

β -Thalassaemia Prototype of a Single Gene Disorder with Multiple Phenotypes

Swee Lay Thein

*Department of Haematological Medicine, Guy's, King's & St Thomas' School of Medicine,
Denmark Hill Campus, Bessemer Road, London, UK*

Abstract

As the defective genes for more and more genetic disorders become unravelled, it is clear that patients with the same genotype can have many different clinical conditions even in monogenic disorders. The remarkable phenotypic diversity of the β thalassaemias is prototypical of how the wide spectrum in disease severity can be generated. The most reliable and predictive factor of disease phenotype is the nature of the mutation at the β -globin locus itself. However, relating phenotype to genotype is complicated by the complex interaction of the environment and other genetic factors at the secondary and tertiary levels, some implicated, and others, as yet unidentified. This article reviews the clinical and haematological diversity encountered in β thalassaemia and their relationship with the underlying genotypes.

β Thalassaemia

β thalassaemia belongs to the family of inherited haemoglobin disorders, the commonest monogenic disease. It is characterised by a quantitative deficiency of functional β globin chains. Although it is defined as a reduction in the synthesis of β globin, some forms result from structural haemoglobin variants that are ineffectively synthesised or are so unstable that they result in a functional deficiency of the β chains and a thalassaemia phenotype [1].

1. The β globin gene

β globin is encoded by a structural gene found in a cluster spanning a region of 70 kb on the short arm of chromosome 11 with the other β like genes; (5'- ϵ - γ - δ - β -3') [2]. The general structure of the β globin gene is typical of other globin loci, and consists of three exons (coding regions) interrupted by two non-coding intervening sequences (introns or IVS). Exon 2 encodes the residues involved in haem binding and $\alpha\beta$ dimer formation, while many of the amino acids involved in globin subunit interactions required for the Bohr effect, and 2,3-DPG binding, are found in exon 3.

Sequences important for gene function are found in the 5' promoter region, at the exon-intron junctions, and in the 3' untranslated region (3UTR) at the end of the mRNA sequences. The β globin gene promoter includes 3 positive *cis*-acting elements: TATA box (positions 28 to 31), a CCAAT box (positions 72 to 76) and duplicated CACCC motifs (proximal at positions 86 to 90, and distal at position 101 to 105). While the CCAAT and TATA elements are found in many eukaryotic promoters, the CACCC sequence is found predominantly in erythroid cell-specific promoters. Binding of the Erythroid Krüppel Like Factor (EKLF) to the CACCC motif appears crucial for normal adult β globin expression. The importance of these various 5'-flanking sequences for normal gene expression is underscored by β thalassaemia arising from point mutations in these sequences specifically in and around the TATA box and the CACCC motifs in the 80 to 100 region. An enhancer is also found in intron 2 and 3' of the globin gene, 600 to 900 bp downstream of the poly (A) site.

Upstream of the entire β complex is the β locus control region (β LCR) which consists of five DNase I hypersensitive sites (designated HS 1-5) distributed between 6 to 20 kb 5' of the β globin gene. The LCR

plays a critical role in globin gene expression, it maintains an 'open' globin locus domain and acts as a powerful enhancer of globin gene transcription, in the absence of which the level of gene expression is low. Four of the sites (HS 1-4) are erythroid specific, encompassing binding sequences for erythroid-restricted transcription factors (GATA-1 and NF-E2), while HS5 is ubiquitous.

The developmental regulation of the globin genes reflect their sequential activation in a 5'-3' direction. While the α -like genes undergo a single developmental 'switch' (embryonic→fetal/adult), the β -like genes undergo two 'switches' (embryonic→fetal→adult). Transcription of the ϵ gene in the embryonic yolk gene switches after the 6th week of gestation to the transcription of the two γ genes in the fetal liver, and then around the prenatal period, to that of the δ (minor adult) and β (major adult) genes. At 6 months after birth, Hb F comprises less than 5% of the total haemoglobin and continues to fall reaching the adult level of <1% at 2 years of age. It is at this stage that mutations affecting the β gene become clinically apparent.

The tissue and developmental specific expression of the individual globin genes is governed by the direct physical interactions between the globin promoters and the LCR, the interaction mediated through binding of tissue restricted and ubiquitous transcription factors. However, this precise developmental expression which relies on two mechanisms, gene silencing and gene competition, is not fully understood [3].

2. Pathophysiology and Clinical Diversity of β Thalassaemia

The underlying pathophysiology of β thalassaemia relates to the deficiency of functional β -globin chains, which leads to an imbalanced globin chain production and an excess of α -globin chains [1,4]. The latter aggregate in red cell precursors forming inclusion bodies, causing mechanical damage and their premature destruction in the bone marrow i.e. ineffective erythropoiesis. Red cells that survive to reach the peripheral circulation are prematurely destroyed. Anaemia in β thalassaemia thus results from a combination of ineffective erythropoiesis, peripheral haemolysis, and an overall reduction in haemoglobin synthesis. It is quite clear that the severity of β thalassaemia is directly related to the severity of chain imbalance. Thus factors which reduce the degree of chain imbalance and the magnitude of α chain excess in the red cell precursors, will have an ameliorating effect on the phenotype. At the primary level, this is related directly to the severity of the β thalassaemia mutation itself. At the secondary level, the severity of globin chain imbalance is influenced by variability at two loci: α globin and γ globin genes. Co-inheritance of α thalassaemia reduces chain imbalance with an ameliorating effect, while the presence of extra α globin genes will have an adverse effect. Similarly an inherent capacity for producing γ chain which combine with the excess α to form HbF,

will have an ameliorating effect.

A direct effect of the anaemia is the increased production of erythropoietin which leads to intense proliferation and expansion of the bone marrow with the resulting skeletal deformities. To a large extent these secondary complications of bone disease, splenomegaly, endocrine and cardiac damage can be related to the severity of anaemia and the iron loading that results from the increased gastro-intestinal absorption and the blood transfusions. Recently, however, it has become apparent that these complications of β thalassaemia may be genetically modified by variability at other loci (see below).

The clinical manifestations of β thalassaemia are extremely diverse, spanning a broad spectrum from the transfusion-dependent state of thalassaemia major to the asymptomatic state of thalassaemia trait. The most severe end of the spectrum is characterised by the complete absence of β globin production and results from the inheritance of two β^0 thalassaemia alleles, homozygous or compound heterozygous states. This condition is referred to as β thalassaemia major and, at their worst, the patients present within 6 months of life, and if not treated with regular blood transfusions, die within their first two years. Conversely, many patients who have inherited two β thalassaemia alleles have a milder disease, ranging from a condition that is only slightly less severe than transfusion-dependence through a spectrum of decreasing severity to one that is asymptomatic and often mistaken as β thalassaemia trait. This diverse collection of phenotypes between the two extremes of thalassaemia major and trait, constitute the clinical syndrome of thalassaemia intermedia. The underlying genotypes are equally heterogeneous, resulting from the interaction of other genetic variables with the inheritance of one or two β thalassaemia alleles.

β thalassaemia trait comprises the other end of the phenotypic spectrum of β thalassaemia. It is usually associated with the inheritance of a single β thalassaemia allele, manifested by a mild anaemia with hypochromic microcytic red blood cells, elevated levels of HbA₂ and variable levels of Hb F. However, even the heterozygous states for β thalassaemia show a phenotypic diversity comparable to that of thalassaemia major. In some cases, the β thalassaemia allele can be phenotypically 'silent', with no anaemia or any haematological abnormalities. In others, the heterozygous state causes a phenotype almost as severe as the major forms, that is, the β thalassaemia allele is dominantly inherited.

Although definition of the two extremes of the clinical spectrum of β thalassaemia is easy, assigning the severity of the intermediate form can be problematical. Criteria such as age and level of haemoglobin at presentation, transfusion history, and the requirements for intermittent blood transfusion have been used, but these have their inherent limitations and are highly clinician dependent.

3. Mechanisms Underlying Phenotypic Diversity

Progress in our understanding of the mechanisms underlying the remarkable phenotypic variability of β thalassaemia has been made possible by a combination of the analysis of the molecular basis of the different forms of thalassaemia, family studies and analysis of the genotype/phenotype relationship of the thalassaemia intermediates. Not surprisingly, heterogeneity of the mutations at the β globin locus itself account for a major part of the variability but phenotypic variability can also arise at the secondary level from interactions with other genetic loci, the α - and γ -globin genes, that affect the degree of globin chain imbalance. Finally, at the tertiary level, complications of the disease can be genetically modified by interactions with loci not involving globin chain balance.

3.1. Heterogeneity and Variable Severity of β thalassaemia Alleles

The most common forms of β thalassaemia alleles are those that are prevalent in the Mediterranean, tropical and sub-tropical regions including the Middle East, parts of Africa, Indian subcontinent and South East Asia, but they are by no means confined to those regions [5]. Due to recent population movements, the β thalassaemias have become an important part of clinical practice world-wide. With the exception of a few deletions, the vast majority of β thalassaemia are caused by point mutations within the gene or its immediate flanking sequences. A few β thalassaemia mutations which segregate independently of the β globin cluster have been described in several families [6]; in such cases *trans*-acting regulatory factors are presumably involved [7]. Although more than 200 β thalassaemia alleles have been characterised [8], population studies indicate that probably only 20 account for >80% of the β thalassaemia mutations in the world. This is because in most of the high-frequency areas, only a few [4-6] mutations are common with a varying number of rare ones, and each of these populations has its own spectrum of β thalassaemia alleles [5].

3.1.1. β^0 vs β^+ and β^{++} Thalassaemia Alleles

Functionally the β thalassaemia alleles can be classified as β^0 or β^+ reflecting the resulting phenotype: β^0 thalassaemia in which there is a complete absence of β globin production and the most severe possible, and β^+ thalassaemia in which there is some, although reduced β globin product.

Deletions causing β thalassaemia are rare and result in a complete absence of β globin product. They can be classified as those affecting only the structural β globin gene and the upstream deletions of which 3 have been described. The latter remove all or part of the β -LCR but leave the β gene itself intact, and yet down-regulate the β gene as part of egdb thalassaemia. Four-

teen deletions which involve the structural β globin gene alone, have been described; only the 619 bp deletion at the 3' end of the β gene is common, but even that is restricted to the Sind and Punjab populations of India and Pakistan where it accounts for ~20% of the β thalassaemia alleles. The other deletions, although extremely rare, are of particular functional and phenotypic interest because they are associated with an unusually high level of HbA₂ in heterozygotes. These deletions differ widely in size, but remove in common a region (from positions 125 to +78 relative to the mRNA cap site) in the promoter which includes the CACCC, CCAAT and TATA elements. The mechanism underlying the markedly elevated levels of Hb A₂ and the variable increase in Hb F in heterozygotes for these deletions appears to be related to the removal of the 5' promoter region of the β gene. This removes competition for the upstream LCR leading to its increased interaction with the γ and δ genes in *cis*, enhancing their expression. Although the increases in Hb F are variable, and moderate in heterozygotes, they are adequate to compensate for the complete absence of β globin in homozygotes for these deletions. This mechanism may also explain the unusually high Hb A₂ levels which accompany point mutations affecting the promoter regions.

Transposable elements may occasionally disrupt human genes and result in their inactivation. The insertion of such an element, a retrotransposon of the family called L1 has been reported with the phenotype of β^+ thalassaemia. Despite the insertion of 6-7 kb DNA into its IVS2, the affected gene expresses full length β globin transcripts at a level corresponding to about 15% of normal β globin mRNA [9].

The vast majority of the β thalassaemia alleles result from single base substitutions, minor insertions or deletions of a few bases within the gene or its immediate flanking sequences. They may affect any level of gene expression and are classified according to the mechanism by which they affect gene function: transcription, RNA processing or RNA translation.

Mutations affecting transcription can either involve the conserved DNA sequences that form the β globin promoter or the stretch of 50 nucleotides in the 5'UTR. Generally they result in a mild to minimal deficit of β globin output reflecting the relatively mild phenotype of these β^+ thalassaemias. The C-T mutation at position 101 to the β globin gene appears to cause an extremely mild deficit of β globin such that it is 'silent' in heterozygotes who have normal Hb A₂ levels and normal red cell indices. Several mutations in the 5'UTR e.g. CAP+1A-C, also have a 'silent' phenotype.

Mutations that affect RNA processing can involve either of the invariant dinucleotides (GT at 5' and AG at 3') in the splice junction in which case normal splicing is completely abolished with the resulting phenotype of β^0 thalassaemia. Mutations within the consensus sequences at the splice junctions reduce the efficiency of normal splicing to varying degrees and produce a β^+ phenotype that ranges from mild to severe.

Other splicing mutations involve base substitutions with introns or exons. For example, a cryptic splice site which contains the sequence GT GGT GAG G has been found in exon 1 of the β globin gene, spanning codons 24 to 27. Three mutations within this region activate this cryptic site which acts as an alternative donor site in RNA processing. The mutations in codon 26 (GAC \rightarrow AAE) that gives rise to Hb E (β 26 Gln \rightarrow Lys) also activates this cryptic splice site causing abnormal mRNA processing such that normal splicing which produces Hb E variant is reduced. Since Hb E production is also quantitatively reduced, the compound heterozygous state, Hb E/ β thalassaemia results in a clinical picture closely resembling homozygous β thalassaemia ranging from severe anaemia and transfusion-dependency to thalassaemia intermedia. Other RNA processing mutants affect the polyadenylation signal (AATAAA) and the 3'UTR. These are generally mild β^+ thalassaemia alleles.

Mutations which are expressed at the level of mRNA translation involve either the initiation or extension phases of globin synthesis and are all associated with a β^0 phenotype. Approximately half of the β thalassaemia alleles are characterised by premature termination of β chain extension. They result from the introduction of premature termination codons due to frameshifts or nonsense mutations and nearly all terminate within exon 1 and 2. Mutations that result in premature termination early in the sequence (in exons 1 and 2) are associated with minimal steady-state levels of β mRNA in erythroid cells, due to an accelerated decay of the abnormal mRNA referred to nonsense-mediated mRNA decay (NMD) [10]. In heterozygotes for such cases, no β chain is produced from the mutant allele and only half the normal β globin is present, resulting in a typical asymptomatic phenotype. By contrast, mutations that produce in-phase termination later in the sequence, in exon 3 are not subjected to NMD resulting in substantial amounts of abnormal β mRNA comparable to that of the normal allele [11]. The abnormal mRNA is presumably translated into variant β chains causing a dominant negative phenotype. Hence, these mutants are usually dominantly inherited (see later).

The variable severity of the different β thalassaemia alleles is reflected in their phenotypic effect in heterozygotes, in the degree of hypochromia and microcytosis as indicated by the mean cell haemoglobin (MCH) and mean cell volume (MCV) values respectively. Rund et al [12] showed that the β^0 thalassaemia alleles which are associated with the most severe phenotype, demonstrated a fairly tight range of MCVs (63.1 fl, SD = 3.4) while the β^+ alleles were associated with a wider range of MCVs (69.3 fl, SD = 5.6). The cut-off point between the β^0 and β^+ thalassaemias was 67 fl. The broader range of MCV in β^+ thalassaemia when compared to β^0 thalassaemia, is not surprising given the broad range in the deficit of β globin production, from barely detectable levels at the severe end, to just a little less than normal in the very mild or 'silent' alleles. The 'silent' mutations are normally identified in the com-

ound heterozygous states with a severe β thalassaemia allele which results in thalassaemia intermedia, or in homozygotes who have a typical phenotype of β thalassaemia trait. The 'silent' β thalassaemia alleles are not common, except for the 101 C-T which accounts for a large number of the milder forms of β thalassaemia in the Mediterranean [13].

The mild β thalassaemia alleles are associated with clearly defined changes in heterozygotes and result in disorders of intermediate severity in homozygotes. Interactions with the severe alleles are less predictable due to the wider range of β globin output, and extends from transfusion dependence to intermediate forms of β thalassaemia at the mild end of the spectrum [14,15].

3.1.2. Offsetting Effect of Hb F Due to Nature of β Thalassaemia Alleles

Although some of the phenotypic variability of β thalassaemia can be explained by the differing severity of the β thalassaemia alleles, it does not explain why identical mutations in different ethnic groups sometimes produce a different phenotype or why individuals with β^0 thalassaemia deletions have milder disease despite the complete absence of β chain production. Both situations can be explained by an inherent propensity to produce Hb F for different reasons. Although a given mutation is generally found within one ethnic group, a number of identical mutations have been described in different racial groups. In these cases the mutations may have arisen on different β chromosomal backgrounds, some of which contain the common genetic variant, C-T at position 158 of the ν globin gene, also referred to as the *Xmn1-G ν* polymorphism. Although the *Xmn1-G ν* polymorphism has little effect in normal individuals, under haemopoietic stress, the presence of this site can have the effect of increasing Hb F levels resulting in a milder disease in homozygous β thalassaemia [16]. The increased Hb F output observed in deletions or mutations that involve the promoter sequence of the β globin gene reflect the competition between the ν - and β -globin gene promoters for interaction with the LCR or rate-limiting transcription factors. Hence, although such deletions cause a complete absence of β globin product, the severity of the phenotype is offset by the concomitant increase in haemoglobin F [17].

3.1.3. Dominantly Inherited β Thalassaemia

The common β thalassaemia alleles that are prevalent in the malarious regions, are inherited typically as Mendelian recessives; heterozygotes are clinically asymptomatic and the inheritance of two mutant alleles as homozygotes or compound heterozygotes is required to produce clinical disease. Some forms of β thalassaemia, however are dominantly inherited, in that inheritance of a single β thalassaemia allele in the presence of a normal α globin genotype, results in a clinically detectable disease. Heterozygotes have a thalassaemia intermedia phenotype with moderate anaemia, splenomegaly

and a thalassaemic blood picture. Apart from the usual features of heterozygous β thalassaemia, such as increased levels of HbA₂ and the imbalanced α/β globin biosynthesis, large inclusion bodies similar to those seen in thalassaemia major, are often observed in the red cell precursors, hence the original term of 'inclusion body β thalassaemia'.

This unusual form of β thalassaemia was probably first described in an Irish family in 1973; several members of the family spanning three generations had a thalassaemia intermedia phenotype that was clearly inherited as a Mendelian dominant. The molecular basis was subsequently shown to involve two deletions of 4 and 11 bp in exon 3, interrupted by an insertion of 5 bp [18] which gave rise to a frameshift and the predicted synthesis of an elongated β chain variant with an abnormal carboxy terminal end. Since the first description, more than 30 dominantly inherited β thalassaemia alleles have now been described [19,20]; they include missense mutations, minor deletions leading to the loss of intact codons, frameshifts arising from minor insertions and deletions resulting in elongated β variants with abnormal carboxy terminal ends, and truncated β variants resulting from nonsense mutations. The common denominator of these mutations is the predicted synthesis of highly unstable β chain variants, so unstable that in many cases, they are not detectable and only implicated from the DNA sequence. The predicted synthesis is supported by the presence of substantial amounts of abnormal β mRNA in the peripheral reticulocytes, comparable in amounts to that produced from the normal β allele. Indeed, the large intra-erythroblastic inclusions, that are so characteristic of this form of thalassaemia, have subsequently been shown to be composed of both α - and β -globin chains [21]. In contrast, the inclusion bodies in homozygous β thalassaemia consisted only of precipitated α -globin.

In contrast to the in-phase termination mutations that are recessively inherited, and which lead to termination in exon 1 or 2, those that are dominantly inherited, terminate much later in the sequence of the β globin gene, in exon 3 or beyond. The differential effect of these in-phase termination mutants on the accumulation of mutant mRNA can be explained by a failure of the surveillance mechanism of the Nonsense Mediated Decay (NMD) pathway in the case of the exon 3 mutations. Premature stop codons near the 3' end of the gene, in exon 3 of the β gene, are less likely to trigger NMD leading to an accumulation of the mutant β mRNA leading to the synthesis of truncated β chain variants [22]. These in-phase termination mutations exemplify how shifting the position of a nonsense codon can alter the phenotype of recessive inheritance caused by haplo-insufficiency, to a dominant negative effect due to the synthesis of an abnormal and deleterious protein.

The hyperinstability of these β chain variants relate to the nature and position of the mutations [23], and can be attributed to several mechanisms such as: substitution of the critical amino acids in the hydrophobic heme pocket displacing heme leading to aggre-

gation of the globin variant; disruption of secondary structure due to replacement of critical amino acids; substitution or deletion of amino acids involved in $\alpha\beta$ dimer formation; and elongation of sub-units by a hydrophobic tail [24]. Again, a spectrum in phenotypic severity of this class of β thalassaemia variants is observed, that can be related to variation in the degree of instability of the β globin products. The highly unstable products precipitate in the red cell precursors causing an ineffective erythropoiesis, less unstable variants are able to combine with the α globin sub-units to form haemoglobin tetramers which subsequently precipitate in the peripheral circulation. In this case, red cell production is relatively normal and the features are those of a haemolytic anaemia with inclusion bodies, also known as Heinz Body haemolytic anaemia.

Thus, factors which determine the pathophysiology of mutations involving the β globin gene in the heterozygous state appear to include: whether a β chain variant is synthesised, the stability of the product, its ability to bind haem and to form α/β dimers and tetramers. At the one end of the spectrum are typical β thalassaemia traits due to a 'pure' quantitative defect characterised by a very mild anaemia and hypochromic microcytic red cells, while at the other, there are the predominantly haemolytic anaemias due to instability of haemoglobin tetramers. The dominantly inherited β thalassaemia characterised by the synthesis of highly unstable β chain variants, fall in between the two extremes. They resemble the intermediate forms of β thalassaemia by virtue of the ineffective erythropoiesis, but also have elements of congenital haemolytic anaemias in that there is a variable degree of haemolysis.

Unlike recessive β thalassaemia which is prevalent in malaria-endemic regions, dominant β thalassaemias are rare, occurring in dispersed geographical regions where the gene frequency for β thalassaemia is very low. The vast majority of the dominant β thalassaemia alleles have been described in single families, many as *de novo* events. It is likely that the low frequency of the dominant β thalassaemia alleles is due to the lack of positive selection that occurs in the recessive forms. Clinically, since spontaneous mutations are common in dominant β thalassaemia, it is important that the disorder should be suspected in any patient with a thalassaemia intermedia phenotype even if both parents are haematologically normal and the patient is from an ethnic background where β thalassaemia is rare.

3.2. Secondary Modifiers

The severity of anaemia in β thalassaemia reflects the degree of globin chain imbalance and the excess of α globin chains with all their deleterious effects on the red cell precursors. This globin chain imbalance can be genetically modified by two factors variation in the amount of a globin production and variation in fetal haemoglobin response

3.2.1. α Globin Genotype

In many populations in which β thalassaemia is prevalent, athalassaemia also occurs at a high frequency and hence it is not uncommon to co-inherit both conditions. Homozygotes or compound heterozygotes for β thalassaemia who co-inherit α thalassaemia will have less redundant α globin and tend to have a less severe condition. As with β thalassaemia, the different α thalassaemias which predominate in different racial groups display a wide range of severity. This interaction alone provides the basis for considerable clinical heterogeneity; the degree of amelioration depends on the severity of the β thalassaemia alleles and the number of functional α -globin genes [14,15]. Co-inheritance of a single α gene deletion has very little effect on the phenotype of β^0 thalassaemia while individuals with two α gene deletions and homozygous β^+ thalassaemia may have a mild form of thalassaemia intermedia. At the other extreme, patients who have co-inherited HbH and homozygous β thalassaemia also have thalassaemia intermedia.

Just as co-inheritance of α thalassaemia can reduce the clinical severity of homozygous β thalassaemia, the presence of increased α globin product in β thalassaemia heterozygotes who are normally clinically asymptomatic, tips the globin chain imbalance further, resulting in a thalassaemia intermedia phenotype. In the majority of cases, this is related to the co-inheritance of triplicated β globin genes. Triplicated α genes ($\alpha\alpha\alpha$) occur in most populations at a low frequency. The co-inheritance of two extra α globin genes ($\alpha\alpha\alpha\alpha$) or ($\alpha\alpha\alpha/\alpha$) with heterozygous β thalassaemia results in the thalassaemia intermedia. However, the phenotype of a single extra α gene ($\alpha\alpha\alpha\alpha$) with heterozygous β thalassaemia is more variable and depends on the severity of the β thalassaemia allele [25,26]. There appears to be a critical threshold of globin chain imbalance in each individual above which clinical symptoms appear. This may be related to the efficiency of the proteolytic mechanism of the erythroid precursors.

3.2.2. Variation in Fetal Haemoglobin Production

The role of increased Hb F response as an ameliorating factor becomes evident in the group of homozygous β^0 thalassaemia patients who have a mild disease and are able to maintain a reasonable level of haemoglobin all of which is Hb F. Production of fetal haemoglobin after the neonatal period in β thalassaemia is an extremely complex process and still poorly understood. There appears to be a genuine increase in γ chain synthesis, presumably reflecting the expansion of the ineffective erythroid mass. The effect is augmented by the selective survival of the erythroid precursors that synthesise relatively more γ chains. Hence all β thalassaemias, heterozygous or homozygous, have variable increases in their levels of Hb F. Against this background, there are undoubtedly genetic factors involved.

Recent studies have shown that the level of Hb F and F cells (sub-set of erythrocytes that contain Hb F) are overwhelmingly genetically controlled [27]. About one-third of the genetic variance is due to determinants linked to the β globin gene complex but more than 50% of the genetic variance in F cell levels is due to factors not linked to the chromosome [28]. These *trans*-acting factors presumably play an important role in the fine tuning of γ globin production in adult life.

There are several determinants within the β globin gene cluster that are associated with increased Hb F levels in adult life. The conditions that constitute the group of pancellular hereditary persistence of fetal haemoglobin (HPFH) are clearly defined by increases of 5-30% Hb F in heterozygotes [29]. These are caused by large deletions of the β globin complex or point mutations in the γ globin promoters and are clearly inherited as alleles of the β globin complex in a Mendelian fashion. These variants, however, are rare. Much more common is a genetic variant, C-T polymorphism, at position -158 of the $\epsilon\gamma$ globin gene, also referred to as *Xmn1-G γ* polymorphism. The *Xmn1-G γ* site is present at a frequency of 0.37 and our linkage studies indicate that it accounts for up to 30% of the F cell variance in the general populations [28]. Although the increases in Hb F and F cells are minimal in normal people, clinical studies have shown that, under conditions of haemopoietic stress, for example in homozygous β thalassaemia and sickle cell disease, presence of the *Xmn1-G γ* site favour a higher Hb F response. This could explain why the same mutations on different β chromosomal backgrounds (some with and others without the *Xmn1-G γ* site) are associated with different clinical severity.

Although presence of the *cis Xmn1-G γ* site is a modulating factor, clearly there are some patients who have enhanced Hb F response despite being *Xmn1-G γ* -/- [14,30]. In many cases, family studies have shown that there is an inherent capacity for producing HbF and that the genetic determinant is not linked to the β globin cluster. This is in keeping with our sib-pair studies which showed that >50% of the F cell variance in the general population is accounted for by *trans*-acting factors. Indeed, analysis of a single large family spanning seven generations has assigned one such quantitative trait locus (QTL) for F cell to chromosome 6q (31). Analyses of similar families indicate that there are other QTLs for Hb F and F cells that are not linked to 6q or the β globin gene complex [32]. A genetic determinant which is associated with F Cell variance in sickle cell disease has been assigned to chromosome Xp [33] but its role, if any, in determining the level of Hb F in β thalassaemia is not clear. Recently, using a two locus-genetic model, another QTL for F cell levels has been assigned to chromosome 8q, the effects of this locus are conditional on the *Xmn1-G γ* site [34]. As the genetic basis of the propensity to produce Hb F becomes unravelled it is becoming clear that the conglomeration of the *Xmn1-G γ* polymorphism, the QTLs on 6q, Xp and 8q and others, linked and unlinked to the

β globin complex, constitute the loosely defined syndrome of heterocellular HPFH [35]. Until the different entities become better defined, detection of an inherent capacity for increased Hb F production is, at present, difficult and usually inferred from family studies.

3.2.3. Mosaicism Due to Somatic Deletion of β Globin Gene

This novel mechanism was recently described in an individual who had moderately severe thalassaemia intermedia despite being constitutionally heterozygous for β^0 thalassaemia with a normal α genotype [36]. Subsequent investigations revealed that he had a somatic deletion of a region of chromosome 11p15 including the β globin complex giving rise to a mosaic of cells, 50% with one and 50% without any β globin gene. The sum total of the β globin product is ~25% less than the normally asymptomatic β thalassaemia trait.

This unusual case once again illustrates that the severity of anaemia of β thalassaemia reflects the defective β globin chain production. Furthermore, with respect to potential gene therapy, expression of a single β globin gene in a proportion of the red blood cells appears to be sufficient to redress the chain imbalance to produce a condition mild enough not to need major medical intervention.

3.3. Tertiary Modifiers

With the increasing lifespan of the β thalassaemia patients, subtle variations in the phenotype with regard to some of the complications have become apparent and evidence suggests that they may be affected by genetic variants.

Hyperbilirubinaemia and a propensity to gallstone formation is a common complication of β thalassaemia and is attributed to the rapid turn-over of the red blood cells, bilirubin being a break-down product of haemoglobin. However, varying degrees of jaundice have often been observed in the thalassaemia syndromes from thalassaemia trait through to thalassaemia major [37-39]. Studies have shown that the levels of bilirubin and the incidence of gallstones in β thalassaemia is related to a polymorphic variant (seven TA repeats) in the promoter of the Uridine Diphosphate-Glucosyltransferase IA (UGT1A) gene. Normal individuals who are homozygous for the [TA]₇ variant instead of the usual six, tend to have higher levels of bilirubin (Gilbert's syndrome) [40]. The [TA]₇ variant has also been shown to be associated with increased bilirubin levels in sickle cell disease and other haemolytic anaemias [41]. *In-vitro* studies indicate the variant causes a reduced expression of the UGT1A gene [40].

A common complication of β thalassaemia involve organ damage from iron overload, not just from blood transfusions but also from increased absorption. Preliminary studies suggest that the common mutations C282Y in the *HFE* gene that cause hereditary haemochromatosis, predisposes to iron loading in thalassaemia inter-

media [42]. The co-existence of β thalassaemia trait aggravates and accentuates iron loading in C282Y *HFE* homozygotes [43]. Since the C282Y mutation is rare in populations in which β thalassaemia is common it has a limited role in iron loading amongst these patients [44]. Much more common in the *HFE* gene polymorphism, H63D, whose functional role is still being investigated. As other genes in iron homeostasis become uncovered, it is likely there will be genetic variants in these loci that influence the different degrees of iron loading in β thalassaemia [45].

Similarly, there is increasing evidence that progressive osteoporosis and osteopaenia, another increasingly common complication encountered in young adults with β thalassaemia [46] may be modified by polymorphisms in the genes for the vitaminD and oestrogen receptor, and the *COL1A1* gene that regulates synthesis of type 1 collagen [47]. Genetic variants implicated in other complications of β thalassaemia include: the apolipoprotein E (APOE) ϵ 4 allele in cardiac damage [48]); specific HLA alleles in the tendency to hepatitis and liver cirrhosis; genetic variants in Factor V, prothrombin and MTHFR, and the tendency to thrombosis.

4. Conclusions

There is a spectrum of phenotypes in β thalassaemia, the severity of which relates directly to the degree of chain imbalance and the α globin excess (See Figure 1). Much of the variation can be explained by heterogeneity of the molecular lesions affecting the β globin gene itself *but* it is also clear that variability at the two loci - α and ν globin genes- is important in determining the phenotype, which is extremely encouraging for genetic counselling. However, while genotyping at the β globin and ν globin loci is relative easy to incorporate into the prenatal diagnosis and counselling

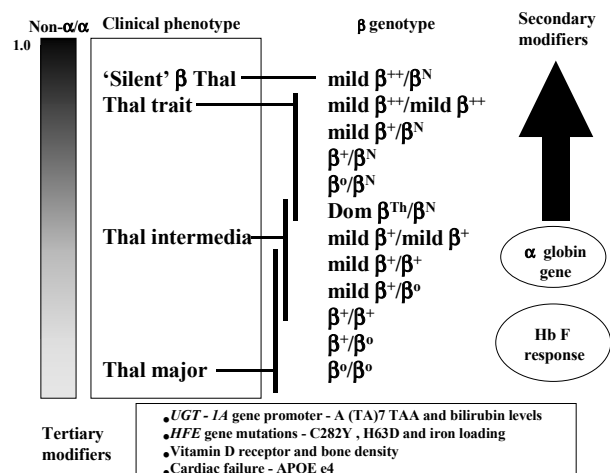


Figure 1. Summary of the Phenotype/Genotype Relationship in thalassaemia

programme, detecting an inherent ability to increase Hb F in response to haemopoietic stress, a major ameliorating factor of β thalassaemia, is still difficult. The presence of such heterocellular HPFH determinants is usually implicated from studies of family members who are often not available. Until the quantitative trait loci for Hb F are better defined, it would appear that it is still not possible to consistently predict phenotype from genotype apart from the two categories of extra α globin genes with heterozygous β thalassaemia, and the inheritance of mild β^+ thalassaemia alleles.

References

- Weatherall DJ, Clegg JB, editors. The Thalassaemia Syndromes. 4th ed. Oxford: Blackwell Science; 2001.
- Forget BG. Molecular Genetics of the Human Globin Genes. In: Steinberg MH, Forget BG, Higgs DR, Nagel RL, editors. Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management. Cambridge, UK: Cambridge University Press; 2001. p.117-130.
- Stamatoyannopoulos G. Molecular and Cellular Basis of Hemoglobin Switching. In: Steinberg MH, Forget BG, Higgs DR, Nagel RL, editors. Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management. Cambridge, UK: Cambridge University Press; 2001. p.131-145.
- Schrier SL. Pathophysiology of thalassaemia. *Curr Opin Hematol.* 2002;9:123-126.
- Flint J, Harding RM, Boyce AJ, Clegg JB. The population genetics of the haemoglobinopathies. In: Rodgers GP, editor. Baillière's Clinical Haematology. London: Bailliere Tindall; 1998. p.1-52.
- Thein SL. Baillière's Clinical Haematology: Beta thalassaemia in sickle cell disease and thalassaemia. In: Rodgers GP, editor. Sickle Cell Disease and Thalassaemia. London: Bailliere Tindall; 1998. p.91-126.
- Viprakasit V, Gibbons RJ, Broughton BC, et al. Mutations in the general transcription factor TFIID result in beta-thalassaemia in individuals with trichothiodystrophy. *Hum Mol Genet.* 2001;10:2797-2802.
- Huisman THJ, Carver MFH, Efremov GD. A Syllabus of Human Hemoglobin Variants. 2nd ed. Augusta, GA, USA: The Sickle Cell Anemia Foundation; 1998.
- Divoky V, Indrak K, Mrug M, Brabec V, Huisman THJ, Prchal JT. A novel mechanism of β thalassaemia: The insertion of L1 retrotransposable element into β globin IVS II. *Blood.* 1996;88:148a.
- Maquat LE. When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. *RNA.* 1995;1:453-465.
- Maquat LE, Carmichael GG. Quality control of mRNA function. *Cell.* 2001;104:173-176.
- Rund D, Filon D, Strauss N, Rachmilewitz EA, Oppenheim A. Mean corpuscular volume of heterozygotes for β -thalassaemia correlates with the severity of mutations. *Blood.* 1991;79:238-243.
- Maragoudaki E, Kanavakis E, Trager-Synodinos J, et al. Molecular, haematological and clinical studies of the -101 C->T substitution in the β -globin gene promoter in 25 β -thalassaemia intermedia patients and 45 heterozygotes. *Br J Haematol.* 1999;107:699-706.
- Ho PJ, Hall GW, Luo LY, Weatherall DJ, Thein SL. Beta thalassaemia intermedia: is it possible to consistently predict phenotype from genotype? *Br J Haematol.* 1998;100:70-78.
- Camaschella C, Maza U, Roetto A, et al. Genetic interactions in thalassaemia intermedia: analysis of β -mutations, α -genotype, γ -promoters, and β -LCR hypersensitive sites 2 and 4 in Italian patients. *Am J Hematol.* 1995;48:82-87.
- Thein SL, Hesketh C, Wallace RB, Weatherall DJ. The molecular basis of thalassaemia major and thalassaemia intermedia in Asian Indians: application to prenatal diagnosis. *Br J Haematol.* 1988;70:225-231.
- Craig JE, Kelly SJ, Barnetson R, Thein SL. Molecular characterization of a novel 10.3 kb deletion causing β -thalassaemia with unusually high Hb A₂. *Br J Haematol.* 1992;82:735-744.
- Thein SL, Hesketh C, Taylor P, et al. Molecular basis for dominantly inherited inclusion body β -thalassaemia. Proceedings of the National Academy of Sciences. USA. 1990;87:3924-3928.
- Thein SL. Dominant β thalassaemia: molecular basis and pathophysiology. *Br J Haematol.* 1992;80:273-277.
- Thein SL. Is it dominantly inherited β thalassaemia or just a β -chain variant that is highly unstable? *Br J Haematol.* 1999;107:12-21.
- Ho PJ, Wickramasinghe SN, Rees DC, Lee MJ, Eden A, Thein SL. Erythroblastic inclusions in dominantly inherited β thalassaemias. *Blood.* 1997;89:322-328.
- Hentze MW, Kulozik AE. A perfect message: RNA surveillance and nonsense-mediated decay. *Cell.* 1999;96:307-310.
- Thein SL. Structural Variants with a β -Thalassaemia Phenotype. In: Steinberg MH, Forget BG, Higgs DR, Nagel RL, editors. Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management. Cambridge, UK: Cambridge University Press, Cambridge, UK; 2001. p.342-355.
- Bunn HF, Forget BG. Hemoglobin: Molecular, Genetic and Clinical Aspects. Philadelphia, PA: W.B. Saunders Company; 1986.
- Camaschella C, Kattamis AC, Petroni D, et al. Different hematological phenotypes caused by the interaction of triplicated α -globin genes and heterozygous β -thalassaemia. *Am J Hematol.* 1997;55:83-88.
- Traeger-Synodinos J, Kanavakis E, Vrettou C, Maragoudaki E, Michael T, Metaxotou-Mavromati A. The triplicated α -globin gene locus in β -thalassaemia heterozygotes: clinical, haematological, biosynthetic and molecular studies. *Br J Haematol.* 1996;95:467-471.
- Garner C, Tatu T, Reittie JE, et al. Genetic influences on F cells and other hematological variables: a twin heritability study. *Blood.* 2000;95:342-346.
- Garner C, Tatu T, Game L, et al. A candidate gene study of F cell levels in sibling pairs using a joint linkage and association analysis. *GeneScreen.* 2000;1:9-14.
- Wood WG. Hereditary Persistence of Fetal Hemoglobin and $\delta\beta$ Thalassaemia. In: Steinberg MH, Forget BG, Higgs DR, Nagel RL, editors. Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management. Cambridge, UK: Cambridge University Press, Cambridge, UK; 2001. p.356-388.
- Galanello R, Dessi E, Melis MA, et al. Molecular analysis of β^0 -thalassaemia intermedia in Sardinia. *Blood.* 1989;74:823-827.
- Craig JE, Rochette J, Fisher CA, et al. Dissecting the loci controlling fetal haemoglobin production on chromosomes 11p and 6q by the regressive approach. *Nature Genetics.* 1996;12:58-64.
- Craig JE, Rochette J, Sampietro M, et al. Genetic heterogeneity in heterocellular hereditary persistence of fetal hemoglobin. *Blood.* 1997;90:428-434.
- Dover GJ, Smith KD, Chang YC, et al. Fetal hemoglobin levels in sickle cell disease and normal individuals are partially controlled by an X-linked gene located at Xp22.2. *Blood.* 1992;80:816-824.
- Garner CP, Tatu T, Best S, Creary L, Thein SL. Evidence for Genetic Interaction between the beta-globin complex and chromosome 8q in the expression of fetal hemoglobin. *Am J Hum Genet.* 2002;70:793-799.
- Thein SL, Craig JE. Genetics of Hb F/F cell variance in adults and heterocellular hereditary persistence of fetal hemoglobin. *Hemoglobin.* 1998;22:401-414.

36. Badens C, Mattei MG, Imbert AM, et al. A novel mechanism for thalassaemia intermedia. *The Lancet*. 2002;359:132-133.
37. Galanello R, Perseu L, Melis MA, et al. Hyperbilirubinaemia in heterozygous β -thalassaemia is related to co-inherited Gilbert's syndrome. *Br J Haematol*. 1997;99:433-436.
38. Galanello R, Piras S, Barella S, et al. Cholelithiasis and Gilbert's syndrome in homozygous β -thalassaemia. *Br J Haematol*. 2001;115:926-928.
39. Sampietro M, Lupica L, Perrero L, Comino A, Martinez di Montemuros F. The expression of uridine diphosphate glucuronosyltransferase gene is a major determinant of bilirubin level in heterozygous β -thalassaemia and in glucose-6-phosphate. *Br J Haematol*. 1997;99:437-439.
40. Bosma PJ, Chowdhury JR, Bakker C, et al. The genetic basis of the reduced expression of bilirubin UCP-glucuronosyltransferase 1 in Gilbert's syndrome. *New England J Medicine*. 1995;333:1171-1175.
41. Passon RG, Howard TA, Zimmerman SA, Schultz WH, Ware RE. Influence of Bilirubin Uridine Diphosphate- Glucuronosyltransferase 1A Promoter Polymorphisms on Serum Bilirubin Levels and Cholelithiasis in Children With Sickle Cell Anemia. *Am J Pediatr Hematol Oncol*. 2001;23:448-451.
42. Rees DC, Luo LY, Thein SL, Sing BM, Wickramasinghe S. Nontransfusional iron overload in thalassaemia: Association with hereditary hemochromatosis. *Blood*. 1997;90:3234-3236.
43. Piperno A, Mariani R, Arosio C, et al. Haemochromatosis in patients with beta-thalassaemia trait. *Br J Haematol*. 2000;111:908-914.
44. Merryweather-Clarke AT, Pointon JJ, Shearman JD, Robson KJH. Global prevalence of putative haemochromatosis mutations. *J Medical Genetics*. 1997;34:275-278.
45. Andrews N. Iron homeostasis: insights from genetics and animal models. *Nature Reviews Genetics*. 2000;1:208-216.
46. Wonke B. Bone disease in β -thalassaemia major. *Br J Haematol*. 1998;103:897-901.
47. Dresner Pollack R, Rachmilewitz E, Blumenfeld A, Idelson M, Goldfarb AW. Bone mineral metabolism in adults with beta-thalassaemia major and intermedia. *Br J Haematol*. 2000;111:902-907.
48. Economou-Peterson E, Aesspopos A, Kladi A, et al. Apolipoprotein E ϵ 4 allele as a genetic risk factor for left ventricular failure in homozygous β -thalassaemia. *Blood*. 1998;92:3455-3459.