such deletions, they can be sufficiently increased to partially compensate for the complete absence of β globin in homozygotes; two individuals homozygous for different

deletions in this group showed a mild disease not transfusion dependent (Schokker et al., 1966; Gilman, 1987; Craig et al., 1992).

εγδβ-thalassemia

Clinically, the εγδβ-thalassemias are characterized in newborns by anemia and hemolysis, which is self-limited, and in adults by the hematological phenotype of β thalassemia trait with normal levels of HbA_2 and HbF (Weatherall & Clegg, 2001). The severity of anemia and hemolysis may be variable (even within a family) and in some cases, blood transfusions are necessary during the neonatal period. Only heterozygotes have been identified; homozygotes, presumably, would not survive early gestation. At the molecular level, the deletions fall into two categories: group I removes all or a greater part of the β-globin complex, including the β-globin gene; group II removes extensive upstream regions leaving the β-globin gene itself intact, despite which, its expression is silenced because of absence of the upstream β-LCR (Thein & Wood, 2009; Rooks et al., 2012). The associated phenotypes of the two groups are similar.

HPFH syndromes and δβ-thalassemia

HPFH and δβ-thalassemia are descriptive terms used for a range of disorders that are characterized by decreased or absent β-globin production and a variable compensatory increase in γ-chain synthesis. In both the $^{G}\gamma^{A}\gamma(\delta\beta)^{0}$-HPFHs (or simply deletional HPFH, dHPFH) and the $^{G}\gamma^{A}\gamma(\delta\beta)^{0}$-thalassemias (shortly δβ-thalassemias) the deletions remove both the δ and β genes and extend a variable degree 3' to the cluster, in some cases for up to 100 kb. The $^{G}\gamma(^{A}\gamma\delta\beta)^{0}$-thalassemias differ in that the 5' end of the deletion partially or totally removes the $^{A}\gamma$ gene in addition to the δ and β genes. The broad classification of the deletion conditions into these two groups, although useful, is rather arbitrary: in fact there is a continuum between δβ-thalassemia and the dHPFHs and that not all conditions fit neatly into one of these groups. dHPFH and δβ-thalassemia were originally distinguished on hematological and clinical grounds. Heterozygous δβ-thalassemia had a similar red cell picture to β-thalassemia, with hypochromic and microcytic erythrocytes, but a normal level of HbA_2 (<3.0%). In addition, there was a raised level of HbF (5%-15%) that had a heterogeneous intercellular distribution. Homozygotes or compound heterozygotes with β-thalassemia had a clinical picture of TI or TM. In contrast, dHPFH heterozygotes, had essentially normal red cell indices, a normal level of HbA_2 and even higher levels of HbF (15%-30%) with a more homogeneous, pancellular distribution. HPFH

homozygotes were clinically normal, albeit with reduced MCV and MCH, whereas compound heterozygotes with β-thalassemia were clinically very mild. As more and more cases and different molecular types of the two conditions were described, these differences became indistinct and it is now clear that there is considerable overlap in many of the parameters that were initially used to differentiate them (Steinberg et al., 2009).

4.4. The α-thalassemias

In contrast to the β-thalassemias, which are usually caused by point mutations of the β-globin gene, the α-thalassemia syndromes are usually caused by the deletion of one or more α-globin genes. Non-deletional forms of α-thalassemia have also been characterized but are relatively uncommon. More than 100 genetic forms of α-thalassemia have thus far been identified, with phenotypes ranging from asymptomatic to lethal. Despite this complexity, the severity of this disorder is usually well correlated with the number of non-functional copies of the α-globin genes. On the basis of the numbers of α-globin genes lost by deletion or totally or partially inactivated by point mutations, the α-thalassemias are classified into two main subgroups: α^+-thalassemia (formerly called α-thalassemia 2), in which one of the genes is deleted or inactivated by a point mutation (-α/αα or $\alpha\alpha^{ND}$/αα, with ND denoting nondeletion), and α^0-thalassemia (formerly called α-thalassemia 1), in which both the α-globin genes on the same chromosome are deleted (--/αα).

In the fetus, defective production of α-chains is reflected by the presence of excess γ-chains, which form γ4 tetramers, called Hb Bart's; in adults, excess β-chains form β4 tetramers, called HbH. Because of their very high oxygen affinity, both tetramers cannot transport oxygen, and, in the case of HbH, its instability leads to the production of inclusion bodies in the red cells and a variable degree of hemolytic anemia.

The homozygous state of α^+-thalassemia and the heterozygous state of α^0-thalassemia (grouped under the term "α-thalassemia minor") are associated with a substantial reduction in the MCV and MCH. In α^+-thalassemia heterozygotes, the MCV and MCH are usually reduced, but there is a small overlap with normal values. In contrast to β–thalassemia minor, HbA_2 levels do not raise, so its identification is based mainly on hematological parameters. Milder forms of α-thalassemia are often misdiagnosed as iron deficiency, although the exact frequency of misdiagnosis is

unknown. Clinically relevant forms of α-thalassemia usually involve α^0-thalassemia, either coinherited with α^+-thalassemia (-α/-- or $\alpha\alpha^{ND}$/--) and resulting in HbH disease or inherited from both parents and resulting in Hb Bart's hydrops fetalis (--/--), which is lethal in utero or soon after birth.

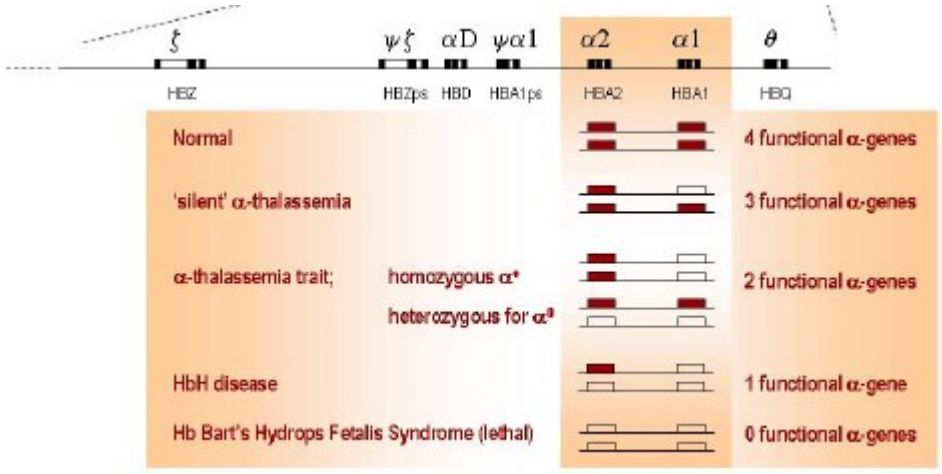


Figure 10. *Classification of α-thalassemia defects.*

HbH disease is often considered to be a relatively mild disorder. Studies have nevertheless highlighted clinically severe phenotypes, notably in non-deletional variants of the disease (Fucharoen & Viprakasit, 2009; Lal et al., 2011). HbH disease is characterized by a wide range of phenotypic characteristics. The form that results from deletions (-α/--) usually follows a relatively mild course, with moderate anemia and splenomegaly. Aside from episodes of intercurrent infection, this form of HbH disease does not require blood transfusions. However, the variety that results from the interactions of a non-deletional α-globin gene mutation together with α^0-thalassemia ($\alpha\alpha^{ND}$/--) follows a much more severe course. This is particularly true when the non-deletional mutation is the α-globin chain termination mutant HbCS (par. 4.1), which is very common in many Asian countries.

Embryos with Hb Bart's hydrops fetalis succumb to severe hypoxia either early in gestation (e.g., in the case of --FIL/--FIL) or during the third trimester (e.g., in the case of --SEA /--SEA). The acronyms FIL and SEA refer to two different deletions that cause α^0-thalassemia and that are

prevalent among Filipinos and persons from Southeast Asia, respectively (Weatherall & Clegg, 2001).

A few children who received an intrauterine transfusion or a transfusion immediately after delivery have survived to 5 years of age. These children require regular transfusions and, when appropriate, iron chelation therapy; they usually have serious clinical complications, congenital anomalies, and delays in cognitive and motor functions (Chui & Waye, 1998).

The hydrops fetalis syndrome is often accompanied by a variety of congenital malformations and maternal complications, including severe anemia of pregnancy, preeclampsia, polyhydramnios, and extreme difficulty in delivery of both the fetus and the hugely enlarged placenta (Weatherall & Clegg, 2001).

Milder variants of α-thalassemia act as genetic modifiers of other inherited conditions, as illustrated by epistatic interactions between α-thalassemia and β-thalassemia (Thein, 2005) or between α-thalassemia and HbS (Williams et al., 2005). Conversely, triplications and quadruplications of the α-globin gene, frequently observed in many populations, can interact with β-thalassemia variants to produce more severe phenotypes. (Ma et al., 2001; Giordano et al., 2009)

4.4.1. Molecular basis of α-thalassemia

All reported α^0-thal deletions remove both linked α-globin genes interstitially along with different lengths of the adjacent chromosomal contents, from subtelomere to the region beyond the α1 gene. Only a few individuals and families have been found to have the phenotype of α^0-thal due to deletions of 40 to 33 kb upstream DNA sequences (including the ζ-globin gene) and minimal consensus sequences (MCS). MCS are highly conserved throughout evolution and are denoted as hypersensitive site HS-40 and -33. This minimal region of 13 kb is thought to contain a remote regulatory element encompassing most, but not all, cognate binding sites for erythroid-specific transcription factors critical for full α-globin gene expression, and removal of this DNA region has been shown to cause a down-regulation of the downstream α-globin genes (Viprakasit et al., 2006).

Therefore, the molecular basis of α^0-thal can be classified into two categories: (1) interstitial deletional α^0-thal caused by deletions that remove both linked α-globin genes and (2) upstream-deletional α^0-thal with both α-globin genes intact.

Since deletional α^0-thal is highly prevalent in the Far East, particularly in the southern part of China, Thailand, the Malaysian peninsula and even in the remote islands of the Philippines with carrier rates of 15%, 2.2% to 9%, 4.5% and 5%, respectively, homozygosity of these deleted alleles is not uncommon (--/--) (Wasi et al., 1974; Chan et al., 1988; Fucharoen et al., 1988; Galanello et al., 1992).

In Southeast Asia and the southern part of China, α^0-thal mostly occurs from a deletion of about 19.3 kb of the DNA, removing both linked α-globin genes but leaving the ζ gene intact (--SEA). Another form of α^0-thal found in this region is the THAI deletion (--THAI), which removes a larger DNA segment (33.5 kb) including the embryonic ζ-globin genes with a significant lower frequency (SEA:THAI = 99:1). Although the different forms of α-thalassemia alleles have a worldwide distribution, the occurrence of Hb Bart's hydrops fetalis is almost solely confined to this region where the highest incidence of α^0-thal exists (Higgs, 2009).

Two common types of deletional α^+-thal have been identified, one involving a deletion of 4.2 kb of DNA (leftward type, -α4.2) and another of 3.7 kb (rightward type, -α3.7). The -α4.2 deletion is present at low frequency in Thailand and Southeast Asian countries, but it is more frequent in Papua New Guinea and Vanuatu in Melanesia. The high frequency and worldwide distribution of both deleted alleles suggested that these mutations are quite ancient in the modern human history, and they have aged enough to acquire additional nucleotide mutations causing several α-globin chain variants occurring on the chromosomes with the -α3.7 or -α4.2 type deletions. Of these, Hb Q-Mahidol (Thailand) in Thai and Chinese, Hb G-Philadelphia in the African-American population and Hb J Tonkarigi in the Pacific Islands were found at a higher frequency than other α-globin variants linked with the α-thal 2 deletion.

The molecular cause of deletional α-thalassemia is explained by DNA sequence analysis: the α–globin genes are embedded within two highly homologous, 4–kb duplication units whose sequence identity appears to have been maintained throughout evolution by gene conversion and unequal crossover events. These regions are divided into homologous subsegments (X, Y, and Z) by non–homologous elements (I, II, and III). Reciprocal recombination between Z segments, which are 3.7 kb apart, produces chromosomes with the $\alpha^{-3.7}$ deletion and chromosomes with three α–genes ($\alpha\alpha\alpha^{anti\ 3.7}$). Recombination between homologous X boxes, which are 4.2 kb apart, also gives rise to the $\alpha^{-4.2}$ gene and a $\alpha\alpha\alpha^{anti\ 4.2}$ chromosome (Figure 11). Further recombination

events between the resulting chromosomes may give rise to quadruplicated and quintuplicated α–genes ($\alpha\alpha\alpha\alpha^{anti\ 3.7}$, $\alpha\alpha\alpha\alpha^{anti\ 4.2}$).

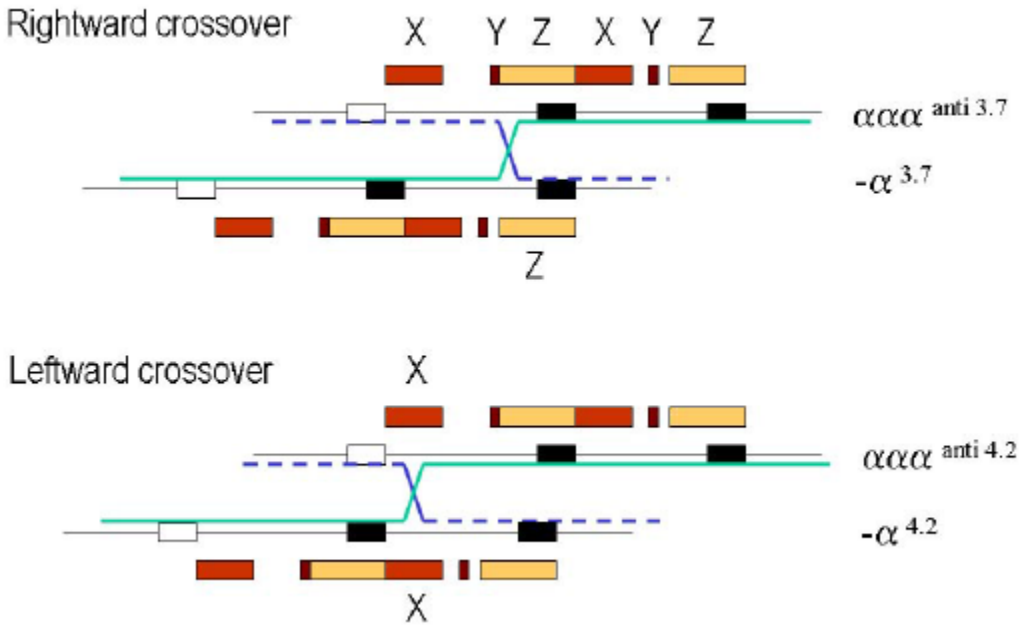


Figure 11. *Deletions that cause α+-thalassaemia. The homologous duplication units X, Y and Z in which the α-genes are embedded are indicated as colored boxes. A cross-over between the mis-paired Z boxes during meiosis gives rise to the -α3.7 and αααanti 3.7 chromosomes. Cross-over between misaligned X-boxes give rise to -α4.2 and αααanti 4.2.*

Nearly 70 different types of non-deletional α^{+}-thalassemia have been reported, most of which are extremely rare and have been reported only in sporadic families. The majority of these mutations have been found to involve the α2 genes located in closer proximity to the regulatory region (HS-40 and HS-33), resulting in a higher expressed output than its downstream counterpart (α1 gene) with a ratio around 3:1. More preferential mutations occurring on α2 genes might not imply that this leftward-most allele is more susceptible to mutagenesis since both α genes are in close proximity, sharing a similar chromosomal context, but it is plausible that due to a higher contribution of the α2 gene on globin synthesis, mutated alleles could have a larger effect on phenotypic expression as affected individuals will present with more severe hematological phenotypes. Although most non-deletional α+-thalassemias are rare, several mutations have been

found to be prevalent within specific populations (Galanello et al., 1992; Ma et al., 2001b; Viprakasit et al., 2002).

Non-deletional mutations, in general, affect the fundamental processes of globin gene expression, from mRNA transcription, splicing and protein translation through creating novel truncated or elongated globin peptides. For example, polyadenylation (poly A) signal mutations (due to different base substitutions or base deletions) in the α2-globin gene are found at high frequencies in Greece, Saudi Arabia and Turkey. Other non-deletional α-thal 2 mutations found at low frequencies in various populations include Hb-CS, Hb Koya Dora and Hb Paksé, which are caused by mutations in the termination codon of the α2-globin gene that result in an elongated polypeptide.

5. Distribution of Hbpathies in Southeast Asia and in Vietnam

Southeast Asia consists of 10 countries with a total population of about 400 million. The ethnic origins of people living in these countries are very heterogeneous. The Mon-Khmer and Tai language-speaking people occupy Thailand, Laos, Cambodia and some parts of Vietnam, Myanmar and Malaysia. The west includes the Burmese (Tibeto-Burman) and the Northeast is the Vietnamese (Austro-Asiatic). The Malayopolynesians (Austronesian) live in Malaysia, Indonesia, Brunei, the Philippines and a number of Pacific island nations. Chinese and Indians are relatively newcomers spread throughout the region.

In Southeast Asia α-thal, β-thal, HbE and HbCS are prevalent. The most common Hbpathy is HbE which is unique to this region. The highest frequency of this variant gene is found in the "HbE triangle" at the borders of Cambodia, Laos and Thailand. In Northeast Thailand, HbE carriership reaches 54%. The severe forms of β-thal are almost equally divided between the β-thal homozygous state and the compound heterozygousity for HbE\β-thal, which reflects the extremely high frequency of these disorders throughout Southeast Asia. The gene frequencies of α-thal reach 30-40% in Northern Thailand and Laos, 4.5% in Malaysia and 5% in the remote island of the Philippines. The frequence of β-thal varies between 4 and 10% reaching 5-8% in Northern Thailand, 3-9% in Laos, 6-10% in Indonesia and 4% in Myanmar. HbCS gene frequencies vary between 1 and 8%.

These abnormal genes in different combinations lead to over 60 different thalassemia syndromes, making Southeast Asia the locality with the most complex thalassemia genotypes. The four

major thalassemic diseases are Hb Bart's hydrops fetalis, homozygous β-thalassemia, β-thalassaemia/Hb E and Hb H diseases (Filon, 2000; Fucharoen & Winichagoon, 2011).

Vietnam is situated along the east coast of the Southeast Asian peninsula, between China in the north, Laos in the northwest and Cambodia in the southeast. Three-quarters of the population of 75 million are of the Kinh ethnic group, and the remainder is divided into 55 different ethnic minority groups (Filon et al., 2000; Blackwell, 1965; Fucharoen & Winichagoon, 1987).

In Vietnam, both β-thalassemia and HbE are prevalent and represent one of the most common forms of hemolytic anemia. The carrier rate for β-thalassemia varies between 1.5% and 25%. The higher incidence is in the ethnic minority groups, especially Tay (11%) and Moung (25%) in the north and Pako (8.33%) and Cotu (14%) in the central region of the country. The prevalence of HbE in Kinh groups is not much different throughout the country, 1–9%. The highest frequencies of HbE were found in the minority groups, Ede (41.7%), Pako (6%), and Vaˆn Kieˆv (23%) who live in the central plateau.

The incidence of β-thalassemia and HbE was less studied in the South of Vietnam. The frequency of β-thalassemia studied in the Kinh people and the Vietnamese refugees in the United States was between 1% and 2%, and HbE in Saigon was about 3.2% (Svasti et al., 2002).

The first screening for β-thalassemia in Vietnam was performed in the North of Vietnam, situated between China in the north and Laos in the northwest (Filon, 2000). Later, the spectrum of β-globin mutations has been investigated in Ho Chi Minh City, South Vietnam (Svasti et al., 2002).

Different β-thalassemia mutations have been so far identified: the most commons, both in the South than in the North of Vietnam, are the frameshift mutation at codons 41/42 (−TCTT), the nonsense mutation at codon 17 (A→T) and the frameshift mutation at codon 95 (+A), which is known as the "Vietnamese" mutation. Other mutations, although less frequent, are present in these countries: the -28 (A→G), the IVS-I-1 (G→T), the frameshift mutation at codons 71/72 (+A) and the IVS-II-654 (C→T) (Svasti et al., 2002).

The prevalence of α-thalassemia it was at first analyzed by O'Riordan et al. (2010) in patients belonging to different ethnic groups (Kinh, Nung, Tay, S'Tieng) from southern and southern central Vietnam. The highest frequencies of α-thalassemia were found among the S'Tieng, in whom the deletions -α3.7 and --SEA were present at frequencies of 0.22 and 0.03, respectively.

The -α3.7 and --SEA deletions were also found in all ethnic groups studied, but at lower frequencies. Only one individual, of Tay background, was found to carry the -α4.2 deletion.

The HbCS variant was also found and showed the higher gene frequency (0.03) in the S'Tieng minority (O'Riordan et al., 2010).

More recent data derived from a community-based survey conducted in the Nam Dong mountainous district, located 60 km westwards of the provincial capital of Hue city. A total of 298 unrelated Cotu ethnic women were randomly recruited. HbCS resulted markedly high with a gene frequency of 0.143. This is the highest frequency of HbCS ever reported in world populations (Nguyen et al., 2014).

AIM OF THE STUDY

In many Asian countries the α- and β-thalassemias and their interaction with several structural Hb variants are producing an increasingly severe health burden. It is therefore important to accurately identify carriers of these disorders and offer the option of preventive measures by prenatal diagnosis to couples at risk of having a child with severe disease. Rapid genotype characterization is fundamental in the diagnostic laboratory and the ability to perform DNA analysis has become an increasingly important requirement. The acquisition of such skills requires the development of suitable training programs involving hematologists, pediatricians, biologists and technical staff. A start has been made in developing and low-income countries and there is already considerable evidence that much more can still be done, including the further development of partnerships between countries where expertise in this field has been developed and countries where no such expertise exists.

This PhD project is part of a longstanding cooperation between Huè College of Medicine and Pharmacy and the University of Sassari which has started in 2001. The purpose of the collaboration was and still is to create positive synergies between the different stakeholders, manage scientific and cooperation interventions and coordinate training and research.

In Vietnam, most of researches on Hbpathies were mainly conducted in the North and the South regions. Recently, Nguyen et al. (2013) performed a preliminary study restricted to a limited number of pregnant women in Thua Thien Hue Province, the central area of Vietnam. Epidemiologic researches of Hbpathies are, in general, not systematic and molecular studies concerning Vietnam are incomplete.

Many of the patients residing in rural areas receive only limited care because of the lack of adequate medical facilities. Prevention is the only way to control the increase in the number of patients. To ensure feasibility of prenatal diagnosis it would be useful to apply simple and inexpensive molecular assay for the detection of mutations.

The aim of the PhD training has been to develop professional skills and competences directed at the avoidance and better management of the Hbpathies in central Vietnam and to provide the groundwork for screening programs.

The main objectives have been as follows:

- acquire professional skills aimed to determine the Hb profile in newborn and adult, to characterize quali-quantitative Hbpthies, to carry out a correct diagnosis of rare Hb variants or complicated associated forms,
- develop and spread the acquired techniques towards Vietnamese medical and biology students for further evolution of education and screening programs;
- determine the spectrum of β-thalassemia mutations in patients and their relatives from Hue province of Vietnam.

MATERIALS AND METHODS

6. Samples collection

Study was done on 353 individuals: a study group of 160 subjects referred to Hue Medicine and Pharmacy College and PhuVang District Hospitals for hematological and clinical evaluation and a control group of 193 individuals, who displayed normal hematology and hemoglobin profile.

Information including age, gender, anamnesis and clinical examination was collected.

In our study group, 89/160 samples (55,6%) showed abnormal hematological parameters and have undergone to Hbpathies screening. Twenty seven of them, which showed qualitative or quantitative alteration in Hb profiles, were included in the DNA analysis.

The discovery of a novel β-thalassemic mutation has been supported by a family study performed on additional five samples.

7. Samples preparation

7.1.Whole blood preparation

Blood samples were collected in EDTA tubes. They should be 50-60% full and not overfill. Specimens were mixed gently by inverting 5-10 times and placed on a rocker for up to 30 minutes, then refrigerate at 2-8°C. Clotted or hemolyzed specimens are unacceptable.

7.2.Preparation of hemolysates

The blood aliquot was washed through centrifugation at 1300 g for 5' in physiological solution (NaCl 0.9%) until the supernatant appeared clean and clear. The addition of one volume of ipotonic solution (tridistilled H_2O, mQ) causes the hemolysis and the consequent suspension of Hb (hemolysate), directly used for electrophoretic analysis. An additional step was required to remove cell membranes: half of the volume of CCl_4 was added to the solution. After centrifugation (20000 g for 2'), the supernatant was recollected and contained Hb in water could be readily used for HPLC analysis or stored at -20°C for a long time. Hb solutions were quantified by spectrophotometric analysis at 576 nm.

8. Hematology and Hemoglobin analysis

Hematological parameters were measured by the Blood Analyzer SYSMEX KX-21 and SYSMEX 800i (Japan Care, CO., LTD).

Hb tetramers separation was carried out by either cation-exchange high-performance liquid-chromatography (CE-HPLC), with the Chromsystems Instrument and Chemicals (GmbH, Germany) and by isoelectric focusing (IEF).

8.1. CE-HPLC analysis

The CE-HPLC analysis was performed by the supports of staffs at Genetic Department of Hue University of Medicine and Pharmacy.

The total analysis is carried out on a binary HPLC gradient system by means of a UV/VIS detector. It has been designed to separate and determine in 5–9 min the area percentages for HbA_2 and HbF and to provide qualitative and quantitative determinations of normal and abnormal Hbs.

8.2. IEF analysis

IEF was performed on a polyacrylamide slab gel (%T:%C = 5.2%:3%) where a pH gradient is formed through ampholytes; every protein was separated according to its isoelectric points (pI), it will move to the position on the gel at which the pH is equal to its pI.

The gel was prepared in a sandwich of treated glass: the mold was brushed with a "Repel solution" (5% dimethyldichlorosilane, in chloroform); the support instead was treated with a "Bind solution" (0.2% Silane A174 in acetone).

The mixture for the gel was prepared by adding ampholytes at pH 6.7 to 7.7 (6.2%), ampholytes at pH 3–10 (0.5%), TEMED (0.15%) and ammonium persulfate (APS, 0.56 mg/ml), and it was casted into the prepared sandwich for 1 h at room temperature.

A pre–run was necessary to form the pH gradient: strips of absorbent paper were wetted with the specific electrolyte (40 mM glutamic acid for anode, 0.1 M NaOH for cathode) and placed onto the gel. A voltage of 400 V at 4°C was applied to the gel until the amperage went below 4 mA.

A total of 20 μg of Hb were loaded and run for 2 h at 1600 V at 4°C.

At the end of the run, the Hbs were fixed in Tricloracetic Acid (20% TCA) for 10'; the gel was then stained for 30' with a BBF staining solution (0.1% BBF, 50% EtOH, 5% Acetic Acid).

Destaining was performed in 30% Ethanol and 6% Acetic Acid (Masala and Manca, 1991).

Yellow stained bands were turned into blue by soaking the gel into water: this allowed a semi–quantitative densitometric analysis at 600 nm.

9. Molecular analysis

Through the information obtained from hematology and hemoglobin profile, direct DNA sequencing of specific PCR products was performed in order to identify the abnormality involved, which may consist in single point mutations or microinsertions/deletions.

The search for large deletions (e.g. δβ-thalassemias or deletional α-thalassemias) has been allowed by the multiplex ligation–dependent probe amplification (MLPA).

9.1. DNA extraction

DNA extraction was performed either with commercial kits (Invisorb® Spin Blood Midi Kit, STRATEC Biomedical AG) or Salting out method (Miller *et al.*, 1988). The salting out protocol was the following:

add cold Lysis Buffer (0.32 M Sucrose, 10 mM Tris/Cl pH 7.5, 5 mM $MgCl_2$, 1% Triton X–100) to 2 ml of whole blood in EDTA, up to a volume of 10 ml. Gently shake. Incubate at 4°C for 15'–30' to lyse the red blood cells.
Spin at 1500 g for 20'. A pellet of white cells should form at the bottom of the tube. Remove the supernatant.

Resuspend the pellet in Physio Buffer (75 mM NaCl, 25 mM EDTA) up to a volume of 10 ml. Spin at 1500 g for 10'. Discard the supernatant. Repeat 2 more times to get a cleaner/whiter/purer pellet.

Lyse the white blood cells in 3 ml of WBC lysis solution (10 mM Tris/Cl pH 8.0, 2 mM EDTA). Add 100 µl of 10% SDS and 400–600 µl of proteinase K (10 mg/ml). Pipette up and down a few times. Incubate 37°C overnight.

Add to the solution 1 ml of NaCl 6 M. Shake vigorously for 15''. Spin at 1500 g for 15'. Transfer the supernatant into a clean tube. Add 1 volume of isopropanol 100%. and gently shake to precipitate the DNA. Spin at 2200 g for 15' and discard the supernatant. Add 1 ml of ethanol 70%. Spin at 2200 g for 15' and discard the supernatant. Air-dry the pellet until all the ethanol has evaporated off.
Dissolve the pellet in an appropriate volume of TE 1X pH 8.

Extracted DNA was quantified with NanoDrop 8000 Spectrophotometer (Thermo Scientific).

9.2. Multiplex Ligation-Dependent Probe Amplification (MLPA) analysis

MLPA analysis is a recent technique developed by the MRC-Holland company for the relative quantitation of up to 40 to 45 nucleic acid targets. The uses of MLPA include detection of aneuploidies, common microdeletion syndromes and subtelomeric copy-number changes, identification of marker chromosomes, and detection of familial copy-number changes in single genes.

The entire assay is performed in a single tube, can be completed within 24 to 48 hours and consists of following steps: hybridization of the probes to the complementary target sequences, ligation of the bound probes, PCR amplification of the ligated probes, followed by capillary electrophoresis and analysis.

The results are analyzed by normalization to internal control fragments followed by normalization to the control samples.

Peak ratios between 0.7 and 1.3 are considered normal. Ratios above 1.3 indicate the presence of a gain of the target sequence while ratios below 0.7 indicate loss of a target. Theoretical ratios for a heterozygous deletion and duplications are 0.5 and 1.5, respectively, although the actual values may range 0.3 to 0.7 for heterozygous deletions and 1.3 to 1.7 for heterozygous duplications.

MLPA P102 HBB or P140 HBA probemix (MRC-Holland) was used for copy number quantification on the β- and α-cluster, respectively.

MLPA reactions were performed according to the provided protocol, by using 80 ng of DNA and a 20 h hybridization step.

Ligation and amplification were carried out on an GeneAmp® PCR System 2700 thermal cycler (Applied Biosystem, Foster City, CA, USA).

MLPA products were separated by ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), quantified with the Coffalyser software (MRC-Holland), and compared with a DNA pool from normal subjects.

9.3.PCR and Sequencing Analysis of the β-globin gene

The complete β gene was amplified from genomic DNA with 3 different PCRs, each with its own Chemical and Thermal protocol.

P1 (GCCAAGGACAGGTACGGCTGTCATC) –**P2** (CCCTTCCTATGACATGAACTTAACCAT) primers amplify a 706 bp region (–140 to +566), encompassing a region from the proximal promoter to the beginning of IVS2.

Chemical Protocol: 1X Buffer, 3.5 mM $MgCl_2$, 250 μM dNTPs, 0.3 μM of each primer, 0.3 μg of genomic DNA, 2.5 U of polymerase.

Thermal file: 3' at 95°C; 35 cycles of 1' at 95°C, 45'' at 65°C and 1' at 72°C; additional final extension was added (4' at 74°C).

β7 (TCCTGATGCTGTTATGGGCAA) –**β8** (AAAAGCAGAATGGTAGCTGGA) primers amplify a 923 bp region (+333 to +1255), including part of exon 2 and IVS2. This fragment contains also the *Ava* II polymorphism for Orkin haplotypes.

Chemical Protocol: 1X Buffer, 3 mM $MgCl_2$, 250 μM dNTPs, 0.3 μM of each primer, 0.3 μg of genomic DNA, 2.5 U of polymerase.

Thermal file: 3' at 95°C; 35 cycles of 1' at 95°C, 1' at 60°C and 1' at 72°C; additional final extension was added (4' at 74°C).

β9 (AAAAACTTTACACAGTCTGCC) –**β10** (ATTAGCTGTTTGCAGCCTCA) primers amplify a 956 bp region (+799 to +1764), from IVS2 to 3'UTR of β–gene.

Chemical Protocol: 1X Buffer, 1.5 mM $MgCl_2$, 250 μM dNTPs, 0.3 μM of each primer, 0.3 μg of genomic DNA, 2.5 U of polymerase.

Thermal file: 3' at 95°C; 35 cycles of 1' at 95°C, 1' at 55°C and 1' at 72°C; additional final extension was added (4' at 74°C).

The HS2 region of the β-LCR was amplified using the primers **HS2-BamHI** (*ggatcc*TAAGCTTCAGTTTTTCCTTAGT) and **HS2-SalI** (*gtcgac*TAGATCTGACCCCGTATGT
GAGCAT), engineered to contain *Bam*HI and *Sal*I restriction site, respectively.

Chemical Protocol: 1X Pfu Buffer, 2 mM $MgCl_2$, 200 μM dNTPs, 0.3 μM of each primer, 0.3 μg of genomic DNA, 3 U of Pfu.

Thermal file: 3' at 95°C; 35 cycles of 1' at 95°C, 1' at 64°C and 2' at 72°C; additional final extension was added (5' at 74°C).

All amplified products were electrophoresed through a 1-1.2% agarose, 1X TAE, etidium bromurated gel at 7.5 volts/cm for 45' in the presence of a molecular weight marker.

DNA was recovered from agarose by means of the Montage Gel Extraction Kit (Merck Millipore). The purified fragments were sequenced by terminator chemistry (BigDye Terminator v3.1 Cycle Sequencing Kit, Applied Biosystem). Reaction mix was purified through the Sigma Spin Post-Reaction Clean-Up Columns (Sigma-Aldrich) and subjected to capillary electrophoresis on an ABI PRISM 3130 Genetic Analyzer (Applied Biosystem). Another Plasmid Editor APE (http://biologylabs.utah.edu/jorgensen/wayned/ape/) was used to align the obtained sequences with the reference (AC #: U01317).

9.4. Plasmid constructions and mutagenesis

The β-globin gene promoter was first cloned in pBluescript II SK (shortly pSK) to perform mutagenesis and then transferred in pGL2 plasmid for luciferase assays. Both plasmids confer ampicillin resistance to the transformed bacteria.

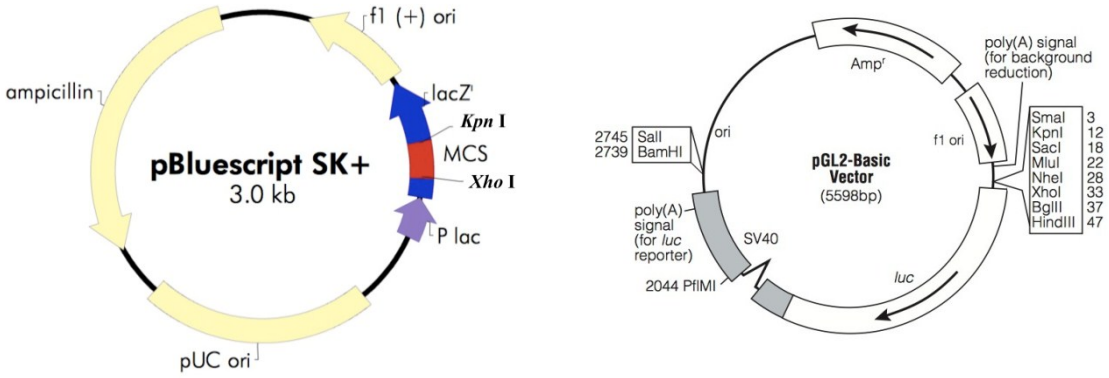


Figure 12: Plasmids used for mutagenesis (pBluescript; pSK) and for lucifearse assay (pGL2 Basic).

In order to clone the fragment containing the β-globin gene promoter in the pSK vector, β_KpnI (*ggtacc*ATCCAGTTTCTTTTGGTTAACCT) and β_XhoI (*ctcgag*TCTGTTTGAGGTTGCTAGTGAACAC) primers, engineered to contain both *Kpn*I and *Xho*I restriction site, respectively, were used.

The PCR reaction mix consisted in: 1X Pfu Buffer, 300 ng of DNA, 300 µM of dNTPs, 0.3 µM of each primer, 3 U of Pfu in a total reaction volume of 50 µl. Thermocycle parameters were:

3’ at 94°C; 35 cycles of 1’ at 94°C, 1’ at 66°C and 4' at 72°C; additional final extension was added (10’ at 74°C).

Two digestions were necessary to prepare the PCR fragment and the plasmid for the ligase reaction. The entire quantity of purified PCR product and 3 µg of plasmid were digested with 3 U of the restriction enzyme (*Kpn* I) for 1 h at 37°C, then purified with phenol:chloroform precipitation and digested again with the second enzyme (*Xho* I).

The phenol:chloroform precipitation protocol was the following:

add mQ water up to 200 µl to the solution containing DNA to be precipitated. Add 200 µl of phenol:chloroform:isoamyl alcohol (25:24:1). Vortex for a few seconds to obtain a homogeneous emulsion. Spin at 20000 g for 5’. Carefully remove the upper aqueous phase and transfer it to a fresh tube. Add the following reagents to the aqueous phase: CH_3COONa 3M, pH 5.5 (0,1 µl x volume of sample), 100% EtOH (2–2.5 volumes), glycogen 10 mg/ml (1 µl). Incubate at -80° C for 30’ or –20°C overnight to precipitate the DNA. Spin at 20000 g for 10’. Carefully remove the supernatant without disturbing the DNA pellet. Add 500 µl of 70% EtOH. Resuspend the pellet. Spin at 20000 g for 5’. Remove as much of the remaining ethanol as possible. Air-dry the pellet and resuspend it in an appropriate volume of sterile mQ water.

The restriction enzyme with the most alkaline pH (*Xho* I) was chosen for the second digestion in order to have a pH at which the alkaline phosphatase works.

The double digested plasmid was incubated 1 h at 37°C with Calf Intestinal Alkaline Phosphatase (CIAP or CIP) simply adding 1 unit of CIAP directly to the second digestion tube. CIAP removes phosphate groups from the 5’ end of the vector so that it cannot ligate, preventing the recircularization of the linearized vector and improving the possibility to transform only with the vector that contains the appropriate insert.

Both double digested insert and dephosphorylated vector were electrophoresed for 30’ in a 1% agarose TAE gel. Bands were excided with sterile scalpel and DNA was purified with Wizard SV Gel and PCR Clean–Up System (Promega).

Purified DNA was quantified by means of NanoDrop and a ligation was set in a 10 µl volume; 25 ng of vector were ligated to the insert (vector:insert 1:3) with 3 units of T4 DNA Ligase by incubating at 16°C overnight. Also a negative ligase control was set, at the same

conditions described above except for the absence of insert. Competent *E. coli* DH5α were transformed, plated on a LB Agar (Sigma) in presence of ampicillin and incubated overnight at 37°C. Also a negative ligase control was used to transform DH5α. It produced colonies only if the recircularized vector was present.
Transformation protocol for DH5α was:

Thaw competent DH5a cells on ice. Gently mix cells with the pipet tip and aliquot 100 µl of cells for each transformation into 1.5 ml tubes that have been pre-chilled on ice. Add ½ volume of the ligase reaction. Incubate on ice for 30'. Heat shock at 37°C for 3'. Incubate for 10' at room temperature. Add 500 µl of LB broth (without antibiotic) using sterile procedures. Incubate at 37°C for 50'. Spin at 1900 g for 5'. Remove 500 µl of supernatant. Resuspend the pellet with the remaining volume of LB. Spread the cells onto LB plates with appropriate antibiotic. Allow plates to dry and incubate inverted at 37°C overnight.

Grown monoclonal colonies were inoculated in 2 ml of LB Broth in presence of ampicillin and incubated overnight at 37°C, in agitation. 1 ml of the mini–inoculum was used to recover circular plasmid DNA with an alkaline lysis combined with SDS technique:

*Transfer 1 ml of mini–inoculation in a 2 ml tube. Spin at 20900 g for 1' and remove the supernatant. Resuspend the pellet in 200 µl of Solution I (*50 mM glucose, EDTA10 mM Tris/Cl pH 8 25 mM. *Add 200 µl of Solution II (*200 mM NaOH, 1% SDS*). Mix gently by inversion. Incubate exactly 5' at room temperature. Add 200 µl of cold Solution III(*7.5 M NH_4–acetate). *Incubate on ice 10'. Spin at 20900 g for 10'. Recover the supernatant. Add 2 volumes of 100% EtOH and incubate 10' at room temperature. Centrifuge at 20900 g for 15'. Resuspend the pellet with 500 µl of 80% EtOH. Centrifuge at 20900 g for 5'. Resuspend in 50 µl of TE containing RNase A (20 mg/ml. Incubate at 37°C for 10'.*

Recovered plasmid DNA was double digested and electrophoresed to confirm the positive transformation: 2 bands are expected and their length should be the same as vector and insert, respectively.

When the mini-inoculum had been confirmed to be positive, a midi–inoculum was performed: 5 µl of mini-inoculum were inoculated in 50 ml of LB Broth (Sigma) in presence of ampicillin and incubated overnight at 37°C, in agitation. 1 ml of midi-inoculum was used for a pure mini-preparation with the commercial kit Wizard Plus SV Miniprep DNA Pur System

(Promega) and 450 ng of insert were sequenced (with the same primers used in PCR) to confirm that no mutations were inserted during the cloning process. A midi-preparation was performed on the midi–inoculum with PureLink HiPure Plasmid Midiprep Kit (Life Technologies) to recover a large number of copies of construct.

By cloning, two different constructs were obtained: the pSK_βWT containing wild type promoter and the pSK_β-72 containing the -72(T→A) mutated one.

Two mutated constructs were also generated by site-directed mutagenesis starting from the pSK_βWT: the pSK_β-87 and the pSK_β-71, containing the -87 C→G and -71 C→T mutation, respectively. The mutagenesis reaction mix consisted in: 1X Pfu Buffer, 20 ng of pSK_βWT, 200 µM of dNTPs, 125 ng of each primer (β87G_F 5'-CTCACCCTGTGGAGCCACAC**G**CTAGGGTTGGCCAATCTAC-3'
β87G_R 5'-GTAGATTGGCCAACCCTAG**C**GTGTGGCTCCACAGGGTGAG-3'
β71T_F 5'-CTAGGGTTGGCCAAT**T**TACTCCCAGGAGCAGG-3'
β71T_R 5'CCTGCTCCTGGGAGTA**A**ATTGGCCAACCCTAG-3'), 3U of Pfu in a total reaction volume of 50 µl. Thermocycle parameters were: 30'' at 95°C; 18 cycles of 30'' at 95°C, 1' at 55°C and 90''/kb at 72°C; additional final extension was added (10' at 72°C).

PCR products were subsequently incubated 1 h at 37 °C with 3 U of *Dpn*I to digest the metilated parental plasmid.

All the plasmid constructions were verified by automated sequencing.

9.5. Cell culture and Luciferase assays

To perform the luciferase assay, wild type and mutated promoters were transferred into the pGL2-Basic Luciferase Reporter Vector containing the HS2 region. We generated the recombinant plasmid pGL2-HS2 inserting the HS2-locus control region into the *BamH*I and *Sal*I restriction sites of pGL2-Basic.

Transient transfection was chosen for the luciferase assay: K562 erytroleukaemia cell line was used.

K562 cells were grown and maintained approximately from 1 to $5 \cdot 10^5$ cells/ml in RPMI 1640 GlutaMAX™ medium (Gibco, Thermo Fisher Scientific), containing 10% fetal bovine serum, 100 mg of streptomycin, 100 U/ml of penicilin at 5% CO_2 and 37°C.

A liposome transfection or lipofection was performed with the Lipofectamine LTX reagent (Life Technologies) according to the manufacturer's protocol; briefly 500 ng of total plasmid DNA were added to 0.5 µl of Plus in 100 µl of Opti–MEM serum and antibiotic free medium (Life Technologies) and incubated for 5' at room temperature; 1.25 µl of LTX was then added. After a 30' incubation, the mixture was added to $1 \cdot 10^5$ cells/500 µl in a 24–well plate and incubated for 36–48 h at 37°C. Cationic lipids are used to convey the DNA inside the cells: they create artificial membrane vesicles (liposomes) that bind plasmid DNA molecules; these complexes adhere to and fuse with the negatively charged cell membrane, bringing the DNA within the cell.

Dual Luciferase Assay (Promega) was performed: basically the inserted β promoter upstream the *Firefly* luciferase gene can drive the expression of the enzyme in K562. Mutations on the cloned promoter were tested to understand their ability to decrease luciferase's expression and hence luminescence. The *Renilla* luciferase expression vector pRL–TK (Promega) was cotransfected with the proband pGL2 construct (1:20 pRL:pGL2) for normalization of transfection efficiency and to provide an internal control for lipofection. Also a negative control (with no pGL2 construct) were used in the experiments.

Cells were then lysed and *Firefly* (from pGL2) and *Renilla* (from pRL) luciferase gene activities were measured through a luminometer (Victor X5, PerkinElmer).

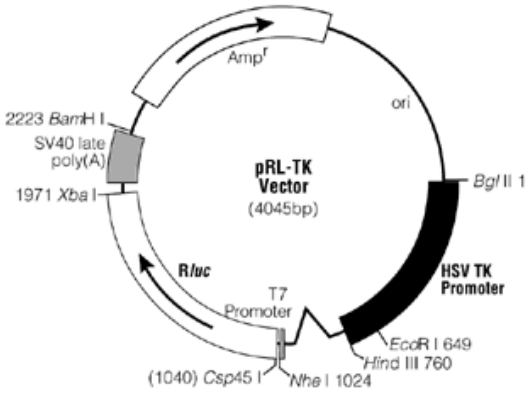


Figure 13: pRL-TK vector used as coreporter in the Dual Luciferase Assay.

The luminescence from the *Firefly* luciferase of each well was normalized to the relative luminescence from the *Renilla* one, in order to obtain comparable values among experiments. The value of the wild type promoter was set as 100% of luminescence.

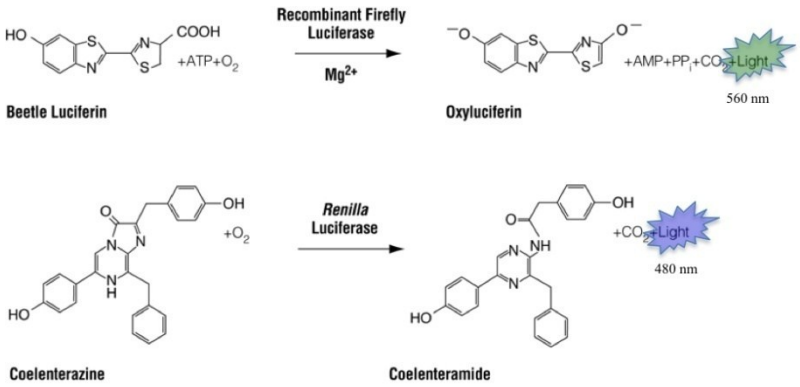


Figure 14: Dual Luciferase Reporter Assay and its bioluminescent reactions catalyzed by firefly and Renilla luciferases.

10. Statistical analysis

Data are presented as the mean ± SD. Results were subjected to the Student's t-test. Differences between groups with $p<0.05$ were considered to indicate a statistically significant. All experiments were performed independently in quadruplicated.

RESULTS

Out of 160 subjects belonging to the study group, 71 (44,4%) showed hematological indices overlapping to the control group, whereas 89 (55,6%) were found to have hypocromic microcytic anemia (Table I).

Table I: Hematological indices for study and control groups

Samples	Hematological indices				
	RBC (x 10^{12}/l) *4-5.5	Hb (g/dl) *12-15	MCV (fl) *80-100	MCH (pg) *28-32	RDW (%) *12-15
Control group (N = 193)	4.79 ± 0.29	13.07 ± 0.84	85.9 ± 3.17	29.4 ± 0.87	13.85 ± 0.74
Study group (N = 89)	5.1 ± 0.49	10.43 ± 2.39	80.13 ±7.12	26.22 ±1.58	16.34 ± 3.32
P value (Student t-test)	NS	P<0.001	P<0.001	P<0.001	NS

NS = non-significant

*Reference values

After selection, the 89 samples have undergone to Hbpathies screening and 27 of them, which showed abnormal qualitative or quantitative Hb profiles, were included in the DNA analysis.

Table II: Hematology and Hb profiles for 27 patients stratified by phenotype or supposed β-genotype.

	SAMPLES			
	Study group (N. 27)			**Control group**
	N = 18	**N = 3**	**N = 6**	**N = 193**
β-genotype *	$\beta^E\beta$	$\beta^E\beta^0$	$(\beta\beta^0)$ Tm or TI	
RBC ($x10^{12}$/l) 4-5.5	4.97± 0.53	5.94; 5.88; 5.76	5.4 ± 0.72	4.79 ± 0.29
Hb (g/dl) 12-15	12.5 ± 1.1	8.5; 8.1; 7.8	10.3 ± 2.09	13.23 ± 0.98
MCV (fl) 80-100	82.8 ± 6.33	71; 62.5; 60.4	69.3 ± 11.97	85.9 ± 3.17
MCH (pg) 28-32	25.6 ± 1.5	21; 18.4; 20.2	24.2 ± 2.82	29.4 ± 0.87
RDW (%) 12-15	14.1 ± 1.23	24.3; 21.7; 19.5	15.9 ± 1.78	13.85 ± 0.74
HbA_2 (%) ≤3.5	3.8 ± 0.38	5.3; 5.8; 6.3	3.9 ± 0.43	2.7 ± 0.3
HbE (%)	24.9 ± 2.13	60.1; 59.2; 58.4	-	-
HbF (%)	< 1.5	34.6; 35; 35.3	1.8 ± 1.2	-

*supposed according to hematology and Hb profiles; Tm: thalassemia minor; TI: thal intermedia

HbE variant was observed in 21 of 27 samples; this is the only structural Hbpathy identified in the study. The Figure 15 shows the chromatographic profile (panel a) and the isoelectrophoretic pattern (panel b) observed in 18 individuals carrying from 22.8 to 27% of the mutated tetramer.

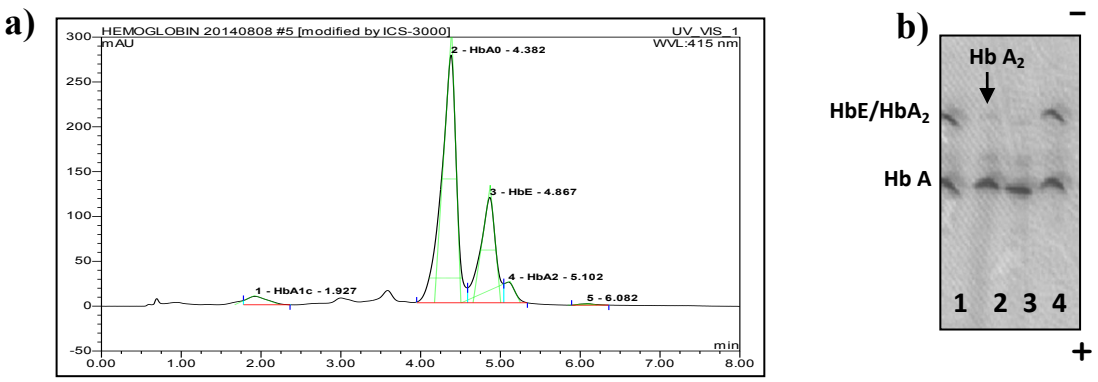


Figure 15: panel a): CE-HPLChromatogram of Hbs in a carrier for HbE variant. Panel b): Isolectrofocusing of Hbs: lanes 1,4 carriers for HbE; lanes 2,3 controls.

These data were in agreement with the $\beta^{E}\beta$ genotype, which was confirmed by sequencing.

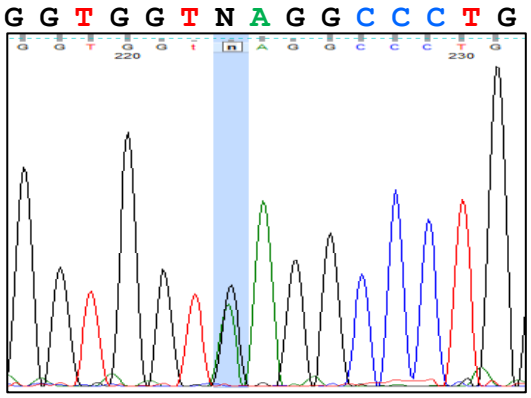


Figure 16: nucleotide sequencing of a region of the β-globin gene from one of the 18 HbE carriers. The GAG(Glu) →AAG(Lys) heterozygosity at codon 26 is shown.

HbE was the predominant circulating Hb fraction in 3 samples (Table II), the remaining being the HbF (~35%). The HbA tetramer was not present.

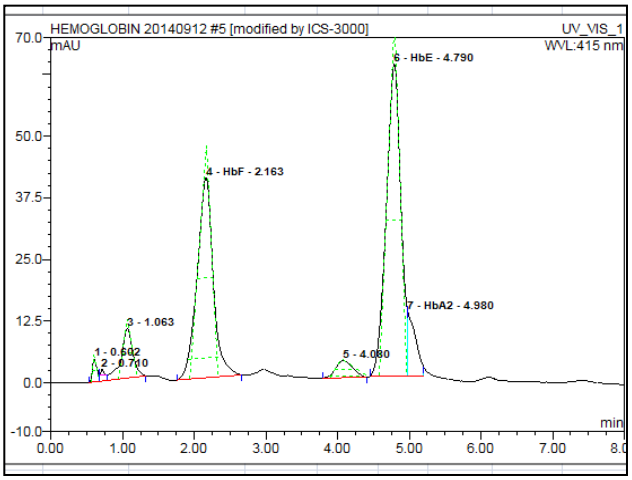


Figure 17: CE-HPLChromatogram of Hbs in a patient supposed as being $\beta^E\beta^0$ compound.

This Hb profile, with a greatly increased production of HbF, is mostly noted in the presence of the $\beta^E\beta^0$ compound state rather than the HbE homozygousity.

Sequencing of the β gene confirmed a compound condition in all 3 cases. In 2 of these, the β^E and the thalassemic β^{017} alleles are coinherited (Fig. 18).

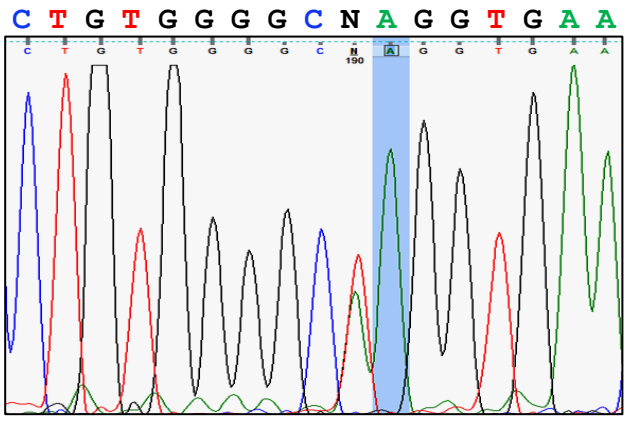


Figure 18: nucleotide sequencing of a region of the β-globin gene showing the AAG (Lys)→TAG (stop) heterozygosity at the 17 codon.

The third sample resulted heterozygous for both the β^E variant and the -TTCT deletion at codons 41/42, which is responsible for a frameshift β^0 mutation resulting in a stop codon at the new codon 59 terminating translation.

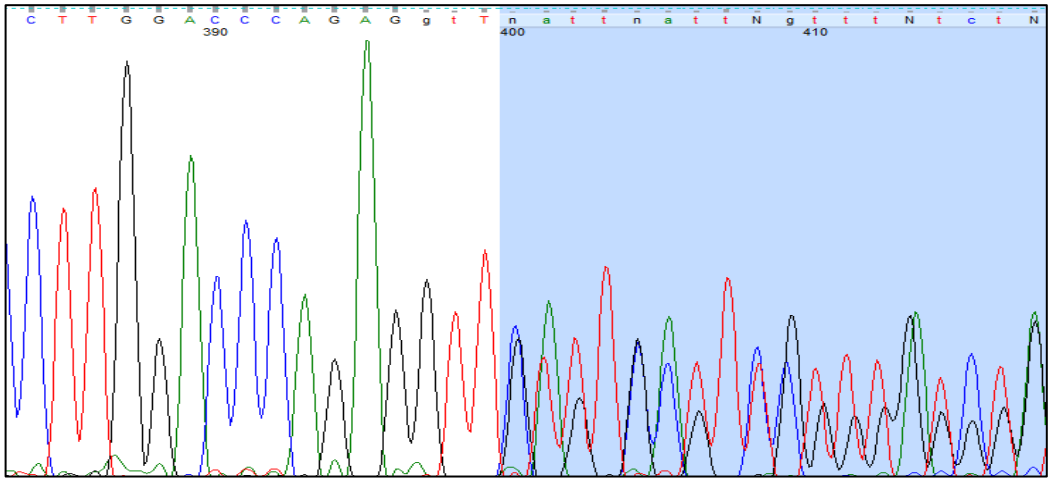


Figure 19: nucleotide sequencing of a region of the β-globin gene showing the frameshift resulting from the -TTCT deletion at codons 41/42.

All of the 6 samples suspected of having β-thalassemia variants (Table II) were undergone to the β-globin gene sequencing. Table III lists the hematological parameters and the assessed β-genotype.

Table III: Hematology and β-genotype in n. 1-6 samples.

Samples	Hb	HbA_2	MCV	MCH	β-genotype*
1	6.8	3.4	81.2	25.6	normal
2	12.4	3.46	73.5	22.4	normal
3	10.0	4.5	52.3	26.5	$\beta\beta^{017}$
4	12.3	4.17	62.7	19.3	$\beta\beta^{017}$
5	9.31	3.7	63.2	25.2	$\beta\beta^{-TTCTdel41/42}$
6	10.8	3.96	83	26.3	$\beta\beta^{-72}$

2 out 6 patients (samples 1 and 2) showed normal β gene sequence, identical to the reference except for several common polymorphisms, listed in Table IV, previously described (Orkin et al., 1984) and not reported as functional changes.

Table IV: Polymorphisms observed in the β-globin gene

Samples	SNPs			
	Codon 2	IVS II nt 16	IVS II nt 74	IVS II nt 666
1	CAT	G	T	C
2	CAT	G	G	C

The homozygosity state observed for all four polymorphisms, the presence of a hypocromic microcytic anemia and borderline HbA_2 levels, prompted us to further investigate the chromosomal arrangement.

The MLPA results have highlighted a normal HBB cluster, both in structure and copy number. The HBA cluster was also investigated and a triplication resulted in the sample n.1 (Figure 16). The presence of Hb Constant Spring was excluded in both patients (HBA MUT CS 135 nt probe).

Samples	HBB cluster	HBA cluster	CS MUT
1	normal	triplicated	Excluded
2	normal	normal	Excluded

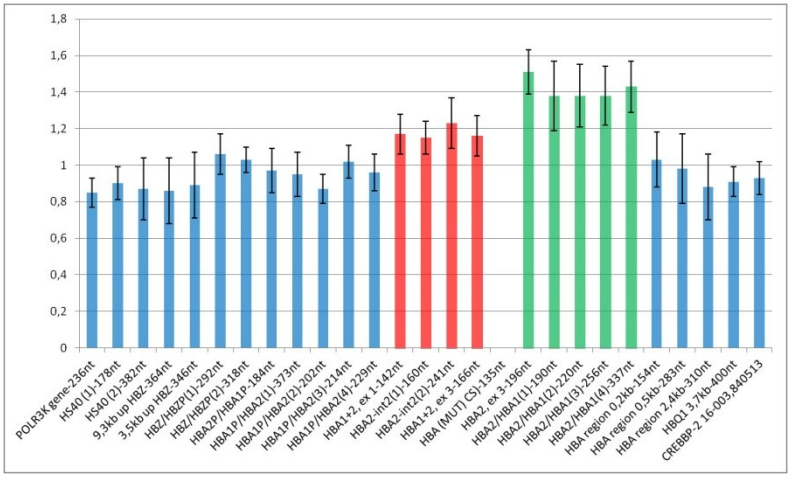
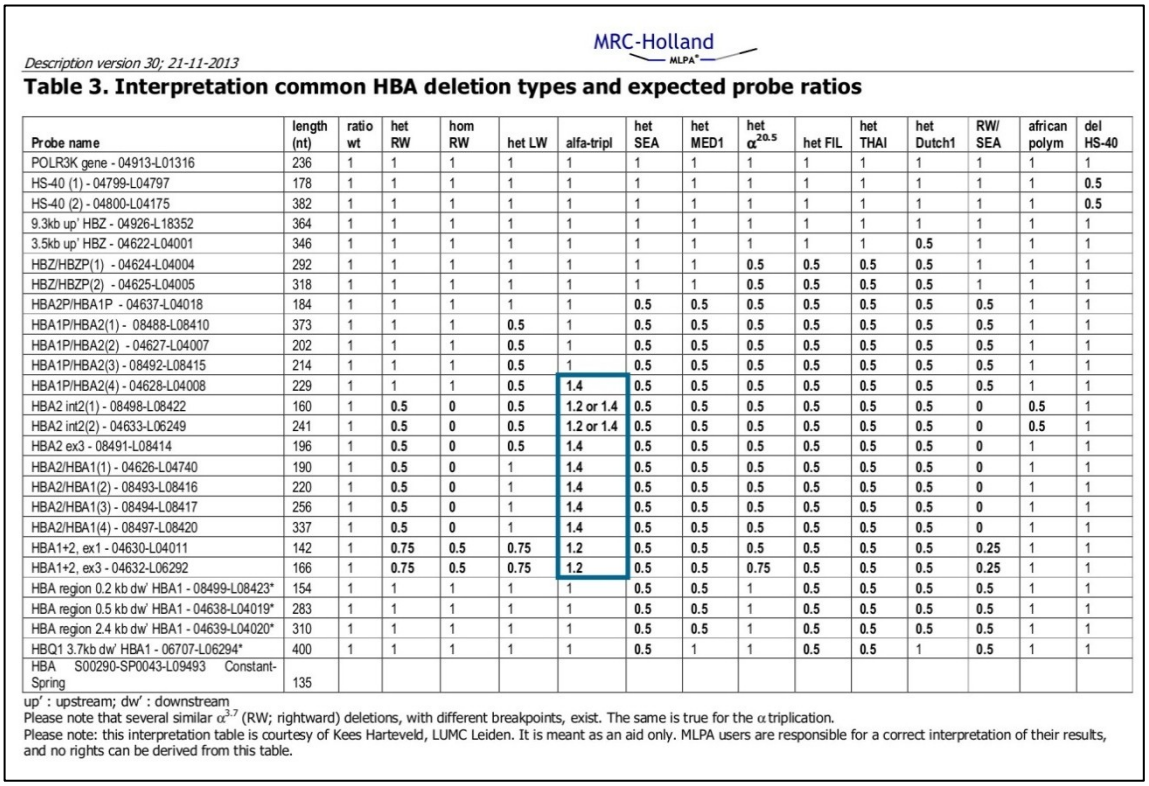


Figure 20: Bar chart for sample n. 1 analyzed with SALSA MLPA kit P140–B4 HBA. In red and green the heterozygous triplicated region.



Description version 30; 21-11-2013

MRC-Holland MLPA®

Table 3. Interpretation common HBA deletion types and expected probe ratios

Probe name	length (nt)	ratio wt	het RW	hom RW	het LW	alfa-tripl	het SEA	het MED1	het $\alpha^{20.5}$	het FIL	het THAI	het Dutch1	RW/ SEA	african polym	del HS-40
POLR3K gene - 04913-L01316	236	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HS-40 (1) - 04799-L04797	178	1	1	1	1	1	1	1	1	1	1	1	1	1	0.5
HS-40 (2) - 04800-L04175	382	1	1	1	1	1	1	1	1	1	1	1	1	1	0.5
9.3kb up' HBZ - 04926-L18352	364	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3.5kb up' HBZ - 04622-L04001	346	1	1	1	1	1	1	1	1	1	1	0.5	1	1	1
HBZ/HBZP(1) - 04624-L04004	292	1	1	1	1	1	1	1	0.5	0.5	0.5	0.5	1	1	1
HBZ/HBZP(2) - 04625-L04005	318	1	1	1	1	1	1	1	0.5	0.5	0.5	0.5	1	1	1
HBA2P/HBA1P - 04637-L04018	184	1	1	1	1	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	1
HBA1P/HBA2(1) - 08488-L08410	373	1	1	1	0.5	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	1
HBA1P/HBA2(2) - 04627-L04007	202	1	1	1	0.5	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	1
HBA1P/HBA2(3) - 08492-L08415	214	1	1	1	0.5	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	1
HBA1P/HBA2(4) - 04628-L04008	229	1	1	1	0.5	1.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	1
HBA2 int2(1) - 08498-L08422	160	1	0.5	0	0.5	1.2 or 1.4	0.5	0.5	0.5	0.5	0.5	0.5	0	0.5	1
HBA2 int2(2) - 04633-L06249	241	1	0.5	0	0.5	1.2 or 1.4	0.5	0.5	0.5	0.5	0.5	0.5	0	0.5	1
HBA2 ex3 - 08491-L08414	196	1	0.5	0	0.5	1.4	0.5	0.5	0.5	0.5	0.5	0.5	0	1	1
HBA2/HBA1(1) - 04626-L04740	190	1	0.5	0	1	1.4	0.5	0.5	0.5	0.5	0.5	0.5	0	1	1
HBA2/HBA1(2) - 08493-L08416	220	1	0.5	0	1	1.4	0.5	0.5	0.5	0.5	0.5	0.5	0	1	1
HBA2/HBA1(3) - 08494-L08417	256	1	0.5	0	1	1.4	0.5	0.5	0.5	0.5	0.5	0.5	0	1	1
HBA2/HBA1(4) - 08497-L08420	337	1	0.5	0	1	1.4	0.5	0.5	0.5	0.5	0.5	0.5	0	1	1
HBA1+2, ex1 - 04630-L04011	142	1	0.75	0.5	0.75	1.2	0.5	0.5	0.5	0.5	0.5	0.5	0.25	1	1
HBA1+2, ex3 - 04632-L06292	166	1	0.75	0.5	0.75	1.2	0.5	0.5	0.75	0.5	0.5	0.5	0.25	1	1
HBA region 0.2 kb dw' HBA1 - 08499-L08423*	154	1	1	1	1	1	0.5	0.5	1	0.5	0.5	0.5	0.5	1	1
HBA region 0.5 kb dw' HBA1 - 04638-L04019*	283	1	1	1	1	1	0.5	0.5	1	0.5	0.5	0.5	0.5	1	1
HBA region 2.4 kb dw' HBA1 - 04639-L04020*	310	1	1	1	1	1	0.5	0.5	1	0.5	0.5	0.5	0.5	1	1
HBQ1 3.7kb dw' HBA1 - 06707-L06294*	400	1	1	1	1	1	0.5	1	1	0.5	0.5	1	0.5	1	1
HBA S00290-SP0043-L09493 Constant-Spring	135														

up' : upstream; dw' : downstream

Please note that several similar $\alpha^{3.7}$ (RW; rightward) deletions, with different breakpoints, exist. The same is true for the α triplication.

Please note: this interpretation table is courtesy of Kees Harteveld, LUMC Leiden. It is meant as an aid only. MLPA users are responsible for a correct interpretation of their results, and no rights can be derived from this table.

Figure 21: The grid shows the expected probe ratios in the presence of deleted or triplicated HBA cluster (from SALSA MLPA kit P140–B4 HBA product description, MRC-Holland).

As indicated in the blue box of Figure 21, the expected probe ratios in the presence of a triplicated HBA cluster are 1.2 or 1.4. According to this, the 142, 160, 241, and 166 probes (red bars in Figure 15) reach a level of 1.2, and the 196, 190, 220, 256 and 337 probes (green bars in Figure 20) showed an increased value of 1.4.

The normal ratio observed for the 229 probe, which differs from the expected 1.4, denotes that the 5' breakpoint of the triplication is different from that shown in Figure 21.

As shown in Table III, 2 patients (samples 3 and 4) resulted heterozygous for the β^{017} allele, whereas 1 sample (n. 5) showed heterozygosity for the -TTCT deletion at codons 41/42.

Moreover, sequence analysis performed in the n. 6 sample revealed heterozygosity for the novel mutation T→A at position -72 of the β-globin gene promoter, within the conserved CCAAT box.

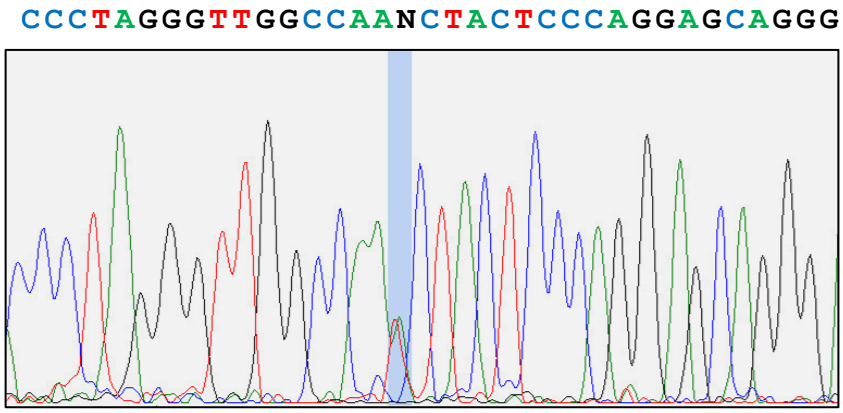


Figure 22 : nucleotide sequencing of the β-globin gene promoter showing the T→A heterozygosity at position -72 from the Cap site, in the CCAAT box.

The index case was a 5-year-old child having red blood cells indices close to normal and slightly increased level of HbA2 (3.96%).

His relatives have been also analyzed: the father's and grandfather's showed similar HbA_2 levels (4.01 and 3.82%, respectively), whereas the other family members displayed normal hematological parameters and Hb components. The hematology data of the proband and his relatives are presented in Table V.

Sequencing revealed that his father and grandfather shared the same $\beta\beta^{-72}$ genotype.

Table V: hematology, β-genotype and MLPA results in the -72 T→A carrier (sample n. 6) and his family members.

Samples	Hematology and genotype						
	Hb	**HbA_2**	**MCV**	**MCH**	**β-genotype***	**HBB cluster#**	**HBA cluster#**
Proband	10.8	3.96	83	26.3	$\beta\beta^{-72}$	normal	Normal
Father	14.1	4.01	96.5	30.5	$\beta\beta^{-72}$	normal	Triplication
Mother	12.0	2.94	96.8	29.7	normal	N.D.	N.D.
Sister	12.1	2.84	85.1	26.9	normal	N.D.	N.D.
Grandfather	13.6	3.82	95.4	29.5	$\beta\beta^{-72}$	normal	N.D.
Grandmother	11.9	2.69	97.5	30	normal	N.D.	N.D.

* defined by sequencing, # defined by MLPA

The common polymorphic sites above-mentioned [codon 2 (CAC→CAT), IVS-II-16 (C→G), IVS-II-74 (G→T), IVS-II-666 (T→C)] were also detected and the following configuration CAT, G, T, C, was found *in cis* to the -72 mutation.

No alteration was observed in HBB cluster of the three -72 carriers, whereas a triplicated HBA cluster, that seems to have no effect on the phenotype, was highlighted in the proband's father. The resulting bar chart was very similar to that of sample n.1 (Figure 20) except for the 229 probe ratio which amounted to 1.55. This value was consistent with that shown in Figure 21.

To assess the contribution of -72 mutation in decreasing β-globin gene expression, wild-type (pGL2-HS2-β-WT) and mutant (pGL2-HS2-β-72) constructs were cotransfected with the pRL control vector into K562 cells.

In order to determine the ability of this system to reproduce the downregulation of the β-gene promoter, we generated and used as positive control the pGL2-HS2-β-87 and pGL2-HS2-β-71 constructs, in which the mutagenized -87(C→G) and -71(C→T) β-promoters were inserted into the pGL2-HS2 vector, respectively. Both -87 and -71 have been already described as β-thalassemia mutations. The β-87 C→G allele is a mild transcriptional mutant described in Mediterranean countries (Huisman & Carver, 1998). It alters the proximal CACCC box, one of the crucial elements for the expression of β-globin gene. Homozygotes (Camaschella et al, 1990) or compound heterozygotes for β-87 and for severe β-thalassemia mutations (Rosatelli et al, 1989) are affected with thalassemia intermedia. Previous expression studies in K562 cells showed a residual activity of 20% to 30% compared with the wild-type promoter (Ristaldi et al., 1995). The -71 C→T mutation occurs one nucleotide immediately downstream of the core CCAAT sequence, which has been shown to be important for the binding of GATA-1 (Al Zadjali et al., 2011). The β-71 allele was found in compound heterozygosity with HbS [β6(A3)Glu>Val] in an Omani family with almost equal expression of HbA and HbS. In addition, molecular screening of a set of subjects with borderline HbA_2 or MCV values revealed the presence of the -71 change in heterozygous state. These results suggested that the -71 C to T mutation may be a mild β-thalassemic allele, although at the time of our study results on gene expression were not present.

In our experiment the expression of the mutant vectors was compared with the pGL2-HS2-β-WT, which was considered to have 100% activity.

The results of the transfection studies are summarized in Figure 23.

Luciferase activities of the mutant controls (-87 and -71) resulted significantly decreased, demonstrating that the system is able to reproduce a down-regulation of the β-globin gene promoter *in vitro*. Relative luciferase activities of -87, -72 and -71 mutated promoters were 32,3 ± 0.7%, 53.7 ± 7.5% and 46.1 ± 4.8%, respectively (Figure 23). These results clarify that -72 mutation, as well as described for -87 and -71, is a mild β-thalassemic allele.

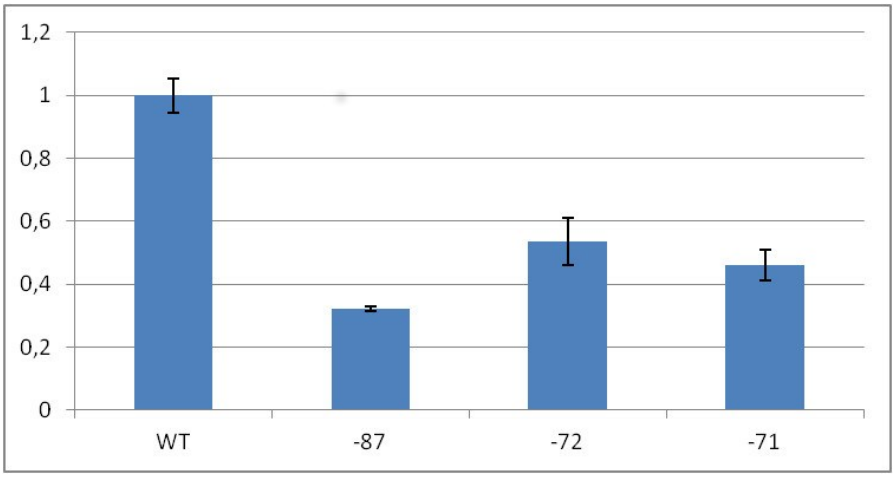


Figure 23: Relative luciferase activity of β-globin promoter in K562 cells.

DISCUSSION

Although Hbpathies were the first human disorders to be characterized as molecular diseases and their pathophysiology understood, after 60 years of significant progress, the management of sickle cell disease and β-thalassemia still depends on supportive care and when required, regular lifelong blood transfusions and iron chelation.

In some developing countries, where due to the lack of resources patients are untreated or poorly transfused, the clinical picture of thalassemia major is characterized by growth retardation, pallor, jaundice, poor musculature, genu valgum, hepatosplenomegaly, leg ulcers, development of masses from extramedullary hematopoiesis, and skeletal changes resulting from expansion of the bone marrow. Other complications are hypersplenism, chronic hepatitis (resulting from infection with viruses that cause hepatitis B and/or C), HIV infection, venous thrombosis, and osteoporosis (Cao & Galanello, 2010).

Approximately 80% of the annual births of babies with severe conditions occur in low-or middle-income countries, many of which have extremely limited facilities for their control and management. Given that the population size of many of them is growing and, as social and public health facilities improve, increasing numbers of these babies will survive to present for diagnosis and treatment. Hence, the Hbpathies will constitute an increasing global health burden.

Programmes directed at the avoidance and better management of these conditions have the potential to significantly improve health indicators in many developing countries (Weatherall et al, 2006; Modell & Darlison, 2008). Identification of new mutations and the update of the mutation spectrum of thalassemia in one ethnic population is a prerequisite for the development of these programs.

In Vietnam, both β-thalassemia and HbE are prevalent and represent an important cause of childhood chronic disease (Svasti et al., 2002).

The Hb E heterozygote is mildly affected and the Hb E homozygosity is a benign disorder with a mild β-globin chain deficit which is comparable to that seen in a β^0-thalassemia heterozygote. However, compound heterozygotes β^E/β^0 are often severely affected.

The Vietnamese population is ethnically highly heterogeneous and the spectrum of β-thalassemia alleles is slowly defining (Filon et al., 2000; Hao et al., 2001; Svasti et al., 2002; O'Riordan et al., 2010; Nguyen et al., 2013). Patients included in these studies had severe clinical symptoms

and were transfusion-dependent. On the whole, six mutations of the β^0 and a mutation of the β^+ type have been identified and observed with distinct incidence in the different areas. The most commons, both in the South and in the North of Vietnam, are the frameshift mutation at codons 41/42 (−TCTT), the nonsense mutation at codon 17 (A→T) and the frameshift mutation at codon 95 (+A), which is known as the "Vietnamese" mutation. Other mutations, although less frequent, are present in these countries: the -28 (A→G), the IVS-I-1 (G→T), the frameshift mutation at codons 71/72 (+A) and the IVS-II-654 (C→T).

The aim of the PhD project has been to develop professional skills and competences directed at the avoidance and better management of the Hbpathies in central Vietnam. The project was focused to provide the groundwork for screening programs determining the spectrum of β-thalassemia mutations in patients and their relatives from Hue province of Vietnam.

A complete blood cell count has been performed on 353 individuals of which 89 showed abnormal hematological parameters like anemia, microcytosis and/or hypochromia.

Hb phenotype determination led to detect altered Hb profiles in 27 samples out of 89 examined.

Twenty one patients showed an Hb profile consistent with the presence of the HbE variant [β26(B8)Glu→Lys]. The mutated tetramer is resolved from the normal counterpart by both IEF and CE-HPLC techniques since its additional positive charge. Nevertheless, elution and comigration of HbE with HbA_2, as occurs for other slow-moving variants, required a molecular DNA approach to carry out a correct diagnosis.

Sequencing confirmed the presence of the GAG→AAG mutation at 26 codon and allowed to define the β-genotype which resulted $\beta^E\beta$ in 18 samples, and $\beta^E\beta^0$ in the remaining 3. In these last, sequencing revealed two previously described mutations in addition to the β^E allele: the nonsense change at codon 17 (AAG→TAG), observed in 2 samples, and the frameshift mutation at codon 41/42(-TTCT). The -TTCT deletion is the most common mutation in South and Southeast China where its frequency reaches above 40%. The β^{017} allele has a high prevalence in Southern Chinese provinces as well as in North Thailand (Filon et al., 2000). These two mutations also had highest frequencies in North and South Vietnam (Filon et al., 2000; Svasti et al., 2002).

The compound heterozygous state for HbE and β^0-thalassemia results in a remarkably variable phenotype: approximately one half of the patients are phenotypically similar to patients with TM who require regular transfusion therapy, and the other half resembles TI (Fucharoen et al., 2000;

Gibbons et al., 2001). Notable are variations in anemia, growth, development, hepatosplenomegaly, and transfusion requirements (Fucharoen et al., 2000).

The cause of the striking variability remains largely unknown. Despite seemingly identical genotypes, patients of the same family may show significant differences in clinical severity. HbF level is the strongest predictor of morbidity. However, the basis of increased HbF is usually unknown.

In a review of 378 patients with HbE- β^0-thalassemia from Thailand, the Hb concentrations ranged from 3 to 13 g/dl, with an average of 7.7 g/dl (Fucharoen et al., 2000). In our study, the mean value of Hb concentration was 8.13 g/dl, with HbF markedly elevated. None of the 3 patients undergoes a regular transfusion program. They have occasionally received transfusions and the level of persistent HbF synthesis appears to be the most important factor in their clinical mildness.

The presence of a hypocromic microcytic anemia and increased or borderline HbA_2 levels was suggestive of β-thalassemia variants for 6 of 27 samples.

Two of them (samples 3 and 4) resulted heterozygous for the β^{017} allele, whereas one sample (n. 5) showed heterozygosity for the -TTCT deletion at codons 41/42.

Conversely, any functional mutation was observed in the β gene of samples 1 and 2, but only common polymorphisms were revealed: the C→T silent mutation at codon 2 (His→His), and three changes in IVS II (+16C→G or *Ava* II − → +, +74G→T and +666T→C) (Orkin et al., 1982; Sahoo et al., 2014). The homozygosity state observed for all polymorphisms above-mentioned, despite their hypervariability, prompted us to further investigate the presence of a deletion involving the β-globin gene *in trans*. Indeed, deletional forms of β-thalassemia, mainly caused by unequal crossing over, were described (Thein, 2013). The MLPA results have highlighted a normal HBB cluster, both in structure and copy number. In order to check a possible co-inheritance of α-thalassemia triplication, the HBA cluster was also investigated. In fact, triplications and quadruplications of the α-globin gene, frequently observed in many populations, can interact with β-thalassemia determinants, to produce moderate to severe phenotypes. A triplicated arrangement was found in sample n.1, whereas sample n. 2 revealed a normal HBA. These results do not explain the clinical features of the two samples, especially for sample n. 1, a 7-year-old child. He showed a severe anemia state with the following Hb

components: HbA 92.4%; HbF 4.2%; HbA_2 3.4%. After his Hb has fallen below 7.0 g/dl he was transfused and until now he is in regular transfusion program.

It is noteworthy that there is emerging evidence that a set of genes, linked to the HBB cluster and trans-acting, contribute to Hb regulation and inherited hemolytic anemia. For example, rare variants in KLF1 have been described in association with borderline or increased HbA_2 and HbF levels and/or microcytic hypochromic anemia in the absence of mutations in β-globin gene (Perseu et al., 2011; Huang et al., 2015). KLF1 is an essential erythroid transcription factor, which plays multifunctional role during erythropoiesis and Hb switching between fetal and adult states.

Moreover, hereditary hemolytic anemia can involve also different genetic determinants, as RBC membrane disorders or RBC enzyme disorders. Further studies are warranted to elucidate the molecular basis of the observed phenotype.

Sample n. 6, a 5-year-old child, presented a mild microcytic anemia with a Hb level of 10.8 g/dL, MCV of 83.0 fL, MCH of 26.3 pg and a slightly increased HbA_2 level of 3.96%. He had never required transfusions. The father's and grandfather's showed similar HbA_2 levels (4.01 and 3.82 %, respectively). The other family members showed normal hematological parameters and Hb components.

The proband was found to have a novel to literature T→A mutation at position -72 from the Cap site, in the CCAAT box of the β promoter region. Sequencing revealed that his father and grandfather had the same genotype $\beta^{-72}\beta$.

The common polymorphic sites above-mentioned (codon 2, IVS-II-16, IVS-II-74, IVS-II-666) were also detected and the following configuration CAT, G, T, C, was found *in cis* to the -72 mutation. The coupling of specific β-globin gene mutations with neutral changes has been widely described (Orkin et al., 1982b; Sahoo et al., 2014); it allowed the detailed characterization of chromosome regions in which mutant β-globin genes reside.

The β-globin gene cluster is a good model to study gene expression and regulation. The efficient transcription of the β-globin gene is dependent upon the fidelity of several conserved DNA motifs within the promoter regions. These motifs in the β-globin gene include a duplicated CACCC box: proximal at positions -90 to -86 and distal at positions -105 to -101 from the Cap site, the CCAAT box at -76 to -72, and the TATA box at -30 to -26. These elements, in particular

the CCAAT box, are important in the regulation of developmental and tissue-specific expression of globin genes (Thein, 2013).

Mutations observed in these conserved sequence motifs, lead to a slight decrease in β-globin gene expression and are associated with relatively mild forms of β-thalassemia and borderline HbA_2 levels (http://www.globin.cse.psu.edu/) (Orkin et al., 1984b; Agarwal et al., 2006; Chen et al., 2007; Basran et al., 2008; Waye et al., 2011).

To date, one mutant at β-promoter, the -28 A→G, has been identified in Vietnam, only in the southern (Hao et al., 2001; Svasti et al., 2002). The same mutation was also reported in the two countries neighboring Vietnam, China and Thailandia (Zhang et al., 1988; Fucharoen et al., 1997). This variant occurs in the TATA box and has been introduced into the HbVar database as β^+-thalassemic allele. Indeed, the mild thalassemic phenotype observed in compound heterozygotes β^E/β^{-28} from Thai population may indirectly indicate that the -28 mutation is a mild thalassemic allele (Fucharoen et al., 1997).

The -72 mutation, identified in our study for the first time, occurs in simple heterozygous state so its responsibility in causing a β-thalassemia phenotype and the mutation type (β^0 or β^+) cannot be deduced. In order to measure the degree of reduction in the promoter activity *in vitro* expression studies were performed.

We used as positive controls two mutant constructs in which the mutagenized -87(C→G) and -71(C→T) β-promoters were inserted into the basic vector, respectively. The β-87 C→G allele is a mild β-thalassemia mutant that alters the proximal CACCC box. Previous expression studies in K562 cells have showed a residual activity of 20% to 30% compared with the wild-type promoter (Ristaldi et al., 1995). The -71 C→T mutation occurs one nucleotide immediately downstream of the core CCAAT sequence. Based on hematological phenotypes in simple heterozygotes, as well as in compound heterozygotes with HbS, the mutation was assigned as a mild β^+-thalassemic allele (Al Zadjali et al., 2011).

Our *in vitro* experiments shows that the transcriptional activity of the mutated promoter is roughly half that of the wild type promoter. This finding suggests that the -72 mutation can be classified as a β^+-thalassemic allele. However, the absence of homozygous or compound heterozygous states does not allow us to precisely predict its clinical impact and, consequently, its relevance in management programs. Indeed, association of β^+-thalassemia with β^0 or β^E

mutations results in a markedly heterogeneous hematological picture, ranging in severity from that of the β-thalassemia carrier state to that of TM (Thein, 2005).

Furthermore, it has been noted that even in the case of mild disease, β-TI patients may still suffer from many complications including a hypercoagulable state and subsequent thrombotic events (Cao & Galanello, 2010; Musallam & Taher, 2011).

The ability to predict phenotype from genotype has important implications for the screening of β-thalassemia carriers, for genetic counseling and prenatal diagnosis and for planning the appropriate treatment regimen. This ability requires in turn a comprehensive knowledge of the spectrum of β-thalassemia mutations, more so in countries, like Vietnam, in which the extraordinary phenotypic diversity presents particular management problems.

Our results further underline the importance of identifying and characterizing new or rare β-thalassemic alleles, even when mild, in carrier screening and prenatal diagnosis in order to reduce the burden of thalassemias, avoid unnecessary transfusions in TI and start early transfusions in TM patients.

REFERENCES

Agarwal S, Arya V, Stolle CA, Pradhan M (2006) A novel Indian beta-thalassemia mutation in the CACCC box of the promoter region. *Eur J Haematol* 77:530-532

Al Zadjali S, Wali Y, Al Lawatiya F, Gravell D, Alkindi S, Al Falahi K, Krishnamoorthy R, Daar S (2011) The β-globin promoter -71 C>T mutation is a β+ thalassemic allele. *Eur J Haematol* 87:457-460.

Basran RK, Reiss UM, Luo HY, Ware RE, Chui DH (2008) Beta-thalassemia intermedia due to compound heterozygosity for two beta-globin gene promoter mutations, including a novel TATA box deletion. *Pediatr Blood Cancer* 50:363-366.

Bender MA, Bulger M, Close J, Groudine M (2000) β-globin gene switching and DNase I sensitivity of the endogenous β-globin locus in mice do not require the locus control region. *Mol Cell* 5:387-393.

Bender MA, Byron R, Ragoczy T, Telling A, Bulger M, Groudine M (2006) Flanking HS-62.5 and 3'HS1, and regions upstream of the LCR, are not required for beta-globin transcription. *Blood* 108:1395-1401.

Blackwell R (1965) Hemoglobin E in Vietnam. *Nature* 207:768.

Bunn HF, Forget BG (1986). Hemoglobin: Molecular, genetic and clinical aspects. W.B. Saunders, Philadelphia.

Bunn HF, Noguchi CT, Hofrichter J, Schechter GP, Schechter AN, Eaton WA (1982) Molecular and cellular pathogenesis of hemoglobin SC disease. *Proc Natl Acad Sci* 79: 7527-7531.

Busslinger M, Moschanas N, Flavell RA (1981) β+- thalassemia: aberrant splicing results from a single point mutation in an intron. *Cell* 27:289.

Camaschella C, Alfarano A, Gottardi E, Serra A, Revello D, Saglio G. (1990) The homozygous state for the -87 C->G beta + thalassaemia. *Br J Haematol* 75:132-133.

Camaschella C, Mazza U, Roetto A, Gottardi E, Parziale A, Travi M, Fattore S, Bacchiega D, Fiorelli G, Cappellini MD (1995) Genetic interactions in thalassemia intermedia: analysis of β-mutations, β- genotype, β-promoters, and β-LCR hypersensitive sites 2 and 4 in Italian patients. *Am J Hematol* 48:82-87.

Cantu-Medellin N, Vitturi DA, Rodriguez C, Murphy S, Dorman S, Shiva S, Zhou Y, Jia Y, Palmer AF, Patel RP (2011) Effects of T- and R-state stabilization on deoxyhemoglobin-nitrite reactions and stimulation of nitric oxide signaling. *Nitric Oxide* 25:59-69.

Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P (2002) Longrange chromatin regulatory interactions in vivo. *Nat Genet* 32:623-626.

Chan V, Chan TK, Todd D (1988) Different forms of Hb H disease in the Chinese. *Hemoglobin* 12:499-507.

Chang JC, Kan YW (1979) β^0Thalassemia, a nonsense mutation in man. *Proc Natl Acad Sci USA 76*:2886-2889.

Chen XW, Mo QH, Li Q, Zeng R, Xu XM (2007) A novel mutation of -73(A → T) in the CCAAT box of the β -globin gene identified in a patient with the mild β -thalassemia intermedia. *Ann Hematol* 86:653-657.

Chui DHK, Waye JS (1998) Hydrops fetalis caused by α-thalassemia: an emerging health care problem. *Blood 91*:2213-2222.

Clegg JB, Weatherall DJ, Milner PF (1971) Haemoglobin Constant Spring: a chain termination mutant? *Nature* 234:337-340.

Clegg JB, Weatherall DJ. (1974) Hemoglobin Constant Spring, and unusual alpha-chain variant involved in the etiology of hemoglobin H disease. *Ann N Y Acad Sci* 232:168-178.

Craig JE, Kelly SJ, Barnetson R, Thein SL. (1992). Molecular characterization of a novel 10.3 kb deletion causing bthalassaemia with unusually high Hb A2. *Br J Haematol* 82: 735-744.

Donze D, Townes TM, Bieker JJ (1995) Role of erythroid Kruppellike factor in human gamma- to beta-globin gene switching. *J Biol Chem* 270:1955-1959.

Efremov GD, Wrightstone RN, Huisman TH, Schroeder WA, Hyman C, Ortega J, Williams KJ (1971) An unusual hemoglobin anomaly and its relation to alpha-thalassemia and hemoglobin-H disease. *Clin Invest* 50:1628-1636.

Evans T, Felsenfeld G (1989) The erythroid-specific transcription factor Eryf1: a new finger protein. *Cell* 58:877-885.

Feng WC, Southwood CM, Bieker JJ (1994) Analyses of beta thalassemia mutant DNA interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor. *J Biol Chem* 269:1493-1500.

Filon D, Oppenheim A, Rachmilewitz EA, Kot R, Truc DB (2000) Molecular analysis of β-Thalassemia in Vietnam. *Hemoglobin* 24:99-104.

Forget BG, Bunn HF (2013) Classification of the disorders of hemoglobin. Cold Spring Harb Perspect Med 3:a011684.

Forrester WC, Thompson C, Elder JT, Groudine M. (1986) A developmentally stable chromatin structure in the human beta-globin gene cluster. *Proc Natl Acad Sci USA* 83:1359-1363.

Forsberg EC, Downs KM, Christensen HM, Im H, Nuzzi PA, Bresnick EH (2000) Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. *Proc Natl Acad Sci USA* 97:14494-14499.

Fucharoen S, Ketvichit P, Pootrakul P, Siritanaratkul N, Piankijagum A, Wasi P. (2000) Clinical manifestation of beta thalassemia/hemoglobin E disease. *J Pediatr Hematol Oncol* 22:552-557.

Fucharoen S, Viprakasit V. (2009) Hb H disease: clinical course and disease modifiers. Am Soc Hematol-*Education Book* 2009:26-34.

Fucharoen S, Winichagoon P (1997) Hemoglobinopathies in Southeast Asia: molecular biology and clinical medicine. *Hemoglobin* 21:299- 319.

Fucharoen S, Winichagoon P, Pootrakul P, Piankijagum A, Wasi P. (1988) Differences between two types of Hb H disease, α-thalassemia 1/α-thalassemia 2 and α- thalassemia 1/Hb Constant Spring. *Birth Defects Orig Artic Series* 23:309-315.

Fucharoen S, Winichagoon P. (1987) Hemoglobinopathies in Southeast Asia. *Hemoglobin* 11:65-88.

Fucharoen S., Winichagoon P. (2011) Haemoglobinopathies in Southeast Asia. *Indian J. Med Res* 134:498-506.

Fukumaki Y, Ghosh PK, Benz EJ Jr, Reddy VB, Lebowitz P, Forget BG, Weissman SM. (1982) Abnormally spliced messenger RNA in erythroid cells from patients with β+ thalassemia and monkey cells expressing a cloned β+-thalassemic gene. *Cell* 28:585-593.

Galanello R, Aru B, Dessì C, Addis M, Paglietti E, Melis MA, Cocco S, Massa P, Giagu N, Barella S (1992) Hb H disease in Sardinia: molecular, hematological and clinical aspects. *Acta Haematol* 88:1-6.

Galanello R, Dessi E, Melis MA, Addis M, Sanna MA, Rosatelli C, Argiolu F, Giagu N, Turco MP, Cacace E (1989) Molecular analysis of beta zero-thalassemia intermedia in Sardinia. *Blood* 74:823-7.

Galanello R, Origa R (2010) Beta-thalassemia. *Orphanet J Rare Dis*. 5:11.

Giardine B, Borg J, Higgs DR, Peterson KR, Philipsen S, Maglott D, Singleton BK, Anstee DJ, Basak AN, Clark B, Costa FC, Faustino P, Fedosyuk H, Felice AE, Francina A, Galanello R, Gallivan MV, Georgitsi M, Gibbons RJ, Giordano PC, Harteveld CL, Hoyer JD, Jarvis M, Joly P, Kanavakis E, Kollia P, Menzel S, Miller W, Moradkhani K, Old J, Papachatzopoulou A, Papadakis MN, Papadopoulos P, Pavlovic S, Perseu L, Radmilovic M, Riemer C, Satta S, Schrijver I, Stojiljkovic M, Thein SL, Traeger-Synodinos J, Tully R, Wada T, Waye JS, Wiemann C, Zukic B, Chui DH, Wajcman H, Hardison RC, Patrinos GP (2011) Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach. *Nat Genet* 43:295-301.

Gibbons R, Higgs DR, Olivieri NF, Wood WG (2001) The β and δβ thalassaemias in association with structural haemoglobin variants. In: Weatherall DJ, Clegg JB, eds. The Thalassaemia Syndromes (Fourth ed). Oxford, United Kingdom: *Blackwell Science* pp. 393-449.

Gilman JG (1987). The 12.6 kilobase DNA deletion in Dutch b0-thalassaemia. *Br J Haematol* 67:369-372.

Giordano PC, Bakker-Verwij M, Harteveld CL (2009). Frequency of alpha-globin gene triplications and their interaction with beta-thalassemia mutations. *Hemoglobin* 33:124-131.

Gong Q, McDowell JC, Dean A (1996) Essential role of NF-E2 in remodeling of chromatin structure and transcriptional activation of the epsilon-globin gene in vivo by 5'hypersensitive site 2 of the beta-globin locus control region. *Mol Cell Biol* 16:6055-6064.

Grosveld F, van Assendelft GB, Greaves DR, Kollias G. (1987) Position independent, high-level expression of the human beta- globin gene in transgenic mice. *Cell* 51:975-985.

Groudine M, Kohwi-Shigematsu T, Gelinas R, Stamatoyannopoylos G, Papyannopoulou T (1983) Human fetal to adult hemoglobin switching: changes in chromatin structure of the beta-globin gene locus. *Proc Natl Acad Sci USA* 80: 7551-7555.

Hamid M, Akbari MT. (2011). A 13-bp deletion in the 30 untranslated region of the β-globin gene causes β-thalassemia major in compound heterozygosity with IVSII-1mutation. *Med Princ Pract* 20:488-490.

Hao LT, Pissard S, Van PH, Lacombe C, Hanh TD, Goossens M, Kiet TD (2001) Molecular analysis of β-thalassemia in South Vietnam. *Hemoglobin* 25:305-309.

Hardison R, Slightom JL, Gumucio DL, Goodman M, Stojanovic N, Miller W (1997) Locus control regions of mammalian beta-globin gene clusters: combining phylogenetic analyses and experimental results to gain functional insights. *Gene* 205:73-94.

Harju S, Navas PA, Stamatoyannopoulos G, Peterson KR (2005). Genome architecture of the human beta-globin locus affects developmental regulation of gene expression. *Mol Cell Biol* 25:8765-8778.

Hattori Y, Yamashiro Y, Matsuno Y, Ohba Y, Miyaji T, Yamamoto K, Yamamoto K, Omine M, Shimada I. (1988) A beta+-thalassemia (codon 24, GGT->GGA) found in a Japanese. *Hemoglobin* 12:655-660.

Higgs DR (2009) The molecular basis of α thalassemia in "Disorders of Haemoglobin: Genetics, Pathophysiology, and Clinical Management (eds. Steinberg MH, Forget BG, Higgs DR, Weatherall DJ,. Cambridge University Press, Cambridge), pp. 241-265.

Ho PJ, Hall GW, Luo LY,Weatherall DJ, Thein SL. (1998) Beta thalassaemia intermedia: is it possible to predict phenotype from genotype? *Br J Haematol* 100:70-78.

Ho PJ, Rochette J, Fisher CA, Wonke B, Jarvis MK, Yardumian A, Thein SL (1996) Moderate reduction of β-globin gene transcript by a novel mutation in the 5' untranslated region: a study of its interaction with other genotypes in two families. *Blood* 87:1170-1178.

Ho PJ, Wickramasinghe SN, Rees DC, Lee MJ, Eden A, Thein SL (1997) Erythroblastic inclusions in dominantly inherited beta thalassemias. *Blood* 89:322-328.

Huang J, Zhang X, Liu D, Wei X, Shang X, Xiong F, Yu L, Yin X, Xu X (2015) Compound heterozygosity for KLF1 mutations is associated with microcytic hypochromic anemia and increased fetal hemoglobin. *Eur J Hum Genet* 23:1341-1348.

Huang S-C, Benz EJJ (2001) Posttranscriptional factors influencing the hemoglobin content of the red cells. In: Steinberg MH, Forget BG, Higgs DR, Nagel RL, eds. *Disorders of Hemoglobin: Genetics, Pathophysiology and Clinical Management*. Cambridge: Cambridge University Press pp.252-276.

Huisman TH, Carver MF (1998) Recombinant hemoglobin variants. *Hemoglobin.* 22:99-112.

Humphries RK, Ley TJ, Anagnou NP, Baur AW, Nienhuis AW (1984) Beta O-39 thalassemia gene: a premature termination codon causes beta-mRNA deficiency without affecting cytoplasmic beta-mRNA stability. *Blood* 64:23-32.

Jane SM, Gumucio DL, Ney PA, Cunningham JM, Nienhuis AW (1993) Methylation-enhanced binding of Sp1 to the stage selector element of the human gamma-globin gene promoter may regulate development specificity of expression. *Mol Cell Biol* 13:3272-3281.

Lal A, Goldrich ML, Haines DA, Azimi M, Singer ST, Vichinsky EP (2011) Heterogeneity of hemoglobin H disease in childhood. *N Engl J Med* 364:710-718.

Le TH, Pissard S, Pham HV, Lacombe C, Truong DH, Goossens M, Truong DK. (2001) Molecular analysis of β-thalassemia in South Vietnam. *Hemoglobin* 25:305-309.

Li J, Noguchi CT, Miller W, Hardison R, Schechter AN (1998) Multiple regulatory elements in the 5'-flanking sequence of the human epsilon-globin gene. *J Biol Chem* 273:10202-10209.

Lithanatudom P, Wipasa J, Inti P, Chawansuntati K, Svasti S, Fucharoen S, Kangwanpong D, Kampuansai J (2016) Hemoglobin E Prevalence among Ethnic Groups Residing in Malaria-Endemic Areas of Northern Thailand and its lack of association with *Plasmodium falciparum* invasion in vitro. *PLoS One* 11:e0148079.

Ma ESK, Chow EYD, Chan AY, Chan LC (2001b) Interaction between (--SEA) α-thalassemia deletion and uncommon non-deletional α-globin gene mutations in Chinese patients. *Hematologica 86*:539-540.

Ma SK, Au WY, Chan AY, Chan LC (2001) Clinical phenotype of triplicated alphaglobin genes and heterozygosity for beta0- thalassemia in Chinese subjects. *Int J Mol Med 8*:171-175.

Maragoudaki E, Kanavakis E, Traeger-Synodinos J, Vrettou C, Tzetis M, Metaxotou-MavrommatiA, Kattamis C (1999) Molecular, haematological and clinical studies of the −101 C→T substitution of the β-globin gene promoter in 25 β-thalassaemia intermedia patients and 45 heterozygotes. *Br J Haematol* 107:699-706.

Masala B, Manca L (1991) Detection of the common Hb F Sardinia [A gamma (E19)Ile→Thr] variant by isoelectric focusing in normal newborns and in adults affected by elevated fetal hemoglobin syndromes. *Clin Chim Acta* 198:195-202.

Masala B, Manca L, Gallisai D, Stangoni A, Lanclos KD, Kutlar F, Yang KG, Huisman TH (1988) Biochemical and molecular aspects of beta-thalassemia types in northern Sardinia. *Hemoglobin* 12:661-671.

Merika M, Orkin SH (1995) Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Kruppel family proteins Sp1 and EKLF. *Mol Cell Biol* 15:2437-2447.

Metherall JE, Collins FS, Pan J, Weissman SM, Forget BG (1986) β0-thalassemia caused by a base substitution that creates an alternative splice acceptor site in an intron. *EMBO J* 5:2551-2557.

Miller IJ, Bieker JJ (1993) 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* 13:2776-2786.

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215.

Milner PF, Clegg JB, Weatherall DJ (1971). Haemoglobin-H disease due to a unique haemoglobin variant with an elongated alpha-chain. *Lancet* 1:729-732.

Milot E, Strouboulis J, Trimborn T, Wijgerde M, de Boer E, Langeveld A, Tan-Un K, Vergeer W, Yannoutsos N, Grosveld F, Fraser P (1996). Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* 87:105-114.

Modell B, Darlison M (2008) Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ* 86:480-487.

Musallam KM, Taher AT (2011) Thrombosis in thalassemia: why are we so concerned? *Hemoglobin* 35:503-510.

Ney P, Sorrentino B, McDonagh K, Nienhuis A (1990) Tandem AP-1-binding sites within the human beta-globin dominant control region function as an inducible enhancer in erythroid cells. *Genes Dev* 4:993-1006.

Nguyen HV, Sanchaisuriya K, Nguyen D, Phan HT, Siridamrongvattana S, Sanchaisuriya P, Fucharoen S, Fucharoen G, Schelp FP (2013) Thalassemia and hemoglobinopathies in Thua Thien Hue Province, Central Vietnam. *Hemoglobin* 37:333-342.

Nguyen VH, Sanchaisuriya K, Wongprachum K, Nguyen MD, Phan TT, Vo VT, Sanchaisuriya P, Fucharoen S, Schelp FP. (2014) Hemoglobin Constant Spring is markedly high in

women of an ethnic minority group in Vietnam: a community-based survey and hematologic features. *Blood Cells Mol Dis* 52:161-165.

Nienhuis AW, Nathan DG (2012) Pathophysiology and Clinical Manifestations of the β-Thalassemias. *Cold Spring Harb Perspect Med.* 2 a011726.

Nuez B, Michalovich D, Bygrave A, Ploemacher R, Grosveld F (1995). Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature*. 375:316-318.

O'Riordan S., Hien T.T., Miles K., Allen A, Quyen NN, Hung NQ, Anh DQ, Tuyen LN, Khoa DB, Thai CQ, Triet DM, Phu NH, Dunstan S, Peto T, Clegg J, Farrar J, Weatherall D (2010) Large scale screening for haemoglobin disorders in southernVietnam: implications for avoidance and management. *Br J Haematol* 150:359-364.

Orkin SH (1992) GATA-binding transcription factors in hematopoietic cells. *Blood* 80:575-581.

Orkin SH, Antonarakis SE, Kazazian HH Jr (1984b) Base substitution at position -88 in a beta-thalassemic globin gene. Further evidence for the role of distal promoter element ACACCC. *J Biol Chem* 259:8679-8681.

Orkin SH, Antonarakis SE, Loukopoulos D (1984) Abnormal processing of beta Knossos RNA. *Blood* 64:311-313.

Orkin SH, Cheng TC, Antonarakis SE, Kazazian HH (1985) Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human beta-globin gene. *EMBO J* 4:453-6.

Orkin SH, Kazazian HH Jr, Antonarakis SE, Goff SC, Boehm CD, Sexton JP, Waber PG, Giardina PJ (1982b) Linkage of beta-thalassaemia mutations and beta-globin gene polymorphisms with DNA polymorphisms in human beta-globin gene cluster. *Nature* 296:627-631.

Orkin SH, Kazazian HH Jr, Antonarakis SE, Ostrer H, Goff SC, Sexton JP (1982) Abnormal RNA processing due to the exon mutation of beta E-globin gene. *Nature* 300:768-769.

Orkin SH, Nathan DG (1998). The thalassemias. In: Nathan DG, Orkin SH, eds. Hematology of Infancy and Childhood. Philadelphia, PA: Saunders pp.811-886.

Perkins AC, Sharpe AH, Orkin SH (1995) Lethal beta-thalassaemia in mice lacking the erythroid CACCC transcription factor EKLF. *Nature* 375:318-322.

Perseu L, Satta S, Moi P, Demartis FR, Manunza L, Sollaino MC, Barella S, Cao A, Galanello R (2011) KLF1 gene mutations cause borderline HbA(2). *Blood* 118:4454-4458.

Ristaldi MS, Cao A, Casula S (1995) *In vitro* expression of the -87 beta-globin thalassemic mutations. *Blood* 86:2055-2056.

Rooks H, Clark B, Best S, Rushton P, Oakley M, Thein OS, Cuthbert AC, Britland A, Ruf A, Thein SL (2012) A novel 506 kb deletion causing εγδβ thalassemia. *Blood Cells Mol Dis* 49:121-127.

Rosatelli C, Leoni GB, Tuveri T, Scalas MT, Di Tucci A, Cao A. (1987) Beta thalassaemia mutations in Sardinians: implications for prenatal diagnosis. *J Med Genet* 24:97-100.

Rosatelli MC, Dozy A, Faà V, Meloni A, Sardu R, Saba L, Kan YW, Cao A (1992) Molecular characterization of beta-thalassemia in the Sardinian population. *Am J Hum Genet* 50:422-426.

Rosatelli MC, Oggiano L, Battista Leoni G, Tuveri T, Di Tucci A, Scalas MT, Dore F, Pistidda P, Massa A, Longinotti M (1989) Thalassemia intermedia resulting from a mild beta-thalassemia mutation. *Blood* 73:601-605.

Rund D, Dowling C, Najjar K, Rachmilewitz EA, Kazazian HH Jr, Oppenheim A (1992) Two mutations in the β-globin polyadenylylation signal reveal extended transcripts and new RNA polyadenylylation sites. *Proc Natl Acad Sci* USA 89:4324-4328.

Sahoo SS, Biswal S, Dixit M (2014) Distinctive mutation spectrum of the HBB gene in an urban eastern Indian population. *Hemoglobin* 38:33-38.

Schokker RC,Went LN, Bok J (1966) A new genetic variant of beta-thalassaemia. *Nature* 209:44-46.

Stamatoyannopoulos G (2005) Control of globin gene expression during development and erythroid differentiation. *Experiment hematol* 33:259.

Stamatoyannopoulos G, Navas PA, Li Q (2009) Molecular and Cellular Basis of Hemoglobin Switching. In Disorders of Hemoglobin, Genetics, Pathophysiology and Clinical Management (eds Steinberg MH, Forget BG, Higgs DR, Weatherall DJ. Second edition), Cambrige University Press pp.86-100.

Stamatoyannopoulos JA, Goodwin A, Joyce T, Lowrey CH (1995) NF-E2 and GATA binding motifs are required for the formation of DNase I hypersensitive site 4 of the human β-globin *locus* control region. *EMBO J* n14:106-116.

Steinberg MH, Forget BG, Higgs DR, Weatherall DJ. (2009). Disorders of Hemoglobin, Genetics, Pathophysiology and Clinical Management. Second edition, *Cambrige University Press*.

Svasti MLS, Hieu TM, Munkongdee T, Winichagoon P, Van Be T, Van Binh T, Fucharoen S (2002) Molecular Analysis of β-Thalassemia in South Vietnam. *Am J Hematol* 71:85-88.

Takeshita K, Forget BG, Scarpa A, Benz EJ Jr (1984) Intranuclear defect in beta-globin mRNA accumulation due to a premature translation termination codon. *Blood* 64:13-22.

Talbot D, Grosveld F (1991) The 5_HS2 of the globin *locus* control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NFE2 binding sites. *EMBO J* 10:1391-1398.

Tewari R, Gillemans N, Wijgerde M, et al. (1998) Erythroid Kruppel like factor (EKLF) is active in primitive and definitive erythroid cells and is required for the function of 5_HS3 of the betaglobin locus control region. *EMBO J* 8:2334-2341.

Thein SL (1998). Beta thalassaemia. In: Rodgers GP, ed. Bailliere's Clinical Haematology, Sickle Cell Disease and Thalassaemia. London: *Bailliere Tindall* pp.91-126

Thein SL (2005) Genetic modifiers of beta thalassemia. *Haematologica* 90:649-660.

Thein SL (2013) The molecular basis of β-thalassemia. *Cold Spring Harb Perspect Med* 3(5):a011700.

Thein SL, Wood WG (2009) The molecular basis of β thalassemia, δβ thalassemia, and hereditary persistence of fetal hemoglobin. In Disorders of hemoglobin: Genetics, pathophysiology, and clinical management (eds. Steinberg MH, Forget BG, Higgs DR, Nagel RL), Cambridge University Press, Cambridge, UK pp. 323-356.

Thom CS, Dickson CF, Gell DA, Weiss MJ (2013) Hemoglobin variants: biochemical properties and clinical correlates. *Cold Spring Harb Perspect Med* 3:a011858.

Tilghman SM, Tiemeier DC, Seidman JG, Peterlin BM, Sullivan M, Maizel JV, Leder P (1978) Intervening sequence of DNA identified in the structural portion of a mouse beta-globin gene. *Proc Natl Acad Sci USA*.;75:725-729.

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* 10:1453-1465.

Tsai SF, Martin DIK, Zon LI, D'Andrea AD, Wong GG, Orkin SH (1989) Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* 339:446-451.

Tuan D, Solomon W, Li Q, London IM. (1985) The "beta-like-globin" gene domain in human erythroid cells. *Proc Natl Acad Sci USA* 82:6384-6388.

Viprakasit V, Harteveld CL, Ayyub H, Stanley JS, Giodano PC, Higgs DR. (2006) A novel deletion causing α-thalassemia clarifies the importance of the major human alpha globin regulatory element. *Blood* 107:3811-3812.

Viprakasit V, Tanphaichitr VS, Pung-Amritt P, Pung-Amritt P, Petrarat S, Suwantol L, Fisher C, Higgs DR (2002) Clinical phenotypes and molecular characterization of Hb H-Pakse disease. *Hematologica 87*:117-125.

Voet D., Voet J.G. (2011), Biochemistry, John Wiley & Sons Inc., 4th Edition.

Wasi P, Na-Nakorn S, Pootrakul S-N (1974). The α thalassemias. *Clin Haematol* 3:383-410.

Waye JS, Nakamura-Garrett LM, Eng B, Kanavakis E, Traeger-Synodinos J (2011) β+-Thalassemia trait due to a novel mutation in the β-globin gene promoter: -26 (A>C) [HBB c.-76A>C]. *Hemoglobin* 35:84-86.

Weatherall D, Akinyanju O, Fucharoen S, Olivieri N, Musgrove P (2006) Inherited Disorders of Hemoglobin in Disease Control Priorities in Developing Countries. 2nd edition. *Washington (DC) World Bank* Chapter 34.

Weatherall DJ, Clegg JB. (2001) The Thalassaemia Syndromes. 4th ed. Oxford: *Blackwell Science*.

Wijgerde M, Gribnau J, Trimborn T, Nuez B, Philipsen S, Grosveld F, Fraser P (1996) The role of EKLF in human beta-globin gene competition. *Genes Dev*10:2894-2902.

Williams TN, Mwangi TW, Wambua S, Peto TE, Weatherall DJ, Gupta S, Recker M, Penman BS, Uyoga S, Macharia A, Mwacharo JK, Snow RW, Marsh K (2005) Negative epistasis between the malaria-protective effects of alpha+-thalassemia and the sickle cell trait. *Nat Genet* 37:1253-1257.

Wong C, Dowling CE, Saiki RK, Higuchi RG, Erlich HA, Kazazian HH J (1987) Characterization of β-thalassaemia mutations using direct genomic sequencing of amplified single copy DNA. *Nature* 330:384-386.

Zhang JZ, Cai SP, He X, Lin HX (1988) Molecular basis of β-thalassemia in South China. Strategy for DNA analysis. *Hum Genet* 78:37-40.

APPENDIX

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the qualification of any other degree or diploma of a university or other institution of higher learning, except where due acknowledgement is made in the acknowledgements.

Name: Nguyen Quynh Chau Sign:

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my advisor **Professor tutor Monica Pirastru** for her tremendous support of my Ph.D study and related research, for her teaching, enthusiasm, patience, motivation, and immense knowledge. In Italy and in Vietnam, she has followed and helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.The joy and enthusiasm she has for research was contagious and motivational for me, even during tough times in the Ph.D. pursuit. I am also thankful for encouraging my research and for allowing me to grow as a research scientist. Knowledge and laboratory skills that I have learned from her has been an important luggage keeping for my work in Vietnam now and in the future.

I cannot find words to express my gratitude to **Professor Piero Cappuccinelli** who provided me an opportunity to study in a foreigner country. He is a great man that I respect a lot.

I owe my deepest gratitude to **Professor Cao Ngoc Thanh, Professor Nguyen Vu Quoc Huy, Professor Vo Tam, Professor Nguyen Khoa Hung** who introduced, supported and agrreed for me to go abroad to study in Italy. The big chance that they gave me is unvaluable to develope my work.

My thanks also goes to **Professor Nguyen Duy Thang**, my co-tutor, for his help and support during the time I performed the research.

It is with sincere gratitude that I acknowledge the support and help of **Professor Bruno Masala**. I have first met him in Vietnam when he had course for Master students and he gave me many useful advices, brilliant comments and suggestions for my thesis proposal and research. When I was working at lab in Italy, he sometimes came and asked about the procedures that really encouraged me much.

I am indebted to **Professor Paolo Mereu** who were involved in the validation survey for this research project. He taught me the molecular techniquique skills with his sincere patience and tolerance.

It gives me great pleasure in acknowledging the support and help of **Professor Laura Manca** who provided me an opportunity to join her team as intern, and who gave access to the laboratory and research facilities. Thank you for always caring my son, you encouraged me much during the time I stayed far from him in Italy.

I would like to express my sincere thanks to **Professor Leonardo Sechi** and **Professor Claudia Crosio** who facilitated my attendance to the PhD program at Sassari University.

With my profound gratitude I also sincerely thank **Dott. Giovanni Sinni** for his professional support. He has helped me a lot during the time I got troubles related to visa to Italy. Without his

help, I was not able to go to Italy to perform this research. Also thanks to **Giustina Casu** for her kind help during 2 years when I first came to Italy.

My sincere thanks also goes to **Dott. Sandro Trova, Dott. Luca De Candia, Valentina Sanna, Stephania Mulas, Sandro** who shared the work at lab with me and helped me for many skills and techniques. Without their precious support it would not be possible to conduct this research.

I would like to express my deep gratitude to **Tran Thi Kim Lien**, **di Hue**, the very kind ladies who helped me to take samples during the time I performed the research in Vietnam even at night and during the rainy days...

My thanks also goes to everybodies at ***Hematology Department*** in Hue, Vietnam for sharing my work and helped me to do the duties when I was in Italy. Thanks to **Dr. Phan Thi Thuy Hoa**, **Dr. Ha Nu Thuy Duong, Dr. Le Phan Minh Triet** for supporting me to perform the research.

Especially thanks to ***Genetic Department***, **Professor Nguyen Viet Nhan**, **Dr. Ha Thi Minh Thi**, **Le Tuan Linh**, **Tran Thi Nhu Nga**, **Doan Huu Nhat Binh**, **Phan Quoc Bao**, **Nguyen Thi Cam Van** for pemitting me to worked at and allowed me to perform CE-HPLC.

I want to send my warm greetings and best wishes to my dear professors, colleages and friends who were besides me and encouraged me during the most difficult time.Thanks to **Dr. Le Thi Dieu Anh, Dr. Truong Phuoc Anh Thu**, **Dr. Phan Tu Khanh Phuong, Dr. Phung Thi Hoang Yen, Dr. Ngo Van Trung, Nguyen Thi Tuyet Tam, Professor Antonella Pantaleo, Elalim Zen Vukovic, Miranda Zen, Fillipo Dossi, Nice, Ivan, Dott. Maria Carmia, Danilo Puddu, Dott. Mauro Rassu, Dott. Gabriele Nieddu, Dott. Simona Sanna, Dott. Silvia Fancello.**

I owe ***Phu Vang District Hospital*** and ***Hospital of Hue University of Medicine and Pharmacy*** for permitting me to collect samples.

I wish the good things will come to **my patients** who participated enthusiastically to the research. This thesis would not have been possible without their contribution. Thank you very much from the bottom of my heart.

Last but not least, I would like to express my great gratitude to my Family, my beloved husband also my colleage **Dr. Nguyen Minh Quang** who gave me the brillant advices for the research, always beside and encourage me in the hardest situation. I want to send my love to my Mother **Phan Thi Bich Dao**, Father **Nguyen Dao Hang**, my Parents in law **Dr. Nguyen Minh Dung**, **Dr. Hoang Thi Thu Huong** and my Sisters **Nguyen Thi Minh Thu**, **Nguyen Thi Minh Hieu**, Sister in law **Nguyen Hoang Minh Thu** for their kindness, tolerance, great support caring my son while I was far from him. Without their help, I was not able to complete the thesis.

Finally I want to say to my son **Bon Bon** that he is my big motivation enabling me to complete the thesis and making my smile present again when I suffer difficulties in life.

Words can not help to express my gratitude to all of you but I want to say I will never forget everything you have done for me.

I should say again without the enormous support of my Professor Tutor, the thesis could not be completed.

Once again, thank you very much!