DEVELOPMENT OF A HAEMOGLOBINOPATHY GENETIC DIAGNOSTIC SERVICE FOR THE NORTH WEST OF ENGLAND

A thesis submitted to the Manchester Metropolitan University
for the degree of Master of Philosophy in the
Faculty of Science and Engineering

April 2012

Charlotte Button

In collaboration with the Molecular Diagnostics Centre,
Department of Haematology,
Central Manchester NHS Foundation Trust

Abstract

Haemoglobinopathies represent the commonest single gene defects in the world and are found at highest frequency in those countries where malaria is prevalent. It is now recognised that due to the effect of migration they are not confined to specific geographical locations but occur at high frequency in local populations throughout many parts of the world. Although the haemoglobinopathies are highly heterogeneous at the molecular level with more than 1000 mutant alleles described (Hb Variant database), within each ethnic population group common subsets of mutations are found, consisting of a high frequency of several common mutations with a smaller number of rare mutations.

As part of the existing national Antenatal and Newborn Screening Programme, and in order to achieve a high mutation detection rate in the heterogeneous population of the North West of England (NW), and Greater Manchester in particular, a comprehensive suite of genetic diagnostic methodologies were developed to detect the wide spectrum of mutational defects associated with the Haemoglobinopathy disorders.

Between July 2008 and December 2011 a total of 328 referrals from the NW were genotyped at Manchester Royal Infirmary. The total number of haemoglobinopathy diagnoses comprised of: 34% (*n=113*) alpha thalassaemia, 33% (*n=109*) beta thalassaemia, 17% (*n=55*) were compound heterozygous for an alpha and beta thalassaemia mutation and/or variants and 3% (*n=10*) of patients had a mutation resulting in HPFH or δβ thalassaemia. In the remaining 13% (*n=41*) of referrals either no causative mutation was detected or the genotype detected was not enough on its own to account for the patients' phenotype.

A broad spectrum of mutations were detected within our referral population by DNA sequencing, resulting in a wide range of genetic interactions between different alpha and beta thalassaemia associated alleles, reflecting the complex molecular heterogeneity of these disorders. This indicates that a comprehensive mutational screening approach, using multiplex Gap-PCR, DNA sequencing and MLPA is appropriate to detect the mutation spectrum found within the heterogeneous populations of the North-West of England, or in any other ethnically diverse populations.

Acknowledgements

Firstly I would like to thank my supervisors Dr Steve Keeney and Dr Tony Cumming in the Molecular Diagnostics Centre at Manchester Royal Infirmary. I would especially like to thank Steve who has helped throughout this entire project, especially towards the end with final editing. I would also like to thank my supervisor Professor Bill Gilmore at Manchester Metropolitan University for his support throughout this project.

Also thanks to all the staff within the Molecular Diagnostics Centre for being so understanding and sensitive towards me, being heavily pregnant and stressed. Thanks also to all of those that made the lifesaving cups of tea.

I would especially like to thank Megan Sutherland. Her aptitude and concentration during training allowed for a smooth handover of the genetic haemoglobinopathy service, with no disruption and allowed me to focus on my thesis.

The biggest thanks goes to my husband (and personal chef) Dean. I wouldn't have been able to do it without you, those dinners have been a lifesaver!

I would also like to acknowledge the clinicians Dr Kate Ryan and Dr Andrew Will and the Genetic counsellors from the Manchester Sickle Cell and Thalassaemia Centre for sending me some interesting and challenging patient cases.

Furthermore thanks to my unborn baby girl who helped keep everything in perspective by kicking me during the hardest times.

Contents:

Chapter 3: Results

Chapter 4: Discussion

List of Figures:

List of Tables:

List of Abbreviations and Symbols:

Abbreviation	Description
α	Alpha
β	Beta
γ	Gamma
δ	Delta
ε	Epsilon
ζ	Zeta
θ	Theta
ψ	Psi (Pseudo Gene)
2,3-DPG	2,3-diphosphoglycerate
BCL11A	BCL11A gene
B-LCR	Locus control region within the β-globin gene cluster
bp	Base Pair
c.	Nomenclature for cDNA level
°C	Degrees Celsius
CMFT	Central Manchester NHS Foundation Trust
CO_2	Carbon Dioxide
CPA	Clinical Pathology Accreditation
ddNTPs	Dideoxynucleoside triphosphates
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleoside triphosphates
DMSO	Dimethyl Sulphoxide
EDTA	Ethylenediaminetetra-acetic acid
Fe	Iron
Fe^{2+}	Iron (terrous form)
Fe^{3+}	Iron (ferric form)
FOQ	Family Origin Questionnaire
fL	Femtolitre
g.	Nomenclature for genomic DNA level
Hb	Haemoglobin
HbA_2	Haemoglobin A_2
HBA1	Alpha-1 globin gene
HBA2	Alpha-2 globin gene

HBB	Beta globin gene
HBS1L	HBS1L gene
HBG	Gamma Globin gene
Hb F	Fetal Haemoglobin
HGVS	Human Genome Variation Society
HPFH	Hereditary Persistence of Fetal Haemoglobin
HPLC	High Performance Liquid Chromatography
HS-40	DNase I-hypersensitive site within the α-globin gene cluster
IEF	Isoelectric Focusing
IU/dL	International units per decilitre
IU/mL	International units per millilitre
IVS	Intervening sequence (an intron)
kb	Kilobases
KLF1	Kruppel-like factor 1 gene
LSD	Locus specific database
µg	Micrograms
µM	Micromolar
MCH	Mean cell Haemoglobin
MCV	Mean Cell Volume
MDC	Molecular Diagnostics Centre
MRI	Manchester Royal Infirmary
mg/uL	Milligrams per millilitre
mL	Millilitres
$MgCl_2$	Magnesium Chloride
mins	minutes
MLPA	Multiplex Ligation-dependent Probe Amplification
MYB	MYB gene
ng	Nanograms
N13F	N13 forward primer
N13R	N13 reverse primer
NHS	National Health Service
O_2	Oxygen
p.	Nomenclature for protein level
PCR	Polymerase Chain Reaction
pH	Potential Hydrogen

PND	Prenatal Diagnosis
RBC	Red blood cell count
ROX-500	DNA size standard
s	Seconds
SCT	Sickle Cell and Thalassaemia
sH_2O	Sterile Water
SNP	Single Nucleotide Polymorphism
SOX-6	SOX-6 gene
TBE	Tris borate EDTA buffer
UTR	Untranslated Region
w/v	Weight:Volume
ZPP	Zinc Protoporphyrin

Chapter 1: Introduction

1.1 Haemoglobin

Haemoglobin (Hb) is a protein that is contained within all red blood cells (RBC) of the human body. Its main function is the transportation of oxygen (O_2) from the lungs to the body tissues and it also helps to maintain the shape of the red blood cells, it is therefore essential for life.

The structure of the haemoglobin molecule was first described by Max F. Perutz in 1959, it is a tetramer consisting of four polypeptide (globin) chains which are found in pairs; two 141 amino acids long alpha globin chains (α_1, α_2), and two 146 amino acids long beta globin (β_1, β_2) chains (Steinberg, M *et al* 2009).

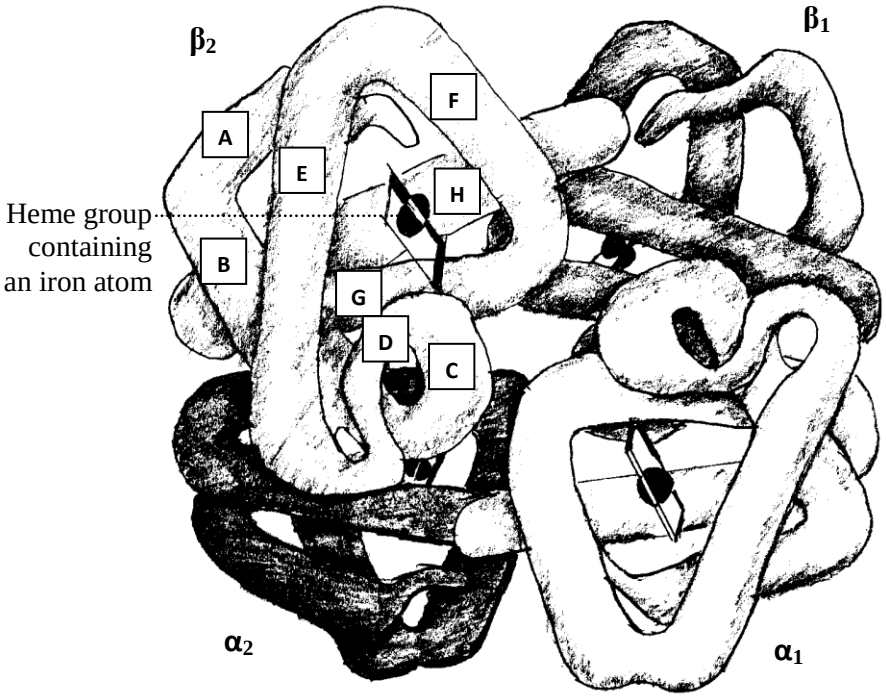


Figure 1.1: Quaternary structure of the haemoglobin A molecule

The four globin chains have a similar secondary and tertiary structure as they can each be divided into eight helical segments, designated A to H which then fold up and are held together by salt bridges and hydrogen bonds to form the precise three-dimensional spherical structure of the haemoglobin molecule which contains four binding sites or surface heme pockets to which up to a total of four oxygen molecules can bind. These heme groups are held between helices E and F, and contain an iron (Fe) atom and a porphyrin ring (comprised of carbon, hydrogen and nitrogen). Subsequent binding or

unloading of O_2 to each of these four iron atoms alters the quaternary structure of the haemoglobin molecule so that when it is in the deoxygenated state (deoxyhaemoglobin) and no O_2 is bound, the Fe atom exists in the ferrous form ($Fe2^+$) and is found out of the plane of the porphyrin ring of heme, whereas in the oxygenated state (oxyhaemoglobin) the iron becomes oxidised by the binding of O_2 and is converted to the ferric form ($Fe3^+$) causing it to move into the plane of the porphyrin ring.

Additionally the haemoglobin molecule is responsible for the transportation of up to 10% of carbon dioxide (CO_2) from the respiring tissues to the lungs, so it can be expired (Bain, B: 2006). Unlike oxygen, carbon dioxide directly binds to the amino group of the polypeptide chains of haemoglobin to form carbaminohaemoglobin. In the deoxygentated state the haemoglobin molecule has a higher affinity for CO_2 unlike oxyhaemoglobin which has a higher affinity for O_2, facilitating the loading and unloading of respiratory gases in the lungs and tissues, forming an efficient gaseous exchange system.

The haemoglobin molecule is an allosteric protein that has more than one shape and as mentioned above, is capable of conformatory changes in its quaternary structure by the interaction between the four globin chains and the heme groups allowing the haemoglobin molecule to reversibly load and unload oxygen and carbon dioxide molecules. The haemoglobin molecule is designed to switch between two alternative structures by structural rotations of the bonds in the side chains between the alpha and beta polypeptides in the F-helix; the deoxy or 'tense' (T) structure and the oxy or 'relaxed' (R) causing the heme molecules to change their positions within the haemoglobin molecule therefore causing the heme molecule to change positions.

The behaviour of the haemoglobin molecule is, to a certain extent, controlled by the partial pressures of respiratory gases found within the blood, thus the ability of the haemoglobin and deoxyghaemoglobin molecules to efficiently load and release the maximum amount of respiratory gases to and from the tissues can be explained by its cooperativity, i.e. its unique structure which is capable of transforming between the T and R structures in the presence or absence of oxygen. In a high O_2 pressure environment such as the lungs, the oxygen binds to deoxyhaemoglobin in the T structure and causes a rotation of the two pairs of polypeptides causing the Fe ions to shift out of the plane of the porphyrin ring and rendering them more accessible to oxygen binding which then transforms the oxygenated

haemoglobin to the R state, which has a higher affinity for oxygen and continues to increase as each successive molecule of oxygen binds.

Conversely, the transformation from the R to the T structure, which happens under conditions of low oxygen pressure such as that found in the tissues, results in the O_2 being released and causes the polypeptides to rotate back and shift the iron atoms so that they become less accessible to oxygen binding, this structural change prevents the irreversible binding of oxygen by allowing the iron atom to form loose bonds, allowing it to be readily released it into the tissues without it becoming permanently oxidised.

The strength by which haemoglobin binds to oxygen is affected by several factors and the haemoglobin molecule is capable of adapting its role of delivery of respiratory gases in response to changes in various physiological conditions, for example ensuring enhanced oxygen delivery to tissues under anaerobic conditions such as hypoxia, fever or as a result of an increase in metabolic rate by bought about by exercising muscle. The change in cooperativity of the globin chains makes it possible for the haemoglobin to alter its affinity for oxygen. Some physiological factors cause a shift in the oxyhaemoglobin dissociation curve to the right, for example a reduction in the pH of the blood caused by increased levels of CO_2 production from respiring tissues causes the haemoglobin molecule to interact with 2,3-diphosphoglycerate (2,3-DPG) which subsequently enhances O_2 delivery by favouring unloading of oxygen. The converse is true in conditions where the pH level is raised such as the lungs in which the haemoglobin molecule has an increased affinity for oxygen, therefore shifting the oxygen dissociation curve to the left.

1.2 Developmental regulation of haemoglobin synthesis

At the molecular level, haemoglobin synthesis is controlled by two multigene clusters on chromosome 16 and chromosome 11, from which a series of structurally distinct globin chains are synthesised at different rates throughout embryonic, fetal and adult life during normal developmental regulation. The genes are arranged on chromosome 11 and 16 in the order in which they are expressed during this development:

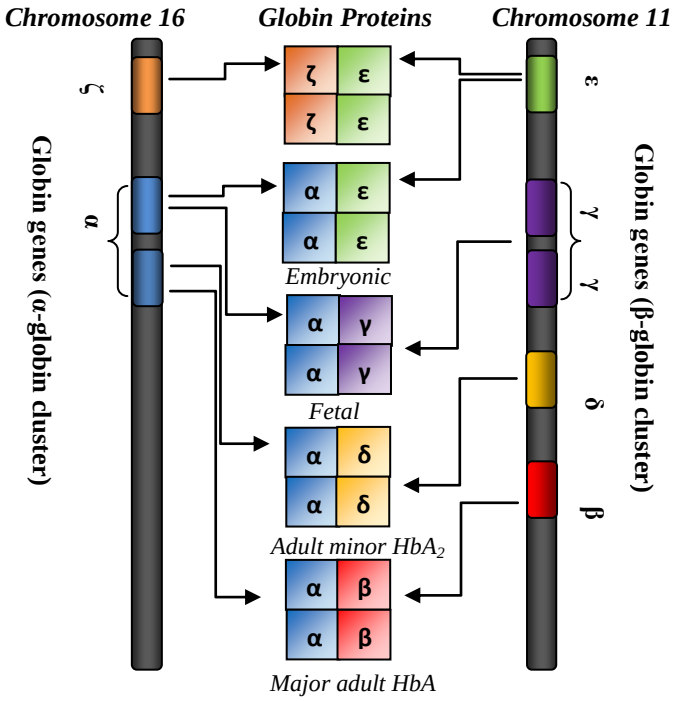


Figure 1.2: Location of the alpha and beta gene clusters.
(Derived from: http://www.ncbi.nlm.nih.gov/books/NBK10023/).

Figure 1.2 shows the alpha (α) gene cluster located on chromosome 16, made up of one embryonic ζ-globin and two highly homologous α-globin genes (*in cis*) 5'-ζ2- α2- α1-3'. The beta (β) gene cluster is located on chromosome 11, and is made up of 5'-ε-^{G}γ-^{A}γ-δ-β-3'. Therefore the genotypes αα/αα and β/β would be designated for a normal adult alpha and beta chain synthesis, respectively.

During the transition between fetal development, the neonatal period and beyond, different globin chains are synthesised and subsequently different haemoglobin tetramers are formed. These different types of haemoglobins found in the red cells at these different stages reflect the different sites at which erythropoiesis occurs; in the early embryonic stage (3-4 weeks gestation) erythropoiesis first occurs in the yolk sac, where specific Gower ($\zeta_2\varepsilon_2$) ($\alpha_2\varepsilon_2$) and Portland ($\zeta_2\gamma_2$) haemoglobins are synthesised. From 10-12 weeks gestation the liver and spleen take over as synthesis sites for a large proportion of fetal or haemoglobin F ($\alpha_2\gamma_2$) and a small proportion of adult or haemoglobin A ($\alpha_2\beta_2$), then later in fetal life and from birth into adulthood the bone marrow is the main site for erythropoiesis in which the predominant adult haemoglobin, haemoglobin A is synthesised together with a small portion of the minor haemoglobin, haemoglobin A_2 ($\alpha_2\delta_2$).

The developmental synthesis of different haemoglobin types between embryonic, fetal and adult life reflects the need for sufficient oxygen delivery from the mother to the fetus across the placental barrier, therefore embryonic and fetal haemoglobins tend to have a higher oxygen affinity than haemoglobin A. This regulatory expression of haemoglobins helps to explain why genetic abnormalities affecting the alpha chains present clinically in early fetal life, whereas abnormalities associated with the beta chain do not present until later in life and can often be difficult to diagnose in the early neonatal period because it is difficult to obtain a reliable quantification of HbA_2 (See Section 1.4.2.) (Bain, B: 2006, Steinberg, M *et al*: 2009, Ryan, K *et al* 2010).

1.3 Haemoglobinopathies

This inherited group of haemoglobin disorders result from a quantitative reduction in the rate of synthesis or absence of constituent globin chains (alpha or beta thalassaemia) or a qualitative structural change of one or more of the globin chains that combine to form haemoglobin resulting in a variant haemoglobin (haemoglobinopathy). Haemoglobinopathy disorders represent the commonest single gene defect in the world (Henderson, S *et al*: 2009, Bain, B: 2006, Clarke, B,E and Thein, S.L: 2004) with the World Health Organisation (WHO) estimating that approximately 7% of the world population are carriers of these disorders.

The thalassaemia syndromes were initially thought to be limited to the Mediterranean and South East Asian regions as the first clinical description was by Thomas B. Cooley in 1925 in infants living in the Mediterranean, but as the number of cases reported across the world increased it became evident that these thalassaemic disorders are were not confined to a specific geographical location and they now occur widely at high frequency throughout many parts of the world including many areas of Europe such as Italy, Spain, Greece and Portugal and also affect individuals within Middle Eastern areas such as Iran and Pakistan as well as those populations in southern China, Thailand, Malaysia and Indonesia. They are also the commonest inherited haemoglobin disorders on the Indian subcontinent (Weatherall, D.J and Clegg, J.B: 2001). The high prevalence of these disorders in some geographical locations reflects the protection these disorders offer against *P.falciparum* malaria (Weatherall, D.J and Clegg, J.B: 2001 and Bain, B:2006), for example in Africa where malaria is endemic, up to one-quarter of individuals are heterozygous for the alpha plus thalassaemia 3.7kb deletion (Bain, B.2006) and sickle cell anaemia is rife. Other factors that contribute to the high frequency and heterogeneity of

these disorders worldwide include general population growth, consanguinity which is common in those regions where thalassaemia is prevalent (Weatherall, D.J: 2011: BSH) and increasing population migration and inter-racial marriages.

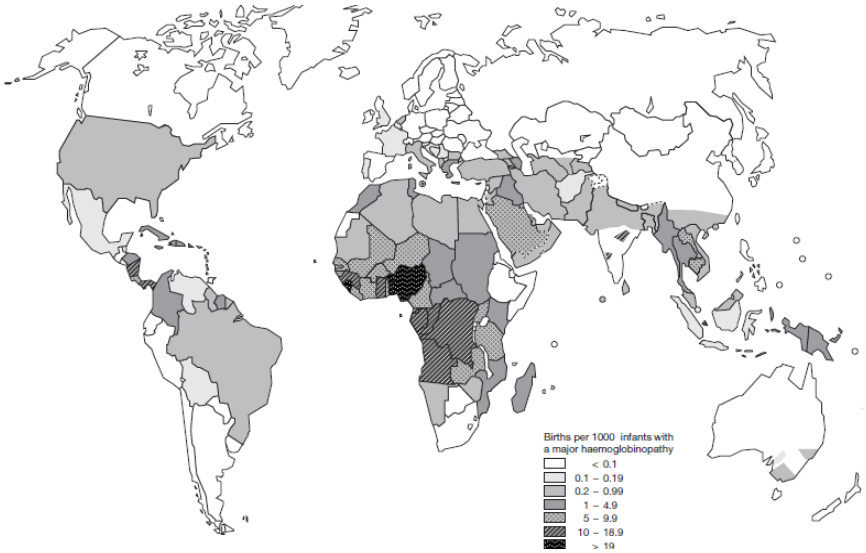


Figure 1.3: Global distribution of haemoglobin disorders (based on the number of births of affected infants per 1000 births).
(Source: http://www.who.int/genomics/public/Maphaemoglobin.pdf)

The incidence at which the two subgroups of thalassaemia (alpha and beta) occur differ within discrete geographical clusters around the world; alpha thalassaemia is found at highest prevalence in African, Southeast Asian and Chinese populations and it rarely occurs in British Caucasian populations, whereas beta thalassaemia occurs mainly within the Mediterranean, Greece and Turkey as well as Italy and the Indian and African subcontinents, and throughout the Middle East.

The haemoglobin disorders present with a broad spectrum of clinical symptoms with varying clinical effects which is governed by which of the globin chain(s) are involved, but many have major implications for those individuals and families that are affected. Different interactions of these genetic disorders can result in death *in utero* if there is a complete abolition of alpha-globin chain synthesis, or severe transfusion dependant haemolytic anaemia if there is a complete absence of beta-globin chain synthesis. This imbalance of globin chain synthesis caused by the thalassaemia syndromes contributes via ineffective erythropoiesis to a range of pathological effects including haemolytic anaemia with excess, unpaired globin chains precipitating in red cell precursors in the marrow (Bain, B. 2006).

The impact of human immigration across the world has led to a major increase in the heterogeneity of these disorders by distributing these alleles outside of the tropical belt and throughout the world (Weatherall, D: 2011). Haemoglobinopathy disorders are common now in the United Kingdom (U.K) and constitute a significant public health problem and present a diagnostic challenge for healthcare professionals. To date more than 1,000 haemoglobin disorders have been reported (NHS Sickle Cell and Thalassaemia Screening Programme: 2009, Old, J.M: 2003), but only a small proportion of these disorders result in a severe disease phenotype and it is only clinical relevant to detect those disorders as identified by the Antenatal screening programme in which Prenatal diagnosis (PND) may be indicated and the offer of termination of pregnancy, where appropriate. For example, haemoglobin Bart's hydrops fetalis syndrome which is incompatible with fetal life and can result in further obstetric complications for the mother.

1.4 Classification of haemoglobinopathies

As discussed in section 1.3, the genetic disorders of haemoglobin are divided into those in which there is either a reduction in the rate of synthesis or absence of constituent globin chains (Thalassaemia) or those in which there is a structural change in the α or non α globin chains that combine to form a variant haemoglobin (Haemoglobinopathy). Thalassaemia syndromes can be further sub-classified based on the gene involved i.e. alpha (α) or beta (β) thalassaemia and on their severity i.e. how much the causative mutation down regulates the amount of protein produced. Alpha or beta plus (α^+, β^+) indicates there is some functional globin protein produced whereas alpha or beta zero (α^o, β^o) indicates that no respective globin protein is produced.

1.4.1 Alpha thalassaemia

Alpha thalassaemia results in a reduced rate of synthesis of alpha globin chains and is associated with anaemia with a variable degree of microcytosis and hypochromia. The normal complement of alpha-globin genes is four (αα/αα) as two alpha genes are inherited on each chromosome 16, therefore this disorder can be classified according to the number of alpha globin genes that are missing or inactive which also dictates the clinically severity of the condition, which can vary from an asymptomatic clinical picture in which one alpha gene is removed or inactivated to a lethal haemolytic anaemia where all four alpha gene are affected (Harteveld, C.L and Higgs, D.R: 2010).

Loss or inactivation of one alpha globin gene (-α/αα) often results in a 'silent' asymptomatic carrier status therefore individuals with this genotype may present with a completely normal red blood cell indices on a Full Blood Count (FBC). Loss of two alpha globin genes from either the same chromosome 16, *in cis* (αα/--) or one from each chromosome, *in trans* (-α/-α) causes a mild to moderate microcytic anaemia. Loss of three genes (-α/αα) results in Haemoglobin H disease, which is associated with marked haemolytic anaemia caused by the excess of unpaired β-chains, that combine to form abnormal haemoglobin tetramers of Hb H (β4) in adult life or Hb Barts (γ4) in fetal life. Hb H disease is a clinical syndrome which is associated with a very variable phenotype, depending on the underlying genetic lesion, the most common genotype being compound heterozygosity for deletional α^{+} and α^{0} thalassaemia. Loss of all four genes (--/--), in which there is a complete absence of functional alpha-globin chains is incompatible with life and leads to intrauterine death as synthesis of haemoglobins F, A or A_2 cannot occur, the clinical syndrome is called haemoglobin Barts hydrops fetalis (Bain, B: 2006 and Weatherall, D: 2001) in which the only haemoglobin present is haemoglobin Barts.

Alpha thalassaemia is predominately caused by large deletions in one or both of the alpha genes (*HBA1* and *HBA2*) on the chromosome, in which more 95% of recognised cases are caused by a deletion (Chong, S *et al*: 2000). Alpha zero thalassaemia is most commonly caused by large deletions that completely or partially remove both of the alpha globin genes.

Figure 1.4 shows a diagrammatic representation of the extent of the five most common deletions (and the British deletion) along the alpha gene cluster that are associated with alpha zero thalassaemia, they are designated --SEA, --FIL, --THAI, --MED and $-(\alpha)^{20.5}$ and –BRIT.

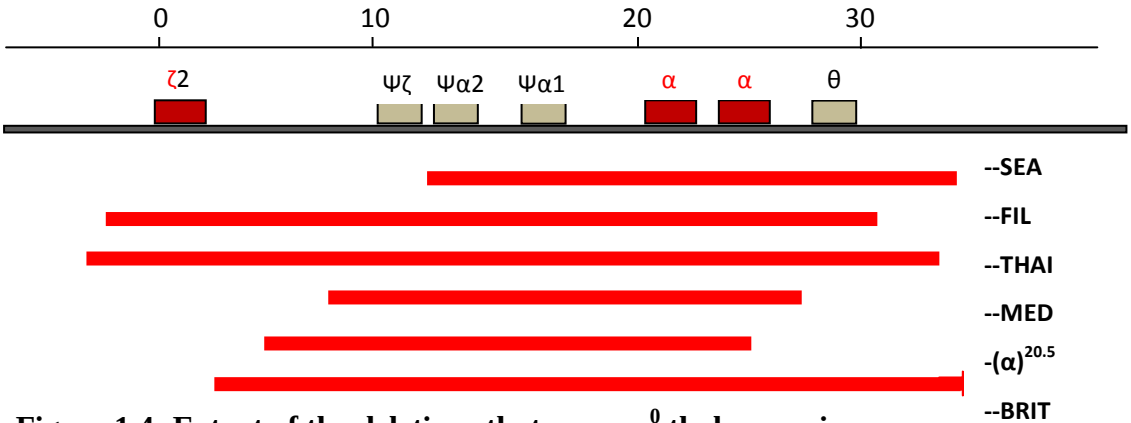


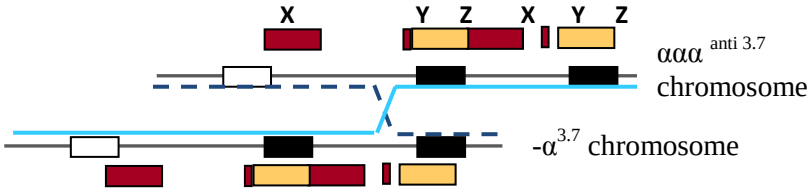
Figure 1.4: Extent of the deletions that cause α^{0} thalassaemia

The extent to which the above deletions span the alpha gene cluster vary, but they all remove or inactivate both of the alpha globin genes on one chromosome therefore no alpha globin chains are synthesised, with the larger deletions also removing the ζ and ψ genes, therefore these genotypes are of major clinical significance as homozygotes or compound heterozygotes for these deletions will exhibit Hb Barts Hydrops Fetalis syndrome.

Detection and quantification of the abnormal haemoglobins associated with alpha thalassaemia (Hb H and Hb Barts) by High Performance Chromatography (HPLC) can be helpful in predicting severe alpha thalassaemia genotypes in newborns, with a cut off of more than 25% Hb Barts being chosen by the Newborn Screening programme to highlight those babies at risk of developing Hb H disease (Sickle Cell and Thalassaemia Handbook for laboratories: http://sct.screening.nhs.uk/policy).

Alpha plus thalassaemia is most commonly caused by the $-\alpha^{3.7}$ and $-\alpha^{4.2}$ deletions, in which 3.7 kb and 4.2 kb of genomic DNA respectively, are removed from within the alpha gene cluster (Embury, S.H *et al*: 1980 and Baysal, E and Huisman, T:1994). It is thought that these are a result of a recombination event that happens during meiosis causing unequal crossing over of genetic material, this happens because each of the alpha globin genes are located within a highly homologous 4 kb area. This can be further subdivided into three homologous subsegments on each alpha-globin gene which are designated X, Y and Z boxes.

Rightward crossover (Z-box):



Leftward crossover (X-box):

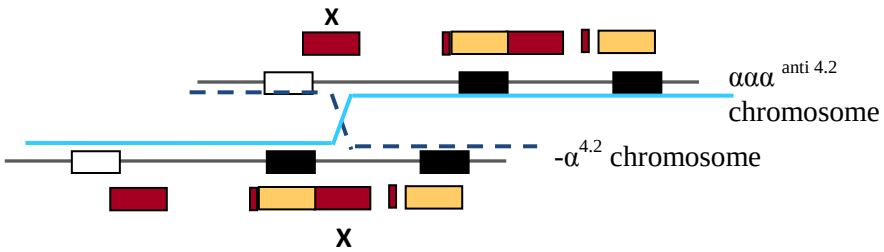


Figure 1.5: Diagrammatic representation of the mutational mechanisms associated with the $-\alpha^{3.7}$ and $-\alpha^{4.2}$ deletions associated with α^{+} thalassaemia. (Derived from Harteveld, C.L. and Higgs, D.R: 2010).

The -$\alpha^{3.7}$ kb deletion is a caused by the unequal cross over between the mispaired Z-boxes (which are 3.7 kb apart) and results in a chromosome with only one functional alpha gene (rightward -$\alpha^{3.7}$ kb deletion) and a chromosome with three alpha genes ($\alpha\alpha\alpha^{anti3.7}$ triplication), causing α^{+} thalassaemia.

Similarly, the $-\alpha^{4.2}$ kb deletion is caused by a non-reciprocal cross over event between the X-boxes (which are 4.2 kb apart) which also gives rise to a chromosome with only one functional alpha gene (leftward $-\alpha^{4.2}$ kb deletion) and a chromosome with three alpha genes ($\alpha\alpha\alpha^{anti4.2}$ triplication), causing α^{+} thalassaemia.

Because both these mutational mechanisms only remove all or part of the *HBA2* gene it does not cause a clinically significant disorder, even when inherited in the homozygous state (-α/-α) it does not cause a severe disease phenotype, subsequently alpha plus thalassaemia presents no risk to the fetus and would therefore not warrant PND analysis (http://sct.screening.nhs.uk/policy).

This cross-over event can also result in a duplication of an alpha globin gene, in which there may be three alpha genes on a single chromosome, known as triple alpha (ααα) (Trent, R.J *et al*: 1981 and Lie-Injo, L.E e*t al*:1981), leading to an increased rate of synthesis of that globin chain. Duplication of this gene does not cause a significant clinical disorder and may actually normalise the globin chain imbalance when it co-exists with a deletion of an alpha globin chain *in trans*, the converse is true when the ααα genotype co-exists with beta thalassaemia in which case it can aggravate the condition.

Co-inheritance of deletional alpha thalassaemia with other genotypes can module the clinical severity of the phenotype, especially in the homozygous state, when found in combination with beta thalassaemia the globin chain imbalance is normalised therefore alleviating the clinical phenotype. However, this can pose a diagnostic challenge in identifying alpha thalassaemia trait because the red cell indices are 'normalised' on a FBC, additionally in the compound heterozygote state with a beta chain variants such as Hb S or Hb E, it lessens proportion of the variant detected by phenotypic tests such as High Performance Liquid Chromatography.

Less commonly α^+ thalassaemia can be caused by a non-deletional mutational mechanism, in which a single-point mutation affects either one of the alpha globin genes and is denoted (α^T) (Weatherall, D.J and Clegg, J.B: 2001). In comparison with the deletional types this group of thalassaemias tend to have a more pronounced effect on the expression and synthesis of the alpha globin chains resulting in a more severe phenotype, this is partly because the majority of these point mutations tend to affect the alpha-2 gene which is responsible for about 70% of alpha chain production (Bain, B: 2006, Weatherall, D.J and Clegg, J.B: 2001, Aguilar-Martinez and Gulbis: 2010, Harteveld, C.L and Higgs, D.R: 2010) and there is no compensatory increase in the expression of the alpha-1 gene, as there is in the deletional types.

Non-deletional forms of alpha thalassaemia are considered to be of clinical significance as in the homozygous or compound heterozygous state with α^0 thalassaemia they can cause particularly severe forms of Haemoglobin H disease or even Hb H hydrops fetalis syndrome (Viprakasit, V *et al* 2002). The mutational mechanism by which they affect the gene varies; those that that commonly affect the polyadenylation signal such as the $\alpha^{PA6:A>G}$ and α^{TSaudi} mutations, reported within Saudi Arabian populations, cause an extension of the alpha globin gene transcripts and interestingly they have been reported to down regulate the synthesis of alpha chains on both of the alpha genes on the same chromosome (Weatherall, D.J and Clegg, J.B: 2001), whereas other mutations such as $\alpha^{IVS1(-5nt)}$ and $\alpha^{IVS1-117(G>A)}$, often found in Mediterranean and some Indian populations affect RNA splicing. Other non-deletional mutations cause alterations to the tertiary structure of the haemoglobin molecule and subsequently result in the formation of highly unstable haemoglobin variants, for example Hb Constant Spring ($\alpha^{CS}\alpha$) (End, L.I *et al*: 1973, Kan, Y.W *et al*: 1974), Hb Seal Rock ($\alpha^{SR}\alpha$) and Hb Quong Sze ($\alpha^{QS}\alpha$) which are reported at highest prevalence within the Middle East, Mediterranean, South East Asia, China and Cyprus can cause chronic haemolytic anaemia due to the secondary damage to the red cells caused by these unstable variants, which are often undetectable by HPLC.

As outlined earlier in section 1.3, alpha thalassaemia is found at highest frequencies in all tropical and subtropical populations were malaria is prevalent (Harteveld, C.L and Higgs, D.R: 2010, Aguilar-Martinez, P and Gulbis, B.:2010) although the spectrum of α^+ and α^0 thalassaemia mutations varies throughout the world. Alpha plus thalassaemia is found at highest prevalence in individuals from Africa, with a carrier frequency of between 10-30% (Ryan, K *et al*: 2010) but it is also found in other ethnic groups such as in South Asia,

Pakistan, India and some Middle Eastern populations such as Iran, Yemen and Kuwait. Alpha zero thalassaemia is most commonly found in those individuals indigenous to the South-Eastern areas of the world including China, Thailand, Malaysia and the Philippines and the Mediterranean regions such as Greece, Turkey and Cyprus (Bain, B: 2006), although at a much lower frequency.

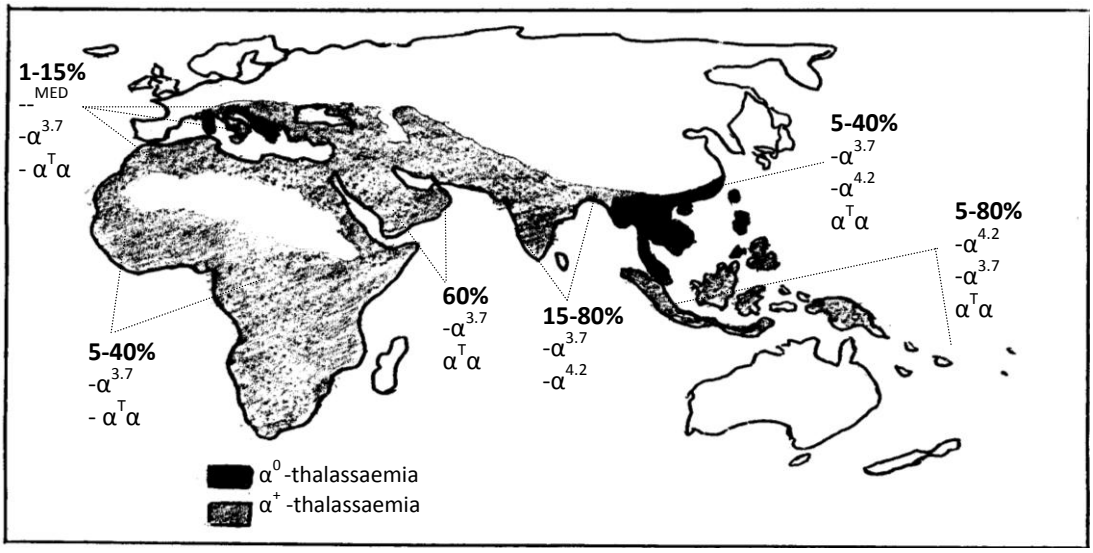


Figure 1.6: Worldwide geographical distribution and frequency of alpha thalassaemia alleles (Derived from Weatherall, D.J: 2001).

Traditionally the distribution of alpha thalassaemia is that as represented in Figure 1.6, with it rarely reported outside of African and South East Asian populations. Over the last few decades the effects of immigration (Henderson *et al*: 2009) has led to demographic changes therefore increasing the 'incidence of alpha thalassaemia in Northern European countries, Northern America (Harteveld, C.L and Higgs, D.R: 2010) and other previously non-endemic areas, including the UK in which a number of cases of alpha zero thalassaemia, caused by the British deletion (--BRIT) have been reported in several families from Wigan (Eng: 2009, Higgs and Bowden: 2001, Higgs *et al*: 1985). Results from a study by Henderson *et al* (2009) highlights the complex ethnic diversity populations within the UK and the great molecular heterogeneity found; out of the total of 67 known non-deletional α^+ mutations described on the Hb Variant Database over 30 of these mutations have been identified within the UK population. The molecular heterogeneity of these disorders highlights the potential for interactions between the different alpha thalassaemia alleles.

1.4.2 Beta thalassaemia

Beta thalassaemia results in a reduced rate of synthesis of beta globin chains and is also associated with a variable degree of microcytosis and hypochromia, resulting a reduction in the red cell indices and a characteristic increase in the haemoglobin HbA_2 percentage.

Unlike alpha thalassaemia, beta thalassaemia is most commonly caused by point mutations, small insertions or deletions within the beta globin gene (*HBB)* or its immediate flanking sequences, in which the underlying genetic lesion causes a reduction in the synthesis of beta globin chains which comprises adult haemoglobin. This leads to an excess of unpaired alpha-globin chains which then precipitate in the erythroblasts and mature red cells causing severe haemolysis, therefore the greater the excess of α-chains the more pronounced the anaemia.

Beta thalassaemia often presents at around 6 months of age when the expression of the major adult haemoglobin profile develops and β-globin chain synthesis replaces γ-chain synthesis, although more severe haematological abnormalities can emerge as early as 3 months of age (Weatherall, D and Clegg, J.B:2001). This group of haemoglobin disorders are extremely heterogeneous at the genetic level, with more than 200 causative mutations described in a wide range of ethnic groups (Weatherall, D.J and Clegg, J.B: 2001, Bain, B: 2006). Beta thalassaemia can be subdivided into different categories based on the varying levels of Hb F and Hb A_2, detected by HPLC and the severity of the clinical presentation.

Normal or silent beta thalassaemia (β^{++}), as the name suggests is a clinically silent haematological clinical condition which is often only identified in those individuals when it is co-inherited with a more severe beta thalassaemia allele. These milder phenotypes are often associated with mutations within the promoter regions of *HBB* for example -101 (C>T), -92 (C>T) -88 (C>T) and -29 (A>G) mutations and are normally associated with normal red cell indices, normal HbA_2 levels but are often associated with increased Hb F percentages of up to 4% (http://globin.bx.psu.edu/hbvar/).

Beta plus (minor) thalassaemia (β^{+}) is also considered to be a mild clinical condition as it does not result in significant haematological changes in the heterozygous state but nevertheless can cause a clinically significant disease in compound heterozygotes and homozygotes. Described beta plus mutations include IVS1-6 (T>C), CAP+1 (A>C) and

IVS 1-5 (G>A). The resultant degree of globin chain imbalance is slightly more pronounced than in the above silent forms with a slight reduction in the red cell indices and the HbA_2 levels are often elevated between 3.6 and 4.2% (Bain, B: 2006).

Even though these milder forms of beta thalassaemia are not considered to be clinically significant they can be difficult to diagnose as the haemoglobin A_2 and haemoglobin F levels are often borderline or even normal, additionally co-inheritance of either alpha zero or homozygous alpha plus thalassaemia can make it even more likely that the diagnosis of β-thalassaemia could be missed as this can result in normal red cell indices.

Furthermore there is a subgroup of β^+ thalassaemia conditions which are associated with a particularly severe clinical course, even in the heterozygous state. These are classified as severe β^+ thalassaemia and are commonly associated with mutations within the 3' end of exon 2 or 3 of *HBB* and splice site junctions such as the IVS1-5 (G>C) and IVSII-654 (C>T) (http://globin.bx.psu.edu/hbvar/) which give rise to elongated unstable β-globin chains that form inclusion bodies within the red cell precursors, subsequently shortening their survival. These mutations are associated with higher HbA_2 levels of between 4 to 5% in heterozygotes which are comparable to those found in beta zero thalassaemia, discussed below. Additionally, in contrast to the above mild recessive forms of the disease, these conditions are inherited in an autosomal dominant manner; therefore patients with dominant beta thalassaemia may require occasional blood transfusions.

Beta zero (β^0) thalassaemia trait, in contrast to β^+ thalassaemia, is caused by a variety of mutational mechanisms for example Codon 30 (G>C), Codon 8/9 (+G) and Codon 39 (C>T) that result in a significant reduction or even complete absence of beta globin chain synthesis which is evident by a marked microcytosis, hypochromia and raised haemoglobin A_2 and F levels. Beta zero thalassaemia is important to identify as in the homozygous or compound heterozygous state it can result in beta thalassaemia major ($\beta^0\beta^0$), a clinical condition in which there is no Hb A (only haemoglobins F and A_2 are present) as there is a complete absence of beta globin chain synthesis, resulting in a severe transfusion dependant haemolytic anaemia.

Furthermore, beta thalassaemia can be further sub-classified to include those patients whose clinical phenotype presents in-between that of mild beta thalassaemia and the

transfusion dependent conditions; these are referred to as beta thalassaemia intermedia and often presents at a later age of between 6-7 years. There are a wide range of genotype combinations that can cause beta thalassaemia intermedia, including homozygosity for β^+ thalassaemia ($\beta^+\beta^+$), compound heterozygosity for β^+ and β^0 thalassaemia (β^+/β^0) and homozygosity for severe β^+ thalassaemia. The combinations in which these beta thalassaemia alleles are inherited dictates the clinical severity, for example, in the milder compound heterozygous states ($\beta^+\beta^+$) Hb A may still be detectable. This group of disorders also encompasses those cases in which the α/β globin chain imbalance in cases of beta thalassaemia is further aggravated by the co-inheritance of triplicated alpha thalassaemia (ααα) or ameliorated by the coinheritance of deletional alpha thalassaemia, particularly α^0 thalassaemia.

Rarer forms of beta zero thalassaemia are caused by deletional mutational mechanisms which remove the portions of the beta globin gene (*HBB*). A small number (*n=14*) of deletions affecting only *HBB* have been described (http://globin.bx.psu.edu/hbvar/) with the most commonly reported deletion being the 619 bp deletion which is isolated to mainly Indian and Pakistani populations, a study by Colah *et al* (2010) reported this deletion was the second common β-thalassaemia mutation found in Gujarat, India. Other deletional types, although are extremely rare are of phenotypic interest as they tend to be associated with unusually high levels of HbA_2 of between 4.6% to 9% in heterozygotes (Weatherall, D.J and Clegg, J.B: 2001).

It is evident that the beta thalassaemia disorders are highly heterogeneous at the molecular level and therefore often presentation with a wide range of clinical manifestations. The large spectrum of beta thalassaemia mutations reported and the diverse range of possible gene-gene interactions makes it challenging for clinicians to predict clinical severity, especially in those cases of beta thalassaemia intermedia, as the condition varies greatly in severity and the severity of the phenotype associated with a specific genotype is not always predictable (Bain, B: 2006, Ho *et al*: 1998). Furthermore the clinical severity of β-thalassaemia is also influenced by other factors including those affecting the α- and γ-globin gene expression (Ho *et al*: 1998)

Beta thalassaemia occurs in a wide range of ethnic groups and each population has its own range of mutations which usually account for most of the thalassaemia in that area, as shown in Figure 1.7 below:

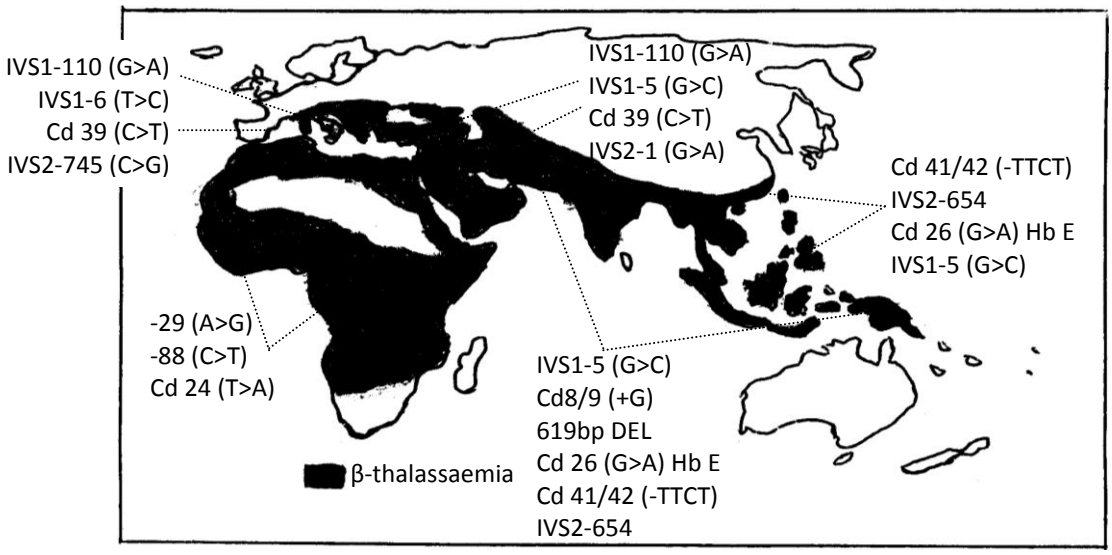


Figure 1.7: Worldwide geographical distribution of beta thalassaemia alleles. (Derived from Weatherall, D.J: 2001).

The prevalence of beta thalassaemia is widespread throughout the Mediterranean regions, India, South-East Asia and African countries in which there are particular mutations clustered around specific geographical locations. The distribution map above shows that Africa has an extremely high prevalence of mild beta thalassaemia mutations in which the promoter gene mutations -29 (A>G) and -88 (C>T) account for over three quarters of the cases of beta thalassaemia (Weatherall, D.J and Clegg, J.B: 2001). β^+ thalassaemia is reported at high prevalence amongst Turkish and Greek Cypriots in which the causative mutation is most commonly IVS1-110 (G>A), this is also commonly reported in the Middle East. The severe β^+ mutation IVS1-5 (G>C) is most commonly found in the Mediterranean, Middle Eastern regions and also India and Sri Lanka (Fisher, C.A *et al*: 2003, Colah, R *et al*: 2010) whereas in South East Asian regions the mutation IVS2-654 (C>T) accounts for a large proportion of thalassaemia. In India most of the mutations causing β thalassaemia are β^0 alleles such as codon 8/9 (+G), codon 41/42(-TTCT) and the 619bp deletion, which accounts for 16% of the thalassaemia there (Weatherall, D.J and Clegg, J.B: 2001), whereas in Spain and Portugal the β^0 mutation IVS1-1 (G>A) mutation is very common.

As a result of recent migration, beta thalassaemia has become a genetically heterogeneous disorder, which is now prevalent worldwide. It is not uncommon to find these disorders in

other white populations living in Northern Europe and the UK (Hickman, M *et al*: 1998). A study by Henderson *et al* (2009) found that the UK had the largest number of different β-thalassaemia mutations (68) in comparison to other countries. Another study by Fisher *et al* (2003) further highlights the considerable heterogeneity of the beta thalassaemia disorders observed in Sri Lanka, even within populations who live in close proximity, which is caused by the migration of various ethnic groups to and from the island. A later study by Colah *et al* (2010) also revealed molecular heterogeneity in Indian sub-groups, in which there was a large number of rare mutations reported which may be a result of several invasions throughout history and migration of populations between countries due to commercial trading, leading to mixing of the gene pool.

1.4.3 Hereditary Persistence of Fetal Haemoglobin disorders

In adult life the proportion of haemoglobin F is usually very low (around less than 1% of the total haemoglobin) but this broad spectrum of disorders is used to describe those conditions in which there is a persistence of raised fetal haemoglobin beyond infancy and in adulthood and arises at around 6 months of age when the switch from fetal to adult haemoglobin normally occurs (Weatherall, D.J and Clegg, J.B: 2001, Bain, B: 2006).

These disorders can be subdivided into Hereditary Persistence of Fetal Haemoglobin (HPFH) disorders and delta-beta (δβ) thalassaemia and Gamma delta beta thalassaemia (γδβ) all of which are caused by genetic defects within the β- and δ-gene clusters and are compensated by the up-regulation of γ-chain synthesis and consequently fetal haemoglobin, the degree to which depends on the underlying mutational mechanism, with the deletional forms being the most common.

The HPFH disorders have been described in a variety of ethnic groups but are most commonly found in African, Mediterranean, Chinese and Indian populations and are relatively benign in the heterozygous state as they are associated with a balanced α/β chain synthesis ratio and normal HbA_2 levels. Even in the homozygous or compound heterozygous state with β-thalassaemia they result in a very mild clinical picture due to a compensatory increase in synthesis of the γ-chains in response to the deletion of the β-gene on the affected chromosome, resulting in relatively high Hb F levels of between 15-30%. Furthermore the high proportion of haemoglobin F associated with these disorders may help to alleviate the clinical severity of some beta haemoglobin variants if they are

co-inherited with deletional forms of HPFH, for example the symptoms of sickle cell anaemia may be alleviated to a very mild sickling disorder.

In contrast, the δβ disorders are associated with a more severe clinical picture even though they are also a result of deletions of both the δ- and β-genes and $^{A}\gamma$ gene in the γδβ disorders. They result in an imbalanced ratio of α/β globin synthesis and are associated with slightly lower elevated Hb F levels than HPFH disorders of between 5-15% and the phenotype often resembles that of beta thalassaemia trait, except the HbA_2 percentage is often normal or decreased. Therefore in the heterozygous state it may result in thalassaemia intermedia and thalassaemia major in the homozygous or compound heterozygous state. It is therefore clinically important to distinguish between these subgroups of disorders.

Known non-deletional forms of HPFH are caused by point mutations and polymorphisms within the promoter regions of either the $^{G}\gamma$- or $^{A}\gamma$-genes and as in the deletional types, these mechanisms are associated with a persistently raised haemoglobin F level into adulthood, however in the non deletional types haemoglobin A and A_2 continue to be synthesised at a reduced rate. The sites at which these polymorphisms and/or mutations are found within the two gamma gene loci (*HBG2* and *HBG1)* may either alter the binding of regulatory factors that affect *HBG* expression or are in linkage disequilibrium with these functional sites (Akinsheye, I *et al* : 2011, Wilber, A *et al*: 2011).

A further interesting group of non-deletional HPFH disorders which are caused by polymorphic changes within the regulatory sequences of the $^{G}\gamma$- or $^{A}\gamma$-genes have also been reported to further influence the percentage of haemoglobin F. A common sequence polymorphism known as the Xmn*1* (C>T) restriction site polymorphism at position -158 upstream of the $^{G}\gamma$ gene has been reported to increase the synthesis of Hb F by a similar mechanism to that described above, although its exact mechanism is unclear. Evidence shows that it has the greatest effect on increasing the Hb F levels in those individuals whose genotypes cause erythropoietic stress, for example β-thalassaemia or sickle cell anaemia, but it has minimal effect in normal individuals, therefore the presence of this polymorphism is likely to have an indirect effect which is also related to the amount of Hb F produced by individual F cells (Sampietro, M *et al*: 1992).

It has been recognised that the production of fetal haemoglobin in individuals may be linked to a large number of other gene variants that are in linkage disequilibrium and

affect the binding of transcription factors within the beta globin cluster, but there are also other genetic loci that have been reported outside of the β-globin gene cluster, that have been shown to influence the production of fetal haemoglobin, including sequence variants in the BCL11A gene, on chromosome 2, which encodes a repressor of γ-globin gene expression and is an important regulator of γ-globin gene silencing, additionally mutations in the erythroid-specific transcription factor, Kruppel-like factor (KLF1) and SOX-6 have shown to play a significant role in the fetal to adult haemoglobin switch, furthermore the two proteins SOX-6 and BCL11A have been reported to cooperate together to silence γ-gene expression, and finally gene variants in the intergenic regions between the *HBS1L* and *MYB* genes on chromosome 6, further influences the expression of the γ-globin gene by the altered expression of the MYB gene (Donze, D *et al*:1995, Craig *et al*:1996, Thein, S.L *et al*:2007, Thein, S.L and Menzel, S: 2009, Galarneau G, *et al*: 2010, Akinsheye, I *et al* :2011, Wilber, A *et al*:2011). As a consequence the discovery of these variants has become a large area of research, mainly for the therapeutic purpose of ameliorating the clinical manifestations of certain thalassaemic conditions such as beta thalassaemia and sickle cell disease.

1.4.4 Haemoglobin Variants (Haemoglobinopathies)

In contrast to the thalassaemia conditions the haemoglobinopathies (or haemoglobin variants) are caused by a qualitative structural changes in either the alpha, beta, gamma or delta globin chains and are most commonly a result of a substitution of a single amino acid within the candidate gene, but rarer mutational mechanisms have also been reported including deletions, in which one or more amino acids are deleted, for example the beta chain variant Hb Leiden, fusion of haemoglobin chains or extensions of the amino acid chain.

The normal complement of globin chains reflects the corresponding percentage at which a variant is found, for example a variant affecting the beta globin chain usually comprises around 12-40% of the total haemoglobin, whereas an alpha chain variant comprises up to 25%, although the actual percentage at which the variant is detected can be altered by a variety of factors; if a variant haemoglobin is highly unstable it may be detected at a much lower percentage or may even go undetected if the variant is rapidly broken down. Additionally the co-inheritance of deletional alpha thalassaemia may proportionally reduce the percentage of a variant on the beta chain, conversely it would increase the percentage

of an alpha chain variant, for example Hb G Philadelphia (*HBA*2 or *HBA1*:c.207C>G or C>A) normally comprises between 20-25%, but in 80% of cases, this variant occurs with a chromosome that carries the 3.7 kb deletion (Bain, B: 2006) further increasing the proportion of the variant to 30-35% and even higher percentages of ~45% if co-inherited *in trans*. Furthermore the proportion of alpha chain variants may vary due to the differences in transcription rates between the alpha 1- and 2-globin genes with the alpha-2 gene being responsible for a larger proportion of globin chain synthesis, therefore an alpha chain variant may actually comprise anywhere between 12.5% and 37.5% (Bain, B: 2006). To date there are over 1000 variant haemoglobins reported on the Hb variant database (http://globin.bx.psu.edu/hbvar/) the majority of which are rare or clinically benign as they present with normal red blood indices. However some of the structural variants can result in a thalassaemic phenotype which is dependent on the underlying mutational mechanism often referred to as thalassaemic haemoglobinopathies which will be explored below. It is important to identify those clinically significant structural variants which occur at high frequencies in those areas where thalassaemia is also prevalent as their co-inheritance is relatively common since they often occur in the same population groups (Weatherall, D.J and Clegg, J.B: 2001) and result in a wide spectrum of disorders of varying clinical severity.

The structural haemoglobin variants that cause thalassaemic phenotypes are often so unstable they are unable to form the normal tetramers of haemoglobin and subsequently cannot transport oxygen efficiently. (Bain, B: 2006, Steinberg, M *et al*: 2009, Steinberg, M and Adams: 1983). They are often caused by a genetic lesion within the intervening/splice site regions of the globin gene which is therefore likely to promote alternative splicing, for example the beta chain variant Hb E (*HBB*:c.79G>A) is considered a thalassaemic haemoglobinopathy as it creates an alternative splice site for messenger RNA (mRNA) which is 5' to the normal splice site at IVS-I. Subsequently the aberrantly spliced mRNA is so unstable that the β^E chain is synthesised at a lesser rate than Hb A and the variant haemoglobin comprises 25-30% in the heterozygote which is less than expected for a beta chain variant. The Hb E variant is associated with a β-thalassaemia phenotype, therefore, if co-inherited with another β-thalassaemia mutation it is likely to result in a severe form of thalassaemia intermedia or major. These conditions form one of the commonest severe forms of thalassaemia in South East Asia and India, but studies have highlighted the varied clinical expression associated with this variant as it tends to present with a diverse range of

clinical symptoms which can be attributable to the degree of globin chain imbalance (Tubsuwan, A *et al*: 2011).

Other variants that result in a similar thalassaemic expression due to the production of aberrantly spliced mRNA can be found on the alpha globin genes, such as Hb Evanston (*HBA2*:c.43T>A) or (*HBA1*:c.43T>C and Hb Fort Worth (*HBA2*:c.83A>G (or *HBA1*), both of these variants are ineffectively synthesised and their instability may account for their low abundance which may give rise to an α^+ phenotype with mild microcytosis and hypochromia. Other alpha chain haemoglobin variants such as Hb Suan Dok (*HBA2*:c.329T>G), Hb Quong Sze (*HBA2*:c.377T>C) and Hb Constant Spring (*HBA2*:c.427T>C) result in hyper-unstable variant globin chains which are unable to form haemoglobin tetramers for a variety of reasons, depending on the underlying genetic mutation. A single base substitution within the termination codon can cause an elongated alpha globin chain, as found in the alpha chain variants Hb Seal Rock (*HBA2*:c.427T>G) and Hb Constant Spring. In this case the alpha globin chain is extended by 31 amino acid residues, comparable to the beta chain variant Hb Florida (*HBB*:c.424delC) in which the beta globin chain is elongated. Other mutational mechanisms may interfere with the α and β polypeptide contact points, as in the Hb Quong Sze variant in which the amino acid substitution disrupts the quaternary structure of the haemoglobin protein. These highly unstable alpha chain variants are also associated with haematological features similar to α^+ thalassaemia, in the heterozygous state. Although alpha plus thalassaemia is considered not to be of any clinical significance, it appears that a "dysfunctional" globin gene (that carries a non-deletional mutation) results in a more severe form of thalassaemia than does a deleted gene. Therefore compound heterozygosity for these unstable variants with deletional α^+ or α^0 thalassaemia may cause Hb H disease, and homozygosity can result in a particularly severe form of Hb H disease as they tend to be inherited in a dominant negative fashion. Interestingly these variants commonly occur in those populations where deletional alpha thalassaemia is also prevalent, for example South East Asia.

Another thalassaemic variant which is found at low frequencies in a variety of ethnic groups and is associated with production of unstable mRNA is Hb Lepore**,** in which there is a non-homologous cross over event between the δ and β genes, resulting in the formation of a fusion chain that is missing a portion of the normal β-globin gene. It is synthesised at a lesser rate causing haematological features similar to β-thalassaemia trait

and homozygosity or compound heterozygosity with β-thalassaemia can result in beta thalassaemia major or intermedia.

There are further subgroups of rare but clinically significant haemoglobin variants, in which the underlying mutational mechanism results in a haemoglobin variant with an increased oxygen affinity, compared to the normal haemoglobin molecule. These variants are of clinical interest as they can subsequently lead to polycythaemia (Bain, B: 2006) or a raised haemoglobin concentration, for example, Hb Malmo (*HBB*:c.294 C>A or 294 C>G), which can be found in those ethnic groups where consanguinity is not uncommon, in which case homozygosity may cause severe erythrocytosis.

Other haemoglobin variants that are found at highest frequency are Hb S (*HBB*:c.20A>T), Hb C (*HBB*:c.19G>A), Hb D-Punjab (*HBB*:c.364G>C) and Hb E (*HBB*:c.79G>A) which has already been discussed above, all of which are found within *HBB*. These variants are considered to be of clinical significance and must be detected as part of the antenatal and newborn screening programme as their co-inheritance with one another or with β-thalassaemia can cause a variety of haematological abnormalities ranging from mild to moderately severe anaemia through to thalassaemia major depending of the nature of the variant.

The haemoglobin S variant is caused by a Glutamic acid to Valine amino acid substitution at codon 6 and can cause sickle cell anaemia, in which the red cells become deformed in shape and sickle, and polymerisation of the sickle haemoglobin occurs under conditions of deoxygenation, therefore preventing the delivery of blood flow through the capillaries and therefore delivery of oxygen to the body tissues.

Compound heterozygosity for this variant with β-thalassaemia results in a highly variable clinical phenotype, depending on whether the variant haemoglobin is co-inherited with a β^{+} or β^{0} mutation and the resultant amount of Hb A present. In the former case the clinical phenotype is a mild to moderate sickle cell anaemia whereas in the latter case there is no haemoglobin A present which results in a clinical course similar to homozygosity for this variant $\beta^{s}\beta^{s}$ (sickle cell anaemia).
Furthermore the clinical severity of sickle cell anaemia is related to the proportional percentage of haemoglobin F. There is a clear relationship in individuals with this

genotype between the Hb F percentage and the haplotype of the chromosome carrying the β^S variant. The beta globin haplotype found in parts of Saudi Arabia and India is associated with unusually high levels of fetal haemoglobin of between 10-25% in adults, whereas those haplotypes found in the Senegal area are associated with relatively high levels of Hb F expression of between 7-10%. Whereas those haplotypes found within the Benin, Bantu, and the Cameroon regions of Africa are associated with lower levels of Hb F of between 6-7%. (Thomas, J.J.*et al*: 1998, Bain, B: 2006, Thein, S.L and Menzel, S: 2009, Akinsheye, I *et al* : 2011). These variations of Hb F percentages observed in the sickle genotype may be due to the association of these haplotypes with non-deletional HPFH variants, as described previously in Section 1.4.3. These are important to consider as the haemoglobin F percentage in sickle cell anaemia can be a useful prognostic factor in predicting the response to hydroxycarbamide therapy.

Haemoglobin C, which is caused by a Glutamic acid to Lysine amino acid substitution at Codon 26 is of very little clinical significance by itself in the heterozygous or homozygous state, but its coinheritance with haemoglobin S lead to a sickling disorder comparable to sickle cell anaemia or when co-inherited another β-thalassaemia mutation, especially β^0, it can lead to thalassaemia intermedia, although it is most commonly found in combination with β^+ mutations as they tend to occur in the same ethnic groups.

Again, these disorders are very heterogeneous as there are a large number of different structural haemoglobin variants reported, all of which have different functional properties that can interact with one another or with thalassaemia, resulting in a diverse range of phenotypes, which therefore makes the potential for predicting individual patient outcomes difficult, especially in those countries where compound heterozygous states are common such as Thailand, Sri Lanka and Africa.

1.5 Guidelines for referral

Population screening and pre-natal diagnosis for the haemoglobinopathy disorders is only cost effective when directed at those ethnic groups which are at highest risk, rather than to the general population. Following the implementation of the national NHS Sickle Cell and Thalassaemia (SCT) Screening Programme in England in 2001, the linked Antenatal and Newborn Screening programme aims to offer timely screening to all antenatal women, couples and newborn babies affected by sickle cell, thalassaemia and other haemoglobin

variants to achieve the lowest possible childhood mortality rates for sickle cell disorders in the newborn period.

The UK antenatal screening programme aims to identify those women deemed to be at 'highest risk' of having an affected pregnancy i.e. offspring with a clinically significant haemoglobinopathy, by offering phenotypic screening to women between 8-10 weeks of pregnancy and, if relevant, to their partners, so the couple can make an informed choice regarding their reproduction i.e. offer of prenatal diagnosis or a termination if the offspring affected. It is worth noting that the screening programme(s) are not designed to diagnose all genotypes associated with the haemoglobinopathy disorders, but only to detect those parental carrier state combinations as listed in Appendix 1 that give rise to significant sickle cell disease or β-thalassaemia in the fetus, these include Hb SS, Hb SC, Hb S/O-Arab, Hb S/D-Punjab, Hb S/β thalassaemia, δβ-thalassaemia, Hb Lepore, Hb E/β thalassaemia, β thalassaemia major and Hb Hydrops Fetalis (Ryan, K: 2010). The antenatal screening programme also aims to detect those who are carriers of alpha zero thalassaemia, as this carries the potential risk of having a fetus with the clinically significant disorder, Hb Barts hydrops fetalis.

Indications for testing as part of the antenatal screening programme (http://sct.screening.nhs.uk/policy) are governed by whether the individual lives in a high or a low prevalence area, which is determined by the fetal prevalence of sickle cell disorders: in high prevalence trusts, such as those found within Manchester, there are more than 1.5 babies born with a sickle cell disorder per 10,000 births (Ryan, K. *et al*. 2010). All antenatal women will be offered screening, (See Appendix 2: Testing algorithm for laboratory screening in high prevalence trusts), whereas individuals in low prevalence areas will be screened based on an assessment of individual risk, determined by use of a Family Origin Questionnaire (FOQ) which asks patients about their family origin. This strategy ensures that a large number of people are not unnecessarily tested, in an attempt to identify a very small proportion of patients with a significant abnormality.

Haemoglobinopathy disorders are predominantly inherited in an autosomal recessive manner, therefore offspring of couples who both carry mutations associated with severe alpha or beta thalassaemia have a one in four chance that their children could be affected by a clinically significant haemoglobinopathy disorder, for example if both parents are

carriers of alpha zero thalassaemia, pre-natal diagnosis is recommended to asses the risk of offspring having Hb Barts hydrops fetalis syndrome.

Recommendations for DNA analysis include definitive diagnosis of the main genetic risk combinations as recommended by the antenatal screening algorithm (Appendix 2) and the BCSH guidelines for the diagnosis of significant haemoglobinopathies (Ryan, K *et al*: 2010). Some cases may need to be assessed on an individual basis, for example where there is a family history of a rare abnormality such a particularly severe form of non-deletional Hb H disease resulting in hydrops fetalis syndrome, or a highly unstable variant.

The molecular heterogeneity of these disorders together with the wide range of gene-gene interactions makes accurate interpretation of the clinical phenotypic screen difficult, for example β-thalassaemia can mask co-existing α^0 thalassaemia by normalising the red cell indices. It is therefore important to test those individuals from high risk ethnic groups such as South East Asia for possible underlying alpha thalassaemia, therefore DNA analysis can be used to clarify the diagnosis and accurately asses the risk to the fetus. Additionally, apart from mass spectrometry, DNA analysis is also the only way a definitive identification of an abnormal haemoglobin variant can be achieved.

1.6 Laboratory methods for diagnosis of haemoglobinopathy disorders

A haemoglobinopathy phenotypic screen is often requested as part of a routine antenatal booking or can form part of an investigation for underlying anaemia, haemolysis or microcytosis, alternatively 'screening may also be preoperative, neonatal, antenatal, pre conceptual, pre-marriage or targeted at specific groups perceived to be at risk' (Bain, B: 2011). A haemoglobinopathy screen includes a Full Blood Count (FBC) followed by Haemoglobin electrophoresis by two independent methods; High Performance Liquid Chromatography (HPLC) and Isoelectric Focusing (IEF).

The resultant imbalance of globin chains in the thalassaemia syndromes causes a form of haemolytic anaemia which is reflected in the FBC by a reduced level of Haemoglobin (Hb), Mean Corpuscular Volume (MCV) and Mean Cell Haemoglobin (MCH) with the degree of microcytosis being proportional to the number of genes affected, the red blood cell count (RBC) may also be increased. Cut-off values indicating heterozygosity for

thalassaemia are an MCV <78 fL and MCH <27 pg, an MCH value below 25 pg is used by the antenatal and newborn screening programme to indicate heterozygosity for α^0 thalassaemia.

HPLC is a chromatographic technique which is used to separate out the different components of haemoglobin within a column under high pressure. The analyte (the blood sample) is contained within a liquid (mobile phase) and is forced through a column which is packed with negatively charged small particles (stationary phase) to which the positively charged components of the haemoglobin are absorbed. The individual components of the haemoglobin molecules elute from the column at different times according to their charge and affinity for the stationary phase, they then can then be detected optically. The variant haemoglobins can then be subsequently identified by their characteristic retention time and quantified by measuring the area under the corresponding peak in the chromatogram.

HPLC is used to accurately quantify the percentage of primary haemoglobins A, Hb A_2 and F, and is used to diagnose cases of β-thalassaemia in which there is a characteristic raised Hb A_2 level of above 3.5%. It can also be used to identify those cases in which there is a raised level of Hb F, in which case HPFH, δβ or thalassaemia major may be suspected.

HPLC is also used for the provisional identification of variant haemoglobins, in which the globin chains produced are abnormal in structure but the α/β globin synthesis ratio is often normal leading to normal MCV and MCH values. The common clinically significant haemoglobin variants such as haemoglobin S, C, D^{Punjab}, E, O^{Arab} and Lepore are easily separated and identified using HPLC, although other variants may have similar retention times and may not be identifiable, therefore only haemoglobin variants that alter the charge relative to the normal haemoglobin molecule can be detected by HPLC (Clarke and Thein: 2004) and those which result in a highly unstable haemoglobin variant may go undetected.

Abnormal results should always be confirmed by an independent technique, such as Isoelectric focusing (IEF) or Sickle solubility testing. In those cases where a variant haemoglobin is suspected, DNA studies may be used for a definitive diagnosis.

HPLC forms an important screening tool used by the Newborn and Antenatal screening programme to quantify the haemoglobin fractions Hb H and Hb Barts, shown below as detected by Bio-Rad Variant V-II HPLC instrumentation:

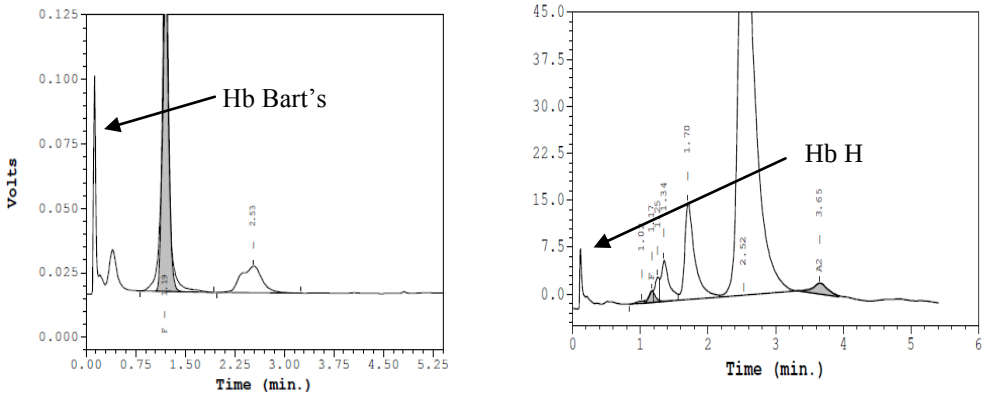


Figure 1.8: Chromatograms showing Hb Bart's and Hb H variants

The newborn screening programme uses a cut-off value of 25% Hb Barts, below which haemoglobin H disease would not be suspected, whereas a baby with Hb Barts hydrops fetalis would usually have no Hb F or A, only Hb Barts.

In reference laboratories within the UK two types of analysis are recommended for the antenatal screening of the haemoglobinopathy disorders. Therefore a second line test must be used in order to validate and confirm the presumed identity of an abnormal variant. Sickle solubility testing may be used for the confirmation of an initial screen suggesting the presence of sickle haemoglobin, alternatively IEF may be used for the confirmation of other common haemoglobin variants. Isoelectric focusing (IEF) is an electrophoretic technique which is used for the identification of haemoglobin variants on an agarose gel, whereby their separation is achieved based on the characteristic Isoelectric point (pl) of each variant i.e. they separate according to their differences in net charge, at a specific pH of solution, when an electrical charge is applied. The various haemoglobin variants form discrete bands that are then visualised by staining and their patterns of migration can be used for the presumptive identification of haemoglobin variants. If this technique is used as a confirmatory technique, then control haemoglobins must also be run on each gel.

Further biochemical tests may be warranted in order to evaluate the patients iron status, including Serum Ferritin, Serum Iron and Zinc Protoporphyrin (ZPP) which are important in order to exclude underlying iron deficiency, as the haematological parameters for thalassaemia and iron deficiency are quite similar (Harteveld, C.L and Higgs, D.R: 2010) and iron studies are important for a differential diagnosis.

Peripheral blood film examination is indicated when an abnormality of globin synthesis is suspected and can be very useful when an initial screening test detects an abnormality. A thalassaemic blood smear usually shows a variable degree of microcytosis and hypochromia, nucleated red blood cells and basophilic stippling in cases of β-thalassaemia. Other characteristic cells may also be observed if the individual has a haemoglobin variant, such as sickled or crescent cells in sickle cell disease, target cells with haemoglobin C variant and basophilic stippling in those with the haemoglobin Constant Spring variant.

Further staining the peripheral blood cells with 1% Brilliant Cresyl Blue can help visualise inclusion bodies that have attached and precipitate into the circulating red blood cells of patients with Hb H disease, this causes damage to the red cell membrane giving them a typical 'golf-ball' appearance (Bain, B: 2006, Aguilar-Martinez and Gulbis: 2010, Harteveld, C.L and Higgs, D.R: 2010).

1.7 Project summary

In those cases where a haemoglobinopathy disorder is suspected, the initial investigation is based on haematological and biochemical analysis methods which aim to identify the thalassaemia syndrome or variant(s) present. However, interpretation of this phenotypic screening data can often be ambiguous where there is scope for complex interactions between genotypes. Providing antenatal couples with an accurate genetic risk assessment for their offspring on a phenotypic basis alone can be difficult. DNA analysis should permit a definitive characterisation of the complex genetic interactions commonly found in these disorders.

Although the haemoglobinopathies are heterogeneous at the molecular level, there are common subsets of mutations associated with geographical regions and ethnic groupings, categorised as the Mediterranean, Asian Indian, Southeast Asian and African populations. Sub populations within each of these areas have their own spectrum of abnormal variants and mutations, as well as a number of rarer mutations found at much lower gene frequencies.

Traditionally the strategy for genetic identification of alpha or beta thalassaemia mutations within a diagnostic DNA laboratory has been 'ethnic-led' i.e. targeted diagnosis by allele-

specific Polymerase Chain Reaction (PCR) according to the individuals' ethnicity. The increased racial heterogeneity of the UK population has significantly increased the spectrum of haemoglobinopathy mutations present, resulting in a higher number of α- and β- thalassaemia defects than in any other of the 60 countries with a published spectrum of mutations (Henderson, S. *et al*. 2009). This increases the number of different mutations and possible interactions which a genetic diagnostic service needs to detect and the 'ethnic-led' screening approach becomes less effective, as it means that rarer or previously uncharacterised mutations may be missed. To illustrate: a study by Fisher *et al* (2003) highlights the importance of detecting underlying ααα alpha thalassaemia in those individuals with beta thalassaemia and the implications of having only a targeted mutational screening approach in complex ethnic populations, in which a screening programme in Sri Lanka, if based on the detection of the three most common mutations would leave 'up to 25% of alleles unidentified' (Fisher *et al*: 2003).

In order to achieve a high mutation detection rate in the heterogeneous population of the North West of England, and Greater Manchester in particular, a comprehensive range of laboratory assays are required to ensure that a wide spectrum of defects can be detected, ranging from large deletions to point mutations. In this study, a range of PCR based methods were employed for the detection of large deletions associated with alpha thalassaemia or the HPFH and δβ disorders. Direct DNA sequencing of the alpha and beta globin genes will also be developed for the detection of point mutations associated with thalassaemia and haemoglobin variants, eliminating the need for lengthy targeted mutation screening protocols.

1.8 Project aim and hypothesis

To develop a comprehensive range of laboratory methodologies to detect different classes of mutations associated with alpha and beta thalassaemia, as part of a haemoglobinopathy genetic diagnostic service, in the heterogeneous population of the North West of England. The spectrum of mutations detected within the referral population will then be analysed in order to determine their frequency and distribution to produce a data set for this group.

The frequency and range of the mutations detected, in those referrals where the ethnicity is known, will then be compared to published data on expected mutation frequencies. The genotypes detected will also be analysed in terms of phenotype and genotype correlations

in order to identify those cases that are not fully explained with a view to suggesting further possible service development. The true genetic heterogeneity of the referral population can then be accurately evaluated and the effectiveness of our more comprehensive first pass screening approach assessed in relation to an alternative, ethnic led, screening strategy.

Chapter 2: Materials and Methods

2.1 Patient selection and Phenotypic Screening

Between July 2008 and December 2011 a range of diagnostic samples, referred as part of the existing national antenatal haemoglobinopathy screening programme were analysed at the Molecular Diagnostics Centre (MDC) within Manchester Royal Infirmary which is CPA accredited. Appropriate consent for genetic analysis was obtained for all patient specimens included as part of the diagnostic service provided for this disorder.

All of the patients within this group had haemoglobinopathy phenotypic screening performed, which included a full blood count (FBC) performed internally at MRI on a Sysmex XE2100 analyser and HPLC on a BioRad Variant II Haemoglobin testing system, run on the β-thal short programme (Variant/Variant II Beta thalassaemia short program kit) or by the external referral centre performed on different instrumentation. Patients were suspected to have a haemoglobinopathy disorder if either a haemoglobin variant was detected by HPLC and/or FBC analysis revealed an MCH of less than 27 pg (normal reference range 27-32 fL) and an MCV of less than 80 fl (normal reference range 80-97 fL). An MCH level of below <25 pg would be indicative for alpha zero thalassaemia. The above red cell changes may also be accompanied by a characteristic increase of Hb A_2 to 3.5% or above (normal reference range 2.3-3.3%) detected by HPLC, in which case beta thalassaemia would be suspected, and in the HPFH and δβ disorders, a persistently raised Hb F level to 4% and above. Where appropriate, in order to accurately assess a patients' phenotype, their iron status was assessed by measurement of their Serum Protein level, in order to exclude iron deficiency anaemia. Further genetic analysis was only carried out on those patients in whom a haemoglobinopathy disorder was suspected and in those that met the criteria as outlined by the SCT testing algorithm in high prevalence areas (see Appendix 2), although exceptions were made in some instances where DNA studies were used to clarify a patient's clinical diagnosis.

2.2 Sample processing for genetic testing

The investigations in this project were carried out on twenty mL of whole blood collected from each adult patient in 4x5 mL Vacutainer tubes (Becton Dickenson, UK) containing 7.2mg of EDTA in each 4 mL tube. For young children and babies only 1mL of whole blood was collected. All samples met the minimum labelling requirements as outlined by

the national screening programme and CMFT trust protocols, including a Family Origin Questionnaire and appropriate consent for further genetic investigation.

2.3 Extraction of genomic DNA

Genomic DNA was extracted from 400µL of whole blood using the semi-automated Promega Maxwell®16 DNA system (Promega, UK). The Maxwell® platform uses individual reagent cartridges and purification plungers that extract DNA by utilising paramagnetic-bead based technology. Firstly the cells were lysed in the presence of chaotropic and detergent agents and then the nucleic acids bind to magnetic silica particles. Then through a series of wash steps the lysed cellular components are removed, along with any other bound particles or contaminants before the nucleic acids are finally eluted into 200µL sterile water. The Maxwell® 16 system was suitable for the rapid (30-40 minutes) semi-automated extraction of DNA from up to 16 whole blood samples at a time. It produced a high quality yield of up to 150ng/µL of DNA, suitable for downstream applications such as those used in this study. For the purpose of this study, when prompted by the machine the protocol selected was "Blood", and the sample type selected was "DNA".

3.0 Deletion analysis by Gap-PCR protocols

Gap-PCR (amplification across the breakpoints of a deletion) provided a simple diagnostic test for the identification of deletions, provided the breakpoint sequences were known such as those associated with thalassaemia including common alpha thalassaemia deletion mutations, alpha gene duplications and other globin gene deletions, such as those associated with the HPFH disorders.

Two primers complementary to the sense and antisense strands regions of the DNA were used that flank the extent of the deletion; amplifying a deletion-specific fragment. The normal allele was detected by amplifying across one of the breakpoints, using one primer complementary to part of the deleted sequence and one complementary to the flanking sequence. For small deletions, typically less than one kilobase in size, the primer pairs generated two products, with the smaller fragment arising from the deletion allele. For larger deletions, the distances between the two flanking primers were too great to amplify the normal DNA fragment and the only product obtained was from the flanking/internal primer combination, representing the deleted fragment. Heterozygosity was detected by

the presence of both the normal and deleted products in a single reaction (See figure 2.2). The different sized bands were visualised by agarose gel electrophoresis.

3.1 Multiplex Gap-PCR for the detection of deletional forms of alpha thalassaemia

Multiplex Gap-PCR provided an essential screening tool for the rapid detection of a range of alpha thalassaemia deletions in the heterogeneous referral population found within the North West of England.

Primers specific for the seven most common alpha thalassaemia deletions were multiplexed so they could be simultaneously detected in a single Gap-PCR reaction. The two methods employed within the MDC were based on adapted protocols derived from Chong, S.S *et al*: 2000 and Liu, Y.T *et al*: 2000. They were capable of detecting heterozygosity, homozygosity and compound heterozygosity for the most common forms of deletional alpha thalassaemia including the alpha plus thalassaemia $\alpha^{-3.7}$ and $\alpha^{-4.2}$ deletions, the alpha zero thalassaemia $(\alpha)^{-20.5}$, $--^{SEA}$, $--^{MED}$, $--^{THAI}$ and $--^{FIL}$ deletions (Arnold, S *et al*: 2001, Baysal, E. and Huisman, T.H.J: 1994, Dodè, C *et al*: 1992, Chong, S.S *et al*: 2000, Bowden, D.K *et al*: 1992). Furthermore these assays encompassed the detection of the British α^{0} deletion (--BRIT), which is prevalent in our referral population (Eng, B *et al*: 2009). Additionally the alpha globin gene duplication (ααα) could also be detected using the Liu, Y.T *et al (*2000) methodology.

Two independent multiplex Gap-PCR assays were used in our genetic screening service so one assay could be used as a confirmatory technique if a deletion is detected by the other.

3.2 Method Validation

Methods adopted were confirmed by the analysis of known control samples carrying relevant deletions obtained from the John Radcliffe Oxford University, Kings College and University College London hospitals. The alpha zero $-\alpha^{20.5}$ deletion could not be independently validated due to a lack of DNA control sample.

3.3 Multiplex Gap-PCR assay based on Liu, Y.T *et al*: 2000 protocol

This protocol was used for the detection of the seven most common forms of deletional alpha thalassaemia and the α-globin gene triplication. Three separate multiplexes were utilised; the first multiplex detected the five most common alpha zero deletions: $-(\alpha)^{20.5}$, --

^{SEA}, --^{MED}, --^{THAI} and –^{FIL}. The second multiplex detected the alpha plus $\alpha^{-3.7}$ deletion and the α-globin gene triplication. The third multiplex detected the alpha plus $\alpha^{-4.2}$ deletion. This protocol helped to overcome the inherent problems associated with the amplification of the GC-rich α-globin gene cluster, by the use of the enhancing agents Betaine and Dimethyl Sulphoxide (DMSO), producing a reliable and reproducible PCR assay suitable for diagnostic use.

Figure 2.1 summarises the positions of the multiplexed primers used in this protocol used for the amplification of specific deletions.

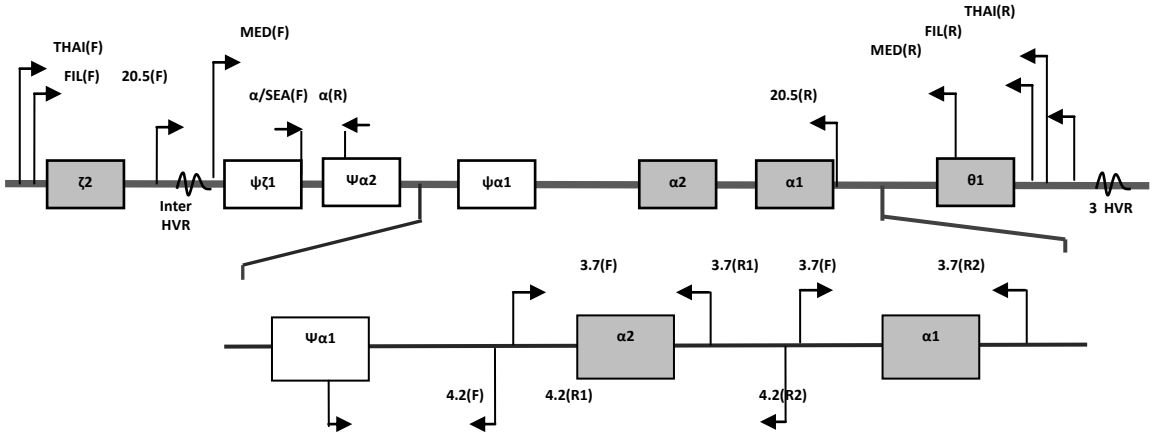


Figure 2.1: Location of the multiplexed PCR primers in the α-globin gene cluster (the genes are represented as grey boxes, derived from Liu, Y.T *et al:*2000).

Table 2.1 (a to c), below summarises the sequence of the primers, used in each of the three multiplex PCR reactions, and their corresponding final concentrations.

(a) Multiplex-1 Primers:

Primer Name	Description	Primer Sequence 5' to 3'	Final Conc. (μM)
1	20.5-F	GGGCAAGCTGGTGGTGGTGTTACACAGCAACTC	0.1
2	20.5-R	CCACGCCCATGCCTGGCACGTTTGCTGAGG	0.1
3	alpha SEA-F	CTCTGTGTTCTCAGTATTGGAGGGAAGGAG	0.3
4	alpha-R	TGAAGAGCCTGCAGGACCAGGTCAGTGACCG	0.15
5	MED-F	CGATGAGAACATAGTGAGCAGAATTGCAGG	0.15
6	MED-R	ACGCCGACGTTGCTGCCCAGCTTCTTCCAC	0.15
7	SEA-R	ATATATGGGTCTGGAAGTGTATCCCTCCCA	0.15
8	FIL-F	AAGAGAATAAACCACCCAATTTTTAAATGGGCA	1.6
9	FIL-R	GAGATAATAACCTTTATCTGCCACATGTAGCAA	1.6
10	Thai-F	CACGAGTAAAACATCAAGTACACTCCAGCC	0.1
11	Thai-R	TGGATCTGCACCTCTGGGTAGGTTCTGTACC	0.1

(b) Multiplex-2 Primers:

Primer	Description	Sequence 5' to 3'	Final Conc. (μM)
12	3.7-F	AAGTCCACCCCTTCCTTCCTCACC	0.3
13	3.7-R1	ATGAGAGAAATGTTCTGGCACCTGCACTTG	0.1
14	3.7-R2	TCCATCCCCTCCTCCCGCCCCTGCCTTTTC	0.1

(c) Multiplex-3 Primers

Primer	Description	Sequence 5' to 3'	Final Conc. (μM)
15	4.2-F	TCCTGATCTTTGAATGAAGTCCGAGTAGGC	0.8
16	4.2-R1	TGGGGGTGGGTGTGAGGAGACAGGAAAGA GAGA	0.4
17	4.2-R2	ATCACTGATAAGTCATTTCCTGGGGGTCTG	0.4

Tables 2.1 (a-c): Sequence of the primers used in: (a) Multiplex-1, (b) Multiplex-2 and (c) Multiplex-3.

3.3a: Multiplex PCR Amplification

1. A separate PCR amplification was performed for each of the three multiplexes using 0.5mL thin-walled PCR tubes (Peqlab, UK).

2. 100-200ng of genomic DNA was added to each reaction.

3. A "no DNA" (blank control) was always included in the PCR to indicate if cross contamination had occurred.

4. A known deletion control (positive control) was included in the PCR if a mutation from a patient sample was being confirmed.

5. A master mix was made for each of the individual multiplex reactions, each containing different reagents, primer concentrations and volumes of water.

Detailed below in Tables 2.2 (a-d) are the reagents used their corresponding volumes for each multiplex:

Multiplex-1 (α^0) PCR:

A total reaction volume of 25μL comprised of:

Reagent	Volume per reaction (μL)
10x PCR Buffer (Life Technologies, UK)	2.5
$MgCl_2$ (Life Technologies, UK)	1.5
Betaine (Sigma, UK)	3.75
DMSO (Sigma, UK)	1.25
Ampli*Taq* Gold DNA Polymerase (Life Technologies)	0.25
dNTP's (Bioline, UK)	0.5
Primers: (*total multiplex volume*):	*10.5*

(Eurofins, Germany)	
20.5-F (5µM)	0.5
20.5-R (5µM)	0.5
alpha SEA-F (5µM)	1.5
alpha-R (5µM)	0.75
MED-F (5µM)	0.75
MED-R (5µM)	0.75
SEA-R (5µM)	0.75
FIL-F (20µM)	2
FIL-R (20µM)	2
Thai-F (5µM)	0.5
Thai-R (5µM)	0.5
Sterile H_2O	3.75
DNA	1

Table 2.2a: Components of the Multiplex-1 PCR reaction master mix.

Multiplex-2 ($\alpha^{-3.7}$) PCR:

The different reagents used for Multiplex 2 reflect the changes made from the original protocol and subsequent assay optimisation needed for the successful amplification of both the $-\alpha^{3.7}$ deletional and normal DNA fragments. Initially an Extensor Hi-Fidelity PCR Enzyme Mix (Thermo Scientific, UK) was used, but the assay reliability was improved by use of a KAPA 2G Robust PCR kit (Kapa Biosystems, Africa). A total reaction volume of 25µL comprised of:

Reagent	**Volume per reaction (µL)**
5X KAPA (2G) GC Buffer (Kapa Biosystems, Africa)	5.0
KAPA2G Robust DNA polymerase (Kapa Biosystems, Africa)	0.5
dNTP Mix *(10mM each dNTP)* (Kapa Biosystems, Africa)	0.5
Primers: (*total multiplex volume*) (Eurofins, Germany)	*2.5*
3.7-F (5µM)	1.5
3.7-R1 (5µM)	0.5
3.7-R2 (5µM)	0.5
Sterile H_2O	14.5

Table 2.2b: Components of the Multiplex-2 ($\alpha^{-3.7}$) PCR reaction master mix.

Multiplex-2 (ααα) PCR:

Multiplex 2 was further optimised in order to overcome the failure of the multiplex to detect the alpha gene duplication (ααα). Therefore a separate PCR methodology was devised using only the two primers specific for this gene duplication.

A total reaction volume of 50µL comprised of:

Reagent	Volume per reaction (µL)
10x PCR Buffer (Qiagen, UK)	5.0
Q-Solution (Qiagen, UK)	10
HotStar *Taq* Plus (Qiagen, UK)	0.25
dNTP's (Bioline, UK)	4.0
Primers: (Eurofins, Germany)	
3.7-F (5µM)	1.5
3.7-R1(5µM)	1.5
Sterile H_2O	25.75
DNA	2.0

Table 2.2c: Components of the Multiplex-2 (*ααα*) PCR reaction master mix

Multiplex-3 ($\alpha^{-4.2}$) PCR:

This multiplex was optimised using the same Qiagen reagents as above for detection of the triplicated alpha genotype (Table 2.2c), due to consistent failure to amplify products within the PCR reaction when the original protocol was used. A total reaction volume of 50µL comprised of:

Reagent	Volume per reaction (µL)
10x PCR Buffer	5.0
Q-Solution	10
HotStar *Taq* Plus	0.25
dNTP's	4.0
Primers:	
4.2-F (5µM)	8.0
4.2-R1 (5µM)	4.0
4.2-R2 (5µM)	4.0
Sterile H_2O	12.75
DNA	2.0

Table 2.2d: Components of the Multiplex-3 ($\alpha^{-4.2}$) PCR reaction master mix

Each of the multiplex PCR reactions were then transferred to a Techne TC-3000 (Bibby Scientific, UK) thermal cycler in order to perform the PCR amplification step. The thermal cycling conditions for each of the reactions are shown below.

Multiplex-1 and Multiplex-3 were run under the same thermal cycling conditions consisting of: Initial 15-minute denaturation at 95°C, followed by 35 cycles of 95°C denaturation for 60 seconds, 65°C annealing for 60 seconds and 72°C extension for 150 seconds. A final 10-minute extension at 72°C completed the reaction.

Multiplex-2 (for the detection of $\alpha^{-3.7}$ deletion) was run by preheating the block to 94°C before inserting the reaction tubes. Then an initial 3-minute denaturation at 94°C was performed, followed by 35 cycles of 94°C denaturation for 60 seconds, 65°C annealing for 60 seconds, and 72°C extension for 150 seconds. A final 10-minute extension at 72°C completed the reaction.

Modified Multiplex-2 (for the detection of ααα alpha) was run by preheating the block to 96°C before inserting the reaction tubes. Then an initial 15-minute denaturation at 96°C, followed by 30 cycles of 98°C denaturation for 45 seconds, 60°C annealing for 90 seconds, and 72°C extension for 135 seconds. A final 5-minute extension at 72°C completed the reaction.

3.3b Agarose gel electrophoresis:

A 1.0% agarose gel (w/v) was prepared by adding 1.0g of molecular grade agarose (Peqlab, UK) to 100mL of 0.6 x TBE buffer in a conical flask and mixed. It was then heated in the microwave on full power for 3 minutes. Three µL of 10mg/mL ethidium bromide was then added to the gel and left to cool to approximately 60°C. A gel tray was prepared with a 1mm 16 well comb and the gel was then poured into the tray and allowed to cool for a minimum of 30 minutes or until it had set.

Once the thermal cycler had finished, the PCR amplification products were prepared for loading on to the gel by addition of 4µL of PCR product to 2µL of loading dye in a separate 0.5 mL tube. A total of 6µL of each sample was loaded into individual wells of the gel and a 1kb UltraRanger DNA ladder (Norgen Biotek, Canada) was added to the wells either side of the PCR product loading. The gel was left to run at 110 volts for a minimum of 1.5 hours. The banding pattern on the gel was then visualised and photographed on a Gene Genius gel documentation system (Syngene, UK). The banding patterns observed for this Gap-PCR assay are shown below.

3.3 c: Interpretation of Banding Patterns Observed

PCR Fragment sizes obtained from the alpha globin gene in Multiplex-1, 2 and 3 are shown below in Figure 2.2 (a-c):

Multiplex-1 (α^0):

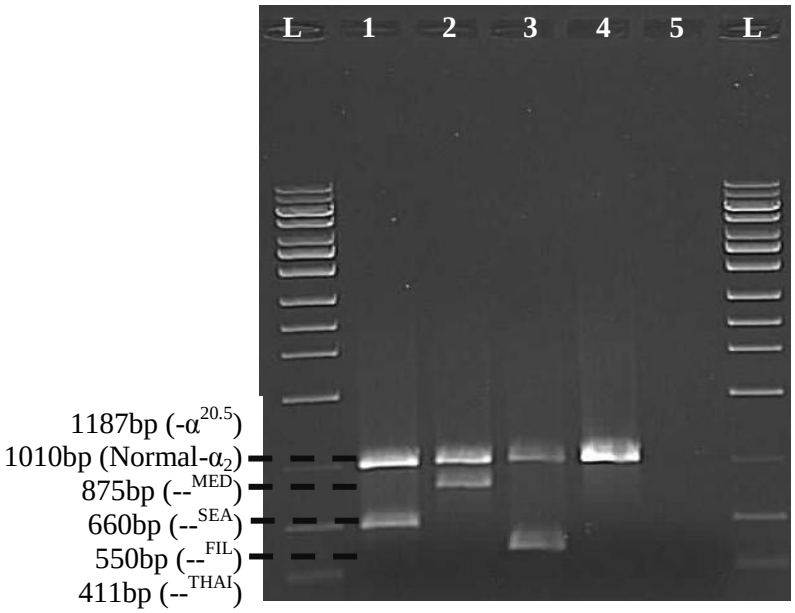


Figure 2.2a: Banding pattern observed for Multiplex-1 (for the detection of α^0 thalassaemia deletions) Run with UltraRanger 1kb DNA ladder (labelled 'L').

The presence of a single band at 1010 bp (for the specific α^2 fragment) indicated that the individual tested negative for the five alpha zero deletions and was therefore was considered normal for these genotypes (as in Lane 4). In the presence of an additional specific deletional fragment, this indicated heterozygosity for that particular deletion, occurring at either of the following positions: 660 bp ($--^{SEA}$) as in lane 1, 875 bp ($--^{MED}$) as in lane 2, 411 bp ($--^{THAI}$), 550 bp ($--^{FIL}$) as in lane 3 and 1187 bp ($-\alpha^{20.5}$) which is not shown in Figure 2.2a above. Lane 5 is a blank 'no DNA' control.

Conversely, the absence of the α^2 fragment and the presence of a single deletion fragment indicated the individual was homozygous or compound heterozygous for deletional alpha thalassaemia.

Multiplex-2 ($\alpha^{-3.7}$):

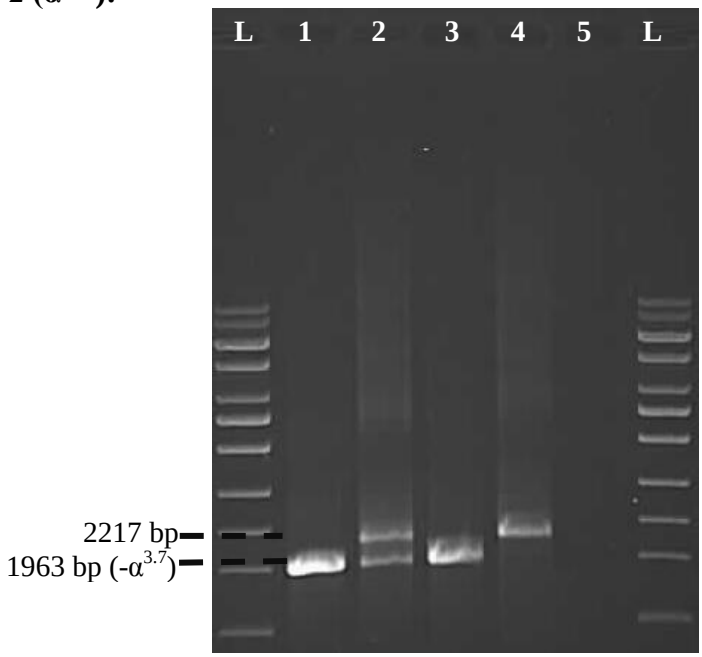


Figure 2.2b: Banding pattern observed for Multiplex-2 (for the detection of the $\alpha^{-3.7}$ deletion). Run with UltraRanger 1kb DNA ladder (labelled 'L')

A single band at 2217bp indicated that the $-\alpha^{3.7}$ deletion was not present, and the individual was normal for this genotype (as in Lane 4).

A band at 1963bp indicated that the deletion was present, and in the presence of the normal band at 2217bp indicated heterozygosity for this deletion (as in Lane 2). When the normal band at 2217bp was absent, this indicated homozygosity for this deletion (as in Lanes 1 and 3). Lane 5 is a blank 'no DNA' control

Multiplex-2 (ααα):

The banding pattern observed for the triplicated alpha genotype is shown below:

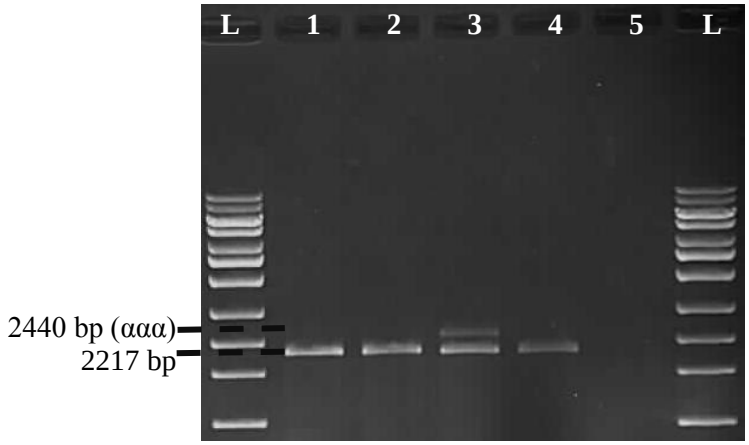


Figure 2.2c: Banding pattern observed for Multiplex-2 (for the detection of ααα alpha). Run with UltraRanger 1kb DNA ladder (labelled 'L')

As in Figure 2.2c a single band at 2217 bp indicated that the gene triplication was not present, and the individual was normal for this genotype (as in Lanes 1, 2 and 4).

A band at 2440 bp indicated the alpha gene triplication was present and when found in the presence of the normal band at 2217 bp this indicated heterozygosity for this genotype (as in Lane 3). When the normal band at 2217 bp was absent, this indicated homozygosity for the alpha gene triplication (not shown). Lane 5 is a blank 'no DNA' control.

Multiplex-3 ($\alpha^{-4.2}$):

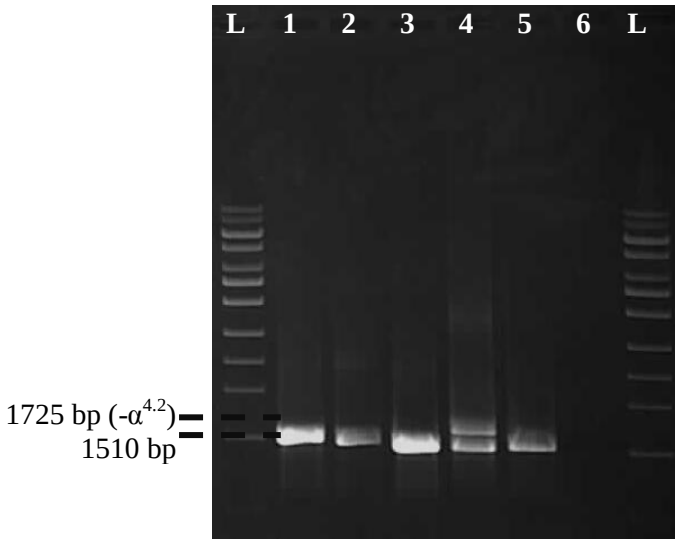


Figure 2.2d: Banding pattern observed for Multiplex-3 (for the detection of the $\alpha^{-4.2}$ deletion). Run with UltraRanger 1kb DNA ladder (labelled 'L').

A single band at 1510 bp indicated the $-\alpha^{4.2}$ deletion was not present, and the individual was normal for this genotype (as in Lanes 1, 2, 3 and 5).

A band at 1725 bp indicated that the $-\alpha^{4.2}$ deletion was present and when found in the presence of the normal band at 1510 bp this indicated heterozygosity for this deletion (as in Lane 4). When the normal band at 1510 bp was absent this indicates homozygosity for this deletion (not shown). Lane 6 is a blank 'no DNA' control.

The following points were considered when interpreting the above banding patterns.

The presence of the specific α^2 fragment indicated at least one intact α^2 gene and thus in the presence of an additional deletional fragment indicated heterozygosity for that specific deletion. Conversely, the absence of the α^2 fragment and the presence of a single deletional fragment indicated that the individual was homozygous for that α-thalassaemia deletion. The presence of two deletional fragments of different size indicated compound

heterozygosity for deletional alpha thalassaemia. The patient was considered negative for the above deletions when the α^2 fragment was present and there were no additional deletional fragments, although this does not entirely exclude alpha thalassaemia as a rare alpha thalassaemia deletion or a non deletional form of alpha thalassaemia may still be present. It is important that these results were interpreted alongside an individual's phenotypic and clinical data.

3.3 d Limitations of the Liu, Y.T *et al*: 2000 protocol

The Liu based protocol did not provide a truly multiplexed assay which was capable of detecting the seven common forms of deletional alpha thalassaemia in a single reaction. All of the three individual multiplex reactions required optimisation and different running conditions. Subsequently the Chong based protocol, discussed below was developed in order to improve and streamline the analysis for alpha thalassaemia.

3.4 Multiplex Gap-PCR assay based on Chong, S.S *et al*: 2000 protocol

This protocol utilised one multiplex reaction for the detection of the seven most common forms of deletional alpha thalassaemia.

As above, this protocol overcomes the difficulties associated with amplification of the alpha globin genes by PCR such as sequence homology and high GC content by use of a Qiagen HotStar *Taq* plus DNA polymerase PCR kit (Qiagen, UK) that contains PCR additives, such as a proprietary "Q-solution" which facilitates the amplification of difficult templates by destabilising secondary structures and reducing the energy required to melt the GC-rich alpha globin region.

In this protocol a multiplex of 16 different primers were used in a single PCR reaction, their sequences and final concentrations are detailed below in Table 2.3:

Primer Name	**Primer Sequence (5'-3')**	**Final Conc.** (μM)
LISI-F	ATACCATGGTTACCCCATTGAGC	0.5
LISI-R	AGGGCTCATTACATGTGGACCC	0.5
α2/3.7-F	CCCCTCGCCAAGTCCACCC	0.2
α2-R	AGACCAGGAAGGGCCGGTG	0.2
20.5-F	GCCCAACATCCGGAGTACATG	0.2
3.7/20.5-R	AAAGCACTCTAGGGTCCAGCG	0.2
4.2-F	GGTTTACCCATGTGGTGCCTC	0.5
4.2-R	CCCGTTGGATCTTCTCATTTCCC	0.5

SEA-F	CGATCTGGGCTCTGTGTTCTC	0.2
SEA-R	AGCCCACGTTGTGTTCATGGC	0.2
FIL-F	TTTAAATGGGCAAAACAGGCCAGG	1.0
FIL-R	ATAACCTTTATCTGCCACATGTAGC	1.0
MED-F	TACCCTTTGCAAGCACACGTAC	0.2
MED-R	TCAATCTCCGACAGCTCCGAC	0.2
THAI-F	GACCATTCCTCAGCGTGGGTG	0.3
THAI-R	CAAGTGGGCTGAGCCCTTGAG	0.3

Table 2.3: Multiplex primer sequences and final concentrations for the Chong Gap-PCR protocol.

3.4a Multiplex PCR Amplification

A PCR reaction was set up (as detailed in steps 1-5 in Section 3.3a). Detailed in Table 2.4 below are the reagents used their corresponding volumes for this multiplex.

A total reaction volume of 50µL comprised of:

Reagent	**Volume per reaction (µL)**
10xPCR Buffer (Qiagen, UK)	5.0
Q-Solution (Qiagen, UK)	10
HotStar *Taq* Plus (Qiagen, UK)	0.25
dNTP's (Bioline, UK)	4.0
Primers *(10µm)* (*total multiplex volume*) (Eurofins, Germany)	*31*
LIS1-F	2.5
LIS1-R	2.5
α2/3.7-F	1.0
α 2-R	1.0
20.5-F	1.0
3.7/20.5-R	1.0
4.2-F	2.5
4.2-R	2.5
SEA-F	1.0
SEA-R	1.0
FIL-F	5.0
FIL-R	5.0
MED-F	1.0
MED-R	1.0
THAI-F	1.5
THAI-R	1.5
DNA	**2.0**

Table 2.4: Components of the "Chong" PCR reaction master mix.

The multiplex PCR reactions were then transferred to a Techne TC3000 thermal cycler (Bibby Scientific, UK) in order to perform the PCR amplification step. The thermal cycling conditions selected started with an initial 15-minute denaturation at 96°C,

followed by 30 cycles of 98°C denaturation for 45 seconds, 60°C annealing for 90 seconds, and 72°C extension for 135 seconds. A final 5-minute extension at 72°C completed the reaction.

Agarose gel electrophoresis was carried out as detailed in step 3.3b.

3.4b Interpretation of Banding Patterns Observed

PCR Fragment sizes observed for the alpha globin gene using the Chong based protocol are shown below in Figure 2.3:

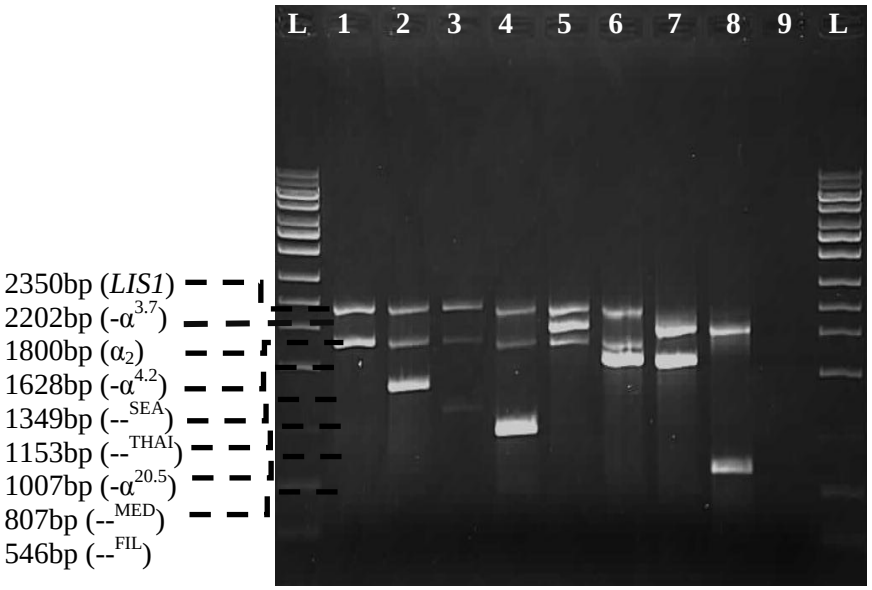


Figure 2.3: Banding pattern observed for 'Chong' PCR. Run with UltraRanger 1kb DNA ladder (labelled 'L')

An individual testing negative for the above 7 deletions will showed two bands: one at 2350bp one at 1800bp (as in Lane 1). In the presence of an additional specific deletional fragment, this indicates heterozygosity for that particular deletion, occurring at either of the following positions: 1349bp (--^{SEA}) as in lane 2, 1153bp (--^{THAI}) as in lane 3, 1007bp (-$\alpha^{20.5}$) as in lane 4, 2202bp (-$\alpha^{3.7}$) as in lane 5, 1628bp (-$\alpha^{4.2}$) as in lane 6 and 546bp (--^{FIL}) (not shown in Figure 3 above). Lane 9 is a blank 'no DNA' control.

The following points were considered when interpreting the above banding patterns. Unlike the Liu based protocols (discussed in Section 3.3c) this protocol is based on the amplification of the α^2 gene which may be totally or partially removed by any of the 7 deletions being tested for. Therefore the presence of the α^2 gene fragment (represented by a band at 1800bp) is used to indicate heterozygosity when any deletion allele is present. Conversely its absence indicates the α^2 globin gene is absent from both chromosomes 16

and can indicate homozygosity or compound heterozygosity for either of the 7 forms of deletional alpha thalassaemia tested for, as in Lane 7 ($-\alpha^{3.7}/-\alpha^{4.2}$) and Lane 8 ($-\alpha^{3.7}/--^{MED}$). Furthermore independent control primers on the *LIS1* gene act as a control for quality of amplification within each individual PCR reaction, and is represented by a band at 2350 bp.

3.5 Gap-PCR assay for the detection of the British (--BRIT) alpha zero deletion based on Eng, B *et al*: 2009 protocol.

This protocol utilised a Gap-PCR reaction for the detection of the --BRIT alpha zero deletion, using the Qiagen HotStar *Taq* plus DNA polymerase PCR kit (as in Section 3.4). Two different primer pairs were used in one reaction: one which spanned the $--^{BRIT}$ breakpoint and amplified the deletion specific fragment and the second which serves a control and amplified a larger fragment from the PAFAH1B1 gene on chromosome 17. Their sequence and final concentrations are detailed below in Table 2.5:

Primer Name	Primer Sequence (5'-3')	Final Conc. (µM)
LIS-1F	ATACCATGGTTACCCCATTGAGC	0.4
LIS1-R BRIT	TTATGTAATGCACATTGCACATCCC	0.4
BRIT-F	CAGGTGTCCATCATCAGGACTAAC	0.4
BRIT-R	CCTTCACCACCACCTGTGTAGG	0.4

Table 2.5: Primer sequences and final concentrations for the British alpha Gap-PCR

3.5 a: Gap-PCR amplification

A PCR reaction was set up (as detailed in steps 1-5 in Section 3.3a).

Table 2.6 below details the reagents used and their corresponding volumes for this assay. A total reaction volume of 50µL comprised of:

Reagent	Volume per reaction (µL)
10xPCR Buffer (Qiagen, UK)	5.0
Q-Solution (Qiagen, UK)	10
HotStar *Taq* Plus (Qiagen, UK)	0.25
dNTP's (Bioline, UK)	4.0
Primers: *(10µm)* (Eurofins, Germany)	
LIS1-F	2.0
LIS1R-BRIT	2.0
BRIT-F	2.0
BRIT-R	2.0
Sterile H_2O	20.75

Table 2.6: Components of the British alpha deletion PCR reaction master mix.

The PCR reaction was then transferred to a Techne TC3000 thermal cycler (Bibby Scientific, UK) in order to perform the PCR amplification step. The thermal cycling conditions selected started with an initial 15-minute denaturation at 96°C, followed by 30 cycles of 98°C denaturation for 45 seconds, 60°C annealing for 90 seconds, and 72°C extension for 135 seconds. A final 5-minute extension at 72°C completed the reaction. Agarose gel electrophoresis was carried out as detailed in step 3.3b.

3.5b Interpretation of Banding Patterns Observed

PCR Fragment sizes observed for the alpha globin gene using the Eng *et al* based protocol are shown below in Figure 2.4:

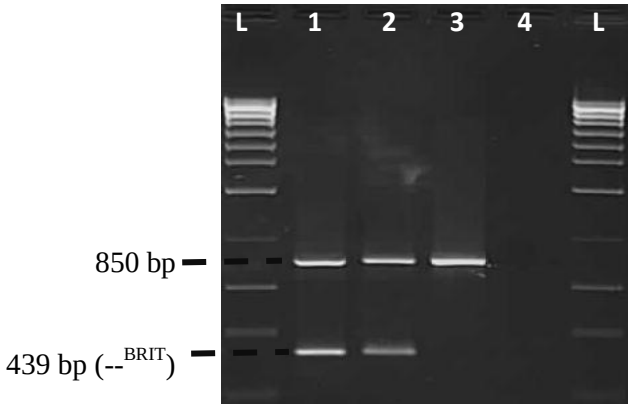


Figure 2.4: Gel image of banding patterns observed for British alpha deletion Gap-PCR.

An individual testing negative for the $--^{BRIT}$ deletion will one band at 850 bp (as in Lane 3). A heterozygote for this deletion will show an additional deletional specific fragment at 439 bp ($--^{BRIT}$) (as in Lanes 1 and 2). Lane 4 is a blank 'no DNA' control.

Gap-PCR methodology was also used for the detection of the nine most common deletions that involve the β-globin gene complex on chromosome 11 and are associated with clinical manifestations of HPFH and δβ-thalassaemia disorders (Craig, R.A *et al*: 1994).

4.0 Gap-PCR for the detection of deletional forms of HPFH and δβ-thalassaemia

A protocol described by Craig, R.A *et al*: 1994 was used for the rapid molecular detection of the nine known deletions associated with the HPFH and δβ-thalassaemia disorders. This was based on Gap-PCR methodology (as described in Section 2.3) in which each specific deletion is detected by amplification products generated by combinations of three specific primers within the same reaction, leading to the production of a unique deletion-specific product in the presence of a deletion and a normal control band of a different size in the

presence of the normal allele. This protocol can detect the following deletions: HPFH-1, HPFH-2, HPFH-3, Spanish $(\delta\beta)^0$, Haemoglobin Lepore, Sicilian $(\delta\beta)^0$, Chinese $^G\gamma(^A\gamma\delta\beta)^0$, Asian-Indian inversion deletion $^G\gamma(^A\gamma\delta\beta)^0$ and the Turkish inversion deletion $(\delta\beta)^0$.

Figure 2.5 summarises the PCR primers used in this protocol and their relative positions along the β-globin gene cluster.

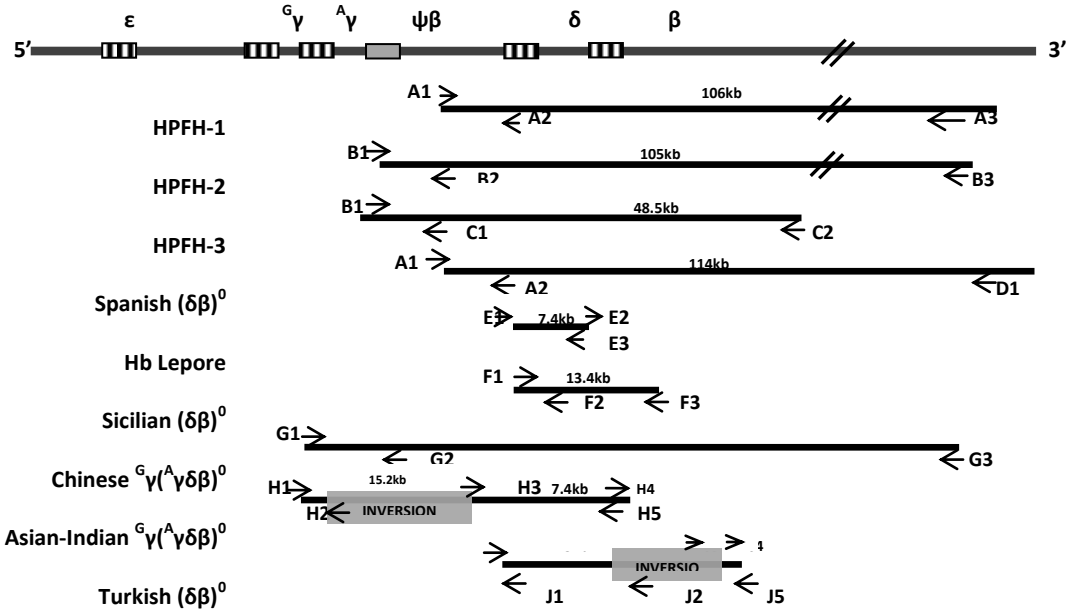


Figure 2.5: Location of the PCR primers along the β-globin gene cluster, showing the location of the primers and extent of deletions detected in the Craig, R.A *et al* Gap-PCR protocol (Derived from Craig, R.A *et al:* 1994).

Table 2.7 below summarises the sequence of the primers used in this protocol.

Primer Name	Sequence (5'-3')
A1	AGAATGTCACACTTAGAATCTG
A2	CACTTTAATTCTGGTCTACCTGAA
A3	ACTGTGATGTTGGAAATGGAC
B1	GACATGGACTATTGTTCAATGA
B2	TGCTATGCCAACTCACTACC
B3	TTTATATATGAAATGCTACTGATT
C1	CTTTGCTGTTCAGGCTTAATTT
C2	GACACAGAGCAGTGATTGGTGCA
D1	AGCCCGCCACTGCACTGTG
E1	GACACACATGACAGAACAGCCAAT
E2	CGATCTTCAATATGCTTACCAAG
E3	CATTCGTCTGTTTCCCATTCTA
F1	TTGGGTTTCTGATAGGCACTG
F2	GTGTCACCCATTAATGCCTTGTAC
F3	TAGATCCCTTTGCCATTATG

G1	GGCATATATTGGCTCAGTCA
G2	CTTGCAGAATAAAGCCTATC
G3	TCAACAAUATCAACATTACACC
H1	ATGCCATAAAGCACCTGGATG
H2	GAGCTGAAGAAAATCATGTGTGA
H3	TAACCATATGCATGTATTGCC
H4	CAATGTATCATGCCTCTTTGCAC
H5	GCAGCCTCACCTTCTTTCATGG
J1	GAAGAGCAGGTAGGTAAAAGAACC
J2	TTCCACTATCTTACTTACACAT
J3	CAAAGCAGCAATACTAAACAGGAG
J4	TGCTGAGGATTGTTTTAGGTC
J5	GAATAGCAGTGGTGAGAGAG

Table 2.7: Primer sequences used in the Craig *et al:* 1994 Gap-PCR protocol.

4.1 Gap-PCR amplification

All PCR reactions were performed using the KAPA2G (GC) robust DNA polymerase PCR kit (Kapa Biosystems, South Africa), which contains a highly robust and sensitive second-generation enzyme and a GC buffer which has been specifically designed for the amplification of difficult templates, therefore providing a reliable assay for the detection of the deletions in Figure 2.5.

1. Separate PCR amplification reactions were performed for each specific deletion being tested for, using 0.5mL thin-walled PCR tubes (Peqlab, UK).
2. 100-200ng of genomic DNA was added to each reaction.
3. A "no DNA" (blank control) was always included in the PCR to indicate if cross contamination had occurred
4. A known deletion control (positive control) was included where possible.
5. A master mix was made for each of the individual multiplex reactions, each containing different reagents, primer concentrations and volumes of water.

All separate PCR reactions were set up in a 25µL total reaction volume containing: 5µL of 5X KAPA (2G) GC Buffer, 0.5µL of KAPA2G DNA polymerase, 0.5µL of dNTP mix (10mM each dNTP), 1.5µL of each of the three deletion specific primers (at a concentration of 5µM, with the exception of the HPFH-2 deletion control primer B2 which was used at a concentration of 1µM). Volumes of $MgCl_2$ and subsequent volumes of sterile water were optimised for each of the individual reactions, as detailed below in Table 2.8a.

PCR Reagents Volume (µL per reaction)			
Deletion	**$MgCl_2$ (25mM)**	**Deletion specific primers**	**sH_20**
HPFH-1	0.5	A1, A2, A3	12
HPFH-2	1.5	B1, B2, B3	11
HPFH-3	2	B1, C1, C2	10.5
Spanish $(\delta\beta)^o$	2.5	A1,A2, D1	10
Hb Lepore	1.5	E1, E2, E3	11
Sicilian $(\delta\beta)^o$	1.5	F1, F2, F3	11
Chinese $^G\gamma(^A\gamma\delta\beta)^0$	1.5	G1, G2, G3	11
Asian-Indian Inv/Del A $^G\gamma(^A\gamma\delta\beta)^0$	1.5	H1, H2, H3	11
Asian-Indian Inv/Del B $^G\gamma(^A\gamma\delta\beta)^0$	1.5	H2, H4, H5	11
Turkish Inv/Del A $(\delta\beta)^o$	1.5	E1, J1, J3	11
Turkish Inv/Del B $(\delta\beta)^o$	1.5	J2, J4, J5	11

Table 2.8a: Component volumes for each of the PCR master mix reaction.

Each PCR reaction was then transferred to a Techne TC3000 thermal cycler (Bibby Scientific, UK) in order to perform the PCR amplification step. The thermal cycling conditions for each of the reactions were then optimised. All reactions commenced with an initial denaturation for 2 minutes, followed by 30 cycles of denaturation, annealing and extension at varying temperatures for each PCR reaction, the details of which are given below in Table 2.8b.

	PCR thermal Cycler Conditions								
	PCR	**Denaturation**		**Annealing**		**Extension**		**Final Ext.**	
Deletion	product size (bp)	Temp (°C)	Time (sec)	Temp (°C)	Time (sec)	Temp (°C)	Time (min)	Temp (°C)	Time (min)
HPFH-1	1,193 *(mutant)* 1,616 *(normal)*	95	20	54	20	72	1	72	1
HPFH-2	1,950 *(mutant)* 2,206 *(normal)*	95	20	55	20	72	1	72	1
HPFH-3	607 *(mutant)* 841 *(normal)*	95	20	58	20	72	1	72	1
Spanish $(\delta\beta)^o$	1,281 *(mutant)* 1,616 *(normal)*	95	60	58	60	72	1	72	10

Hb Lepore	777 *(mutant)* 915 *(normal)*	95	60	58	60	72	2	72	2
Sicilian $(\delta\beta)^o$	1,150 *(mutant)* 1,585 *(normal)*	95	20	60	20	72	1	72	1
Chinese $^{G}\gamma(^{A}\gamma\delta\beta)^0$	508 *(mutant)* 682 *(normal)*	95	60	58	60	72	2	72	2
Asian-Indian Inv/Del A $^{G}\gamma(^{A}\gamma\delta\beta)^0$	327 *(mutant)* 1,195*(normal)*	95	60	58	60	72	1	72	10
Asian-Indian Inv/Del $^{G}\gamma(^{A}\gamma\delta\beta)^0$ B	371 *(mutant)* 665 *(normal)*	95	60	58	60	72	1	72	10
Turkish Inv/Del A $(\delta\beta)^o$	432 *(mutant)* 742 *(normal)*	95	20	65	20	72	1	72	1
Turkish Inv/Del B $(\delta\beta)^o$	489 *(mutant)* 700 *(normal)*	95	20	57	20	72	1	72	1

Table 2.8b: Details of PCR thermal cycler running conditions for each deletion.

The PCR products from each reaction were separated by agarose gel electrophoresis as detailed in step 3.3b.

4.2 Interpretation of Banding Patterns Observed

The PCR fragment sizes observed for each specific deletion are detailed in Table 2.8b above, and illustrated for the HPFH-1, HPFH-2, HPFH-3, Hb Lepore and Chinese $^{G}\gamma(^{A}\gamma\delta\beta)^0$ deletions below in figures 2.6a to 2.6e:

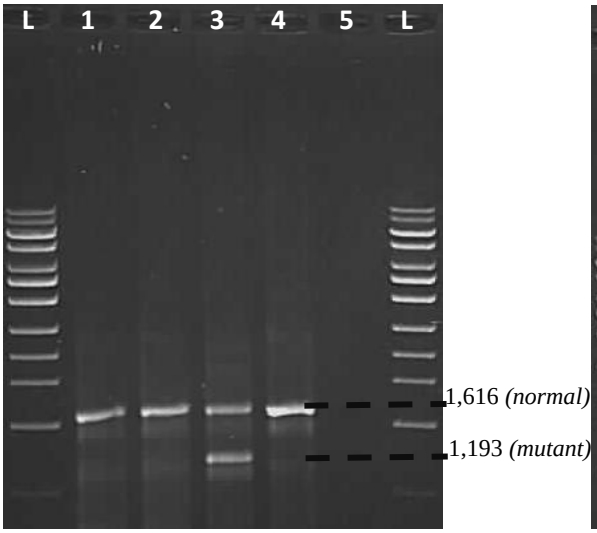


Figure 2.6a: Gel image of banding pattern observed for HPFH-1 deletion.

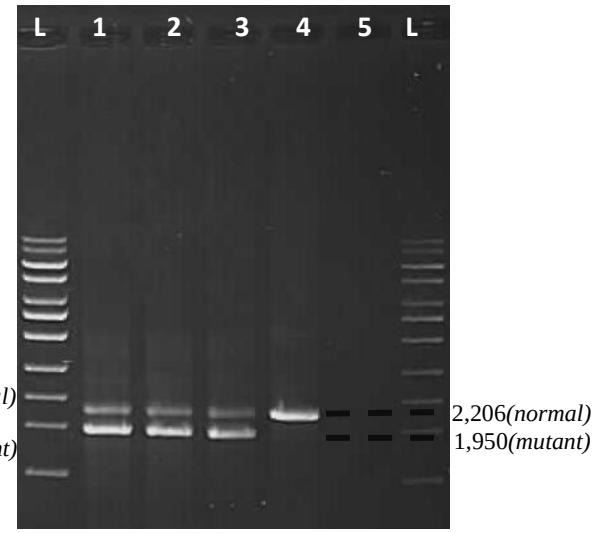


Figure 2.6b: Gel image of banding pattern observed for HPFH-2 deletion.

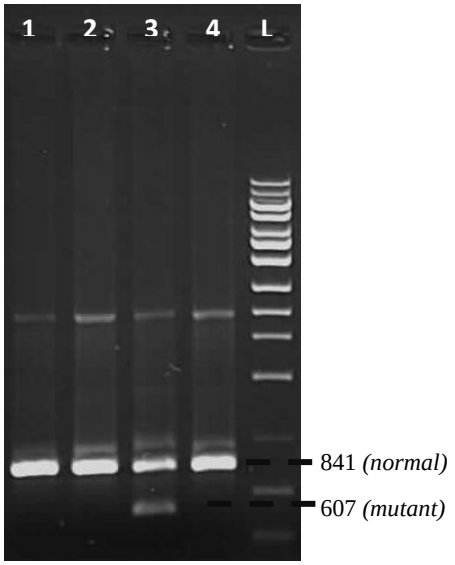


Figure 2.6c: Gel image of banding pattern observed for HPFH-3 deletion.

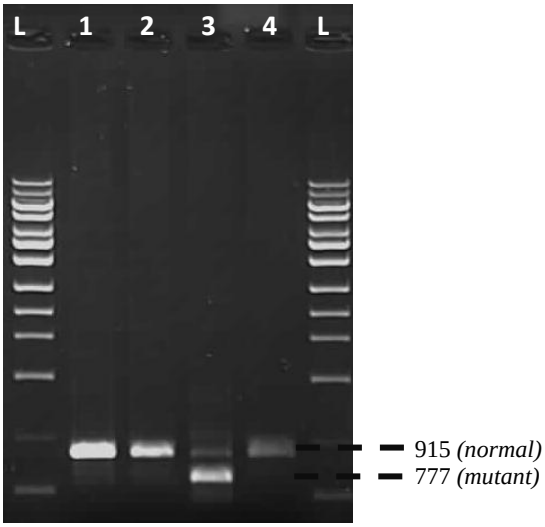


Figure 2.6d: Gel image of banding pattern observed for Hb Lepore deletion

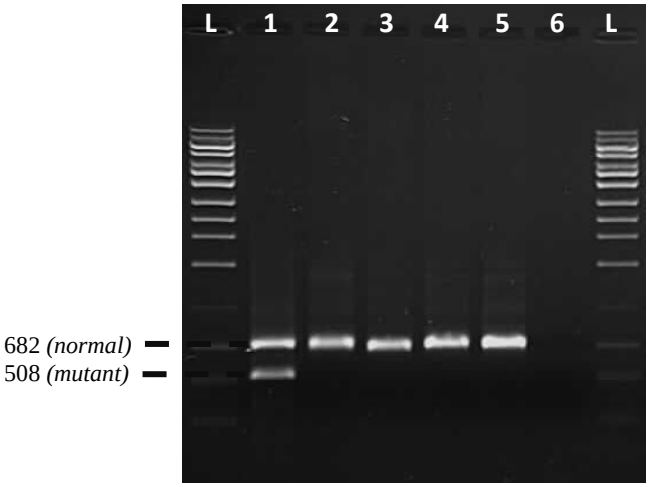


Figure 2.6e: Gel image of banding pattern observed for Chinese $^{G}\gamma(^{A}\gamma\delta\beta)^{0}$ deletion.

5.0 DNA sequencing for the detection of point mutations and small deletions associated with thalassaemia and haemoglobin variants.

DNA sequencing allows the analysis of the exact nucleotide sequence of the gene under study. Direct DNA sequencing of the alpha and beta globin genes represents the "gold standard" for mutation analysis, allowing a single-pass detection method for most of the point mutations associated with thalassaemia and haemoglobin variants, eliminating the need for lengthy targeted mutation screening protocols.

5.1a PCR Amplification for the generation of *HBB* gene specific products

This protocol identified mutations by full DNA sequence analysis of the essential regions of *HBB*, which was amplified in three overlapping fragments using three sets of primer pairs.

As figure 2.7 shows; the 5' promoter region, exon 1, intron 1, exon 2 and the 5' end of intron 2 were amplified in one amplicon (referred to as Product 1). Most of intron 2 were amplified in another amplicon (referred to as Product 2) and the remaining 3' end of intron 2, exon 3 and the 3' UTR were amplified in another amplicon (referred to as Product 3).

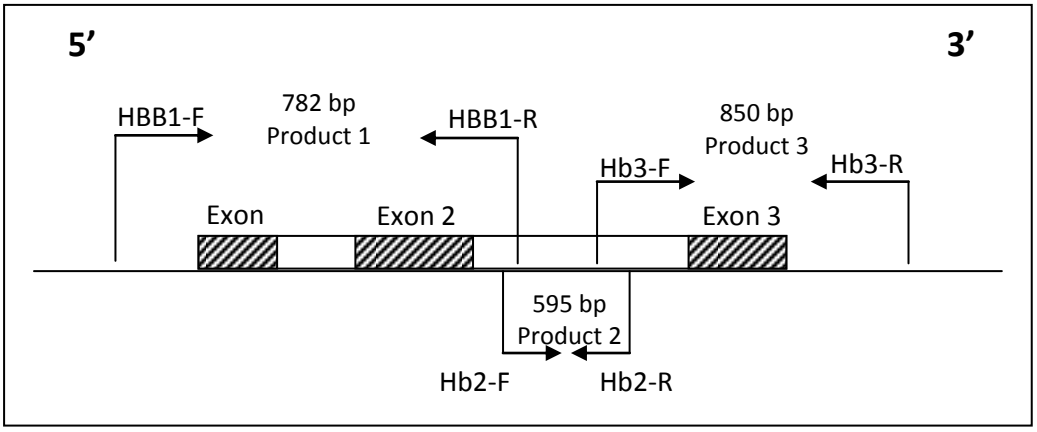


Figure 2.7: Location of PCR primer pairs along the beta-globin gene.

5.1b Primer Design

Three overlapping primers sets were designed using the online Primer 3 software (http://frodo.wi.mit.edu/) to conform to the following parameters: an annealing temperature of between 55-60°C, absence of significant hairpin loop formation, lack of secondary priming sites and low specific binding at the 3' end of each prime pair in order to avoid mispriming.

The picked primers were then checked for underlying SNP's using the online software SNPcheck, Version 2 (http://ngrl.manchester.ac.uk/SNPCheckV2/snpcheck).

The primers were tailed with N13 (modified from M13 universal primer sequences) forward or reverse sequences respectively so that one set of forward and reverse primers could be used in subsequent cycle sequencing reactions.

Their sequences are detailed below in Table 2.9.

***HBB* Frag. No.**	**Primer sequence 5' to 3'**	**Annealing Temp. (°C)**	**Product size (inc. N13 tails) (bp)**
1	**Hb1F**: *gtagcgcgacggccagt*ACTCCTAAGCCAGTGCCAGA **Hb1R**: *cagggcgcagcgatgac*AAAACGATCCTGAGACTTCCA	60	782
2	**Hb2F**: *gtagcgcgacggccagt*TTTCTTTCCCCTTCTTTTCTATGG **Hb2R**: *cagggcgcagcgatgac*AATGGTGCAAAGAGGCATGA	60	595
3	**Hb3F**: *gtagcgcgacggccagt*TGGAATATATGTGTGCTTATTTG CAT **Hb3R**:*cagggcgcagcgatgac*TCTTTATGTTTTAAATGCA CTGACCTC	60	850

Table 2.9: Primer sequences for the amplification of *HBB* and their corresponding annealing temperatures. Common N13F and N13R tailed sequences are shown in italics.

Primer Name	**5' to 3' primer sequence**
N13F	*gtagcgcgacggccagt*
N13R	*cagggcgcagcgatgac*

Table 2.10: N13 forward (N13F) and N13 reverse (N13R) primer sequences for use in cycle sequencing reactions.

See Appendix 3 for *HBB* genomic DNA sequence with primer locations.

5.1.c Full sequence *HBB* PCR set up

A 25 µL PCR reaction was set up using Invitrogen *Taq* DNA polymerase kit (Invitrogen by Life Technologies, UK) in 0.5mL thin-walled PCR tubes to which 2.0µL of patient DNA (100-200ng) was added, except in the "no DNA" (blank control) to which 2.0 µL of sterile H_2O was added.

Three individual master mixes were prepared, for each of three overlapping fragments of the *HBB* gene. Each contained the following: 2.5µL of 10X Invitrogen PCR Buffer, 0.75µL of 50mM $MgCl_2$, 2.0µL of 10mM dNTP mix, 0.1µL Invitrogen *Taq* polymerase, 2.0µL each of forward and reverse primer and 13.65µL of sterile H_2O.

Each of the PCR reactions were then transferred to a Techne TC-3000 thermal cycler (Bibby Scientific, UK) and the 'Beta.cyc' amplification programme was selected. The thermal cycling conditions consisted of an initial 3-minute denaturation at 94°C followed by 30 cycles of 94°C denaturation for 60 seconds, 60°C annealing for 45 seconds, and 72°C extension for 1 minute and a final 5-minute extension at 72°C.

5.1d Agarose gel electrophoresis

Prior to DNA sequencing 4 µL of each PCR product was added to 2µL of loading dye, then analysed on a 1.0% agarose gel (w/v), to check that the amplification had been successful (See Section 3.3b).

5.1e Purification of the PCR products by Microclean

The amplification products, representing the essential regions of *HBB* sequence were then purified using *micro*CLEAN (WebScientific, UK). *Micro*CLEAN is a one-tube DNA clean-up reagent which is used to purify double stranded DNA from reaction buffers, enzymes, dNTPs and primers.

20µL of *micro*CLEAN was added to each PCR DNA product then mixed by vortexing, and left at room temperature for 5 minutes. The tubes were then spun (Jouan SA microcentrifuge) at 70g for 5 minutes. All the supernatant was removed using a pipette and then the tubes were spun down again at 12000 rpm for 2 minutes. The dregs were removed using a pipette and the pellets were re-suspended in 20µL of sterile H_2O and left for a minimum of 5 minutes in order to allow the DNA to rehydrate before proceeding to the next cycle sequencing step.

5.1f Cycle sequencing

Dye terminator cycle sequencing uses dideoxy chain termination chemistry in order to generate terminated DNA strands needed for sequence determination from single stranded DNA templates (Sanger, *et al:* 1977).

During the cycle sequencing reaction the PCR template is terminated at every point along the strand due the random incorporation of a matching dideoxy nucleotide triphosphates (ddNTP). These are A, T, G and C base analogues containing base specific flurophores which are capable of emitting light at distinct wavelengths. These ddNTPs have a hydroxy

(OH) group missing at the 3' position (the location where one nucleotide would normally attach to another to form a chain) therefore in the absence of this OH group additional nucleotides cannot be added to the chain, thus interrupting chain elongation. The Big-Dye® Terminator v3.1 Cycle sequencing Kit (Life Technologies, UK) contains all the components required for cycle sequencing, including ddNTPs, dNTPs, Amli*taq* DNA polymerase FS, $MgCl_2$ and reaction buffer. N13 forward and reverse primers were used (Table 2.10) in the appropriate cycle sequencing reactions in order to sequence the PCR product containing the N13 tails, allowing the PCR products to be sequenced using utilising only one sequencing primer set (forward and reverse).

Cycle sequencing reaction set up

The cycle sequencing reaction was carried out in a 96-well microtitre plate (Thermo Scientific, UK). On ice 4µL of each PCR product was aliquotted from the individual PCR tubes (representing each amplified segment of the gene), into each corresponding well of the plate.

Two separate cycle sequencing reaction master mixes were prepared for each of the forward and reverse cycle sequencing reactions, for the appropriate number of reactions using the 1x reaction mix, to make up a total volume of 11µL (as detailed in Table 2.11).

Reagent	Volume per reaction (µL)
Big Dye Terminator v3.1 (Life Technologies, UK)	1.0
5x Cycle Sequencing Buffer (Life Technologies, UK)	2.5
N13F or N13R primer (2µM)	1.5
sH_20	6.0

Table 2.11: Cycle sequencing reaction mix

11µL of the N13 forward master mix was added to each forward reaction in the plate, containing the DNA templates, similarly 11µL of the N13 reverse master mix was added to each reverse reaction in the plate. The plate was sealed with a silicone septum, briefly centrifuged then transferred to an Primus 96 thermal cycler (MWG Biotech, Germany) run on the 'B.Buffer' cycle, which consisted of an initial 3-minute denaturation at 96°C, followed by 25 cycles of: 96°C denaturation for 15 seconds, 50°C annealing for 10 seconds, and 60°C extension for 2 minutes, then a final hold at 8°C. The samples were then stored, in the fridge at 4°C, until ready to proceed to the next purification step.

5.1g Purification of cycle sequencing products by ethanol precipitation

Purification of the cycle sequencing reaction products was essential prior to sequencing analysis in order to remove unincorporated dye terminators. This was achieved by precipitating the cycle sequencing reactions directly in the 96 well plate using EDTA and ethanol precipitation.

The plate (from step 3.1d) was removed from either the thermal cycler or the fridge and briefly spun at 40g for 30 seconds. Then, using a multi-channel pipette, 5µL of 125mM EDTA (Sigma-Aldrich, UK) (pH 8) and 70µL of 95% ethanol (Severn Biotech, UK) was added to each well. The plate was then sealed with a septum (Life Technologies, UK) then briefly vortexed and then incubated at room temperature for 15 minutes. The plate was then centrifuged (Thermo Jouan BR4i, UK) at 2600g for 30 minutes to precipitate the sequencing products. The seal was then removed and the plate inverted on to a paper towel and centrifuged at 170g for 1 minute to remove the supernatant. The purified products were then washed as follows: 60µL of 70% ethanol (Severn Biotech, UK) was added to each well, the plate was resealed with the septum then centrifuged for 15 minutes at 1600g.

The plate was then again inverted on to a paper towel and centrifuged at 170g for 1 minute to ensure all the supernatant was removed. The plate containing the purified cycle sequencing product was then wrapped in aluminium foil and incubated at room temperature for 15 minutes to ensure all the ethanol had evaporated, leaving only the DNA pellet. The plate was then re-sealed with the silicon septum and stored at -20°C or until ready for sequencing analysis.

5.2 DNA sequencing analysis by Capillary Electrophoresis

The samples were prepared for DNA sequencing by re-suspension of the purified sequencing pellets in "Hi-Di" Formamide (Life Technologies, UK). 15µL of Hi-Di was added to each well containing purified cycle sequencing products and the plate sealed with a septum and then briefly centrifuged at 40g for 30 seconds. The plate was then placed on a thermal cycler in order to denature the sequencing products by performing a 5-minute incubation at 94°C. Immediately after the programme had finished the plate was placed on ice in order to prevent the formation of single stranded DNA secondary structures, then

the plate was briefly centrifuged again at 40g for 30 seconds, ensuring the Formamide was at the bottom of each well and that no bubbles were present.

The 96 well plate containing the denatured cycle sequencing products was loaded onto the 3130*xl* DNA Sequencer (Life Technologies, UK) and the appropriate patient details for the plate were entered in the 'Plate Editor' window. The relevant parameters were selected: Instrument protocol: 'FastSeq_POP7' and analysis method 'Sequence analysis'. During electrophoresis, the dye-labelled DNA fragments are injected into the 16 channel 50cm capillary array by electro-kinetic injection. When an electrical current is applied they move through the capillary, containing POP-7 liquid polymer (Life Technologies, UK) and are separated according to their size. As the terminated cycle sequencing fragments, each containing one of the four base-specific flurophores at the end of their terminated sequence pass through a fixed point in the electrophoretic capillary they are excited by a laser, the detector then records the florescent signal i.e. the characteristic wavelength at which each fluorophore emits light and a base name is assigned accordingly (A,T,G or C) to that point in the DNA sequence. As a result the DNA sequence data from the gene of interest is stored as an electropherogram from which further analysis can be performed.

5.3 DNA Sequence data analysis

Once the sequencing had been completed, the sequencing traces were provisionally reviewed and cropped appropriately using Applied Biosystems DNA sequencing Analysis Software version 5.2 (Life Technologies, UK). These .ab1 data files were then transferred to another computer for further mutation analysis using Mutation Surveyor software (SoftGenetics, USA).

5.4 Analysis of sequencing data using Mutation Surveyor Software

Further mutation detection analysis was performed using Mutation Surveyor version 3.97 software (http://www.softgenetics.com/mutationSurveyor.html). This software was used to simultaneously perform variant analysis and mutation detection from Sanger sequencing .ab1files, generated by the ABI Capillary DNA Sequencer. It employs a range of algorithms to align and compare both the forward and reverse electropherograms, generated from the GenBank reference files against the patient traces, checking for dissimilar and similar peaks that are present. When the software detects a difference in sequence between the reference and patient sequences i.e. a heterozygote or homozygote

mutation(s), it is indicated by a peak in the comparison window, as shown in figure 2.8 below:

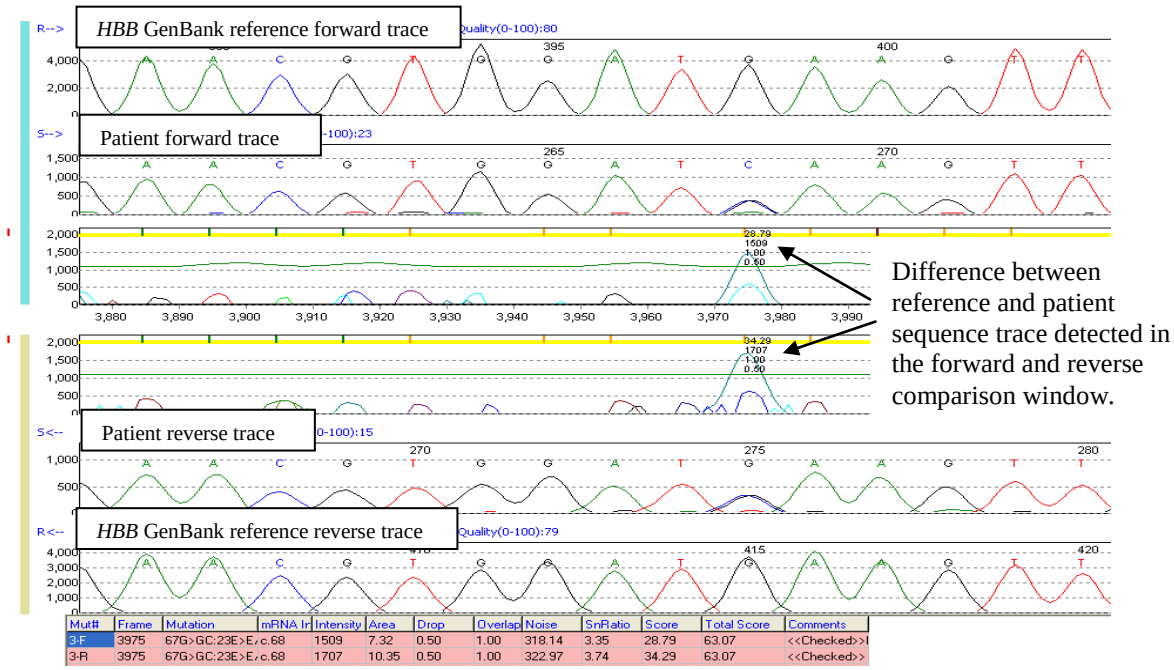


Figure 2.8: Mutation detected using the Mutation Surveyor software

Figure 2.8 illustrates a patient DNA sequence trace compared against a reference DNA sequence trace. The middle rows show the comparison window highlighting the difference between the reference and patient sequence traces, indicating the presence of a heterozygous point mutation.

See Section 5.8 for details of the *HBB* GenBank reference sequence used.

5.5 PCR Amplification for the generation of *HBG2* ($^{G\gamma}$) gene specific products

This protocol was used to identify non-deletional HPFH mutations and the Xmn*1* $^{G}\gamma$ -158 (C>T) restriction site polymorphism, by full DNA sequence analysis of the essential regions of the *HBG2* gene, which was amplified in a single fragment using one primer pair.

See Section 5.8 for details of the *HBG2* GenBank reference sequence used.

5.5a Primer Design

One primer pair (forward and reverse) was designed to amplify the following regions of *HBG2*, the primers were tailed with N13 (modified from M13 universal primer sequences)

forward or reverse sequences respectively (as detailed previously in Table 2.10). Their sequences are detailed below in Table 2.12.

***HBG2* Primer**	**Primer sequence 5' to 3'**	**Annealing Temp. (ºC)**	**Product size (inc. N13 tails) (bp)**
XmnF	*gtagcgcgacggccagt* GAACTTAAGAGATAATGGCCTAA	60	655
XmnR	*cagggcgcagcgatgac* TAGTCCAGACGCCATGGGTCA	60	655

Table 2.12: Primer sequences for the amplification of *HBG2* and their corresponding annealing temperatures. Common N13F and N13R tailed sequences are shown in italics. See Appendix 4 for *HBG2* genomic DNA sequence with primer locations.

5.5b Full sequence *HBG2* PCR set up

A 25µL PCR reaction was set up using Invitrogen *Taq* DNA polymerase kit (Invitrogen by Life Technologies, UK) in 0.5mL thin-walled PCR tubes to which 2.0µL of patient DNA (100-200ng) was added, except in the "no DNA" (blank control) to which 2µL of sterile H_2O was added.

One master mix was prepared, each containing: 2.5µL of 10X Invitrogen PCR Buffer, 0.75µL of 50mM $MgCl_2$, 2.0µL of 10mM dNTP mix, 0.1µL Invitrogen *Taq* polymerase, 2.0µL each of Xmn forward and reverse primer and 13.65µL of sterile H_2O.

Each of the PCR reactions were then transferred to a Techne TC-3000 thermal cycler (Bibby Scientific, UK) and the 'Beta.cyc' amplification programme was selected. The thermal cycling conditions consisted of an initial 3-minute denaturation at 94°C followed by 30 cycles of 94°C denaturation for 60 seconds, 60°C annealing for 45 seconds, and 72°C extension for 1 minute and a final 5-minute extension at 72°C.

Agarose gel electrophoresis was then carried (as detailed in Section 3.3b) in order to check that the amplification had been successful, the PCR products were then purified by *micro*CLEAN (as detailed earlier in Section 5.1e). Cycle sequencing, purification of the cycle sequencing products by ethanol precipitation and then DNA sequencing analysis by Capillary Electrophoresis and DNA Sequence data analysis using Mutation Surveyor Software was carried out as detailed in sections 5.1f to 5.4 above.

See Section 5.8 for details of the *HBG2* GenBank reference sequence used and the primer locations.

5.6 PCR Amplification for the generation of *HBA1* and *HBA2* gene specific products

This protocol identified mutations by full DNA sequence analysis of the essential regions of the *HBA1* and *HBA2* genes. This enabled the identification of non-deletional forms of alpha thalassaemia and alpha globin chain variants. *HBA1* and *HBA2* were amplified as single amplicons using a common 5' forward primer in combination with a specific 3' alpha-1 or alpha-2 reverse primer. Due to the high sequence homology between *HBA1* and *HBA2*, it was essential to check the design of the specific reverse primers to ensure that the correct alpha globin gene was amplified.

The size of the *HBA1* and *HBA2* gene PCR products are shown below in Figure 2.9.

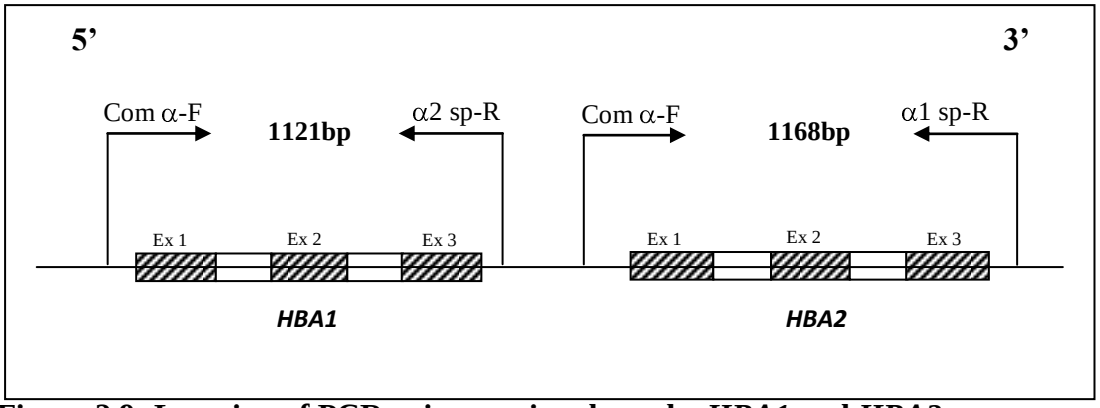


Figure 2.9: Location of PCR primer pairs along the *HBA1* and *HBA2* genes.

The sequences of the primers used are detailed below in Table 2.13.

PCR Primer	Primer sequence 5' to 3'	Annealing Temperature (ºC)
Common α-F	*gtagcgcgacggccagt*GTGGAGGGT GGAGACGTCCTG	66.5
Specific α1-R	*cagggcgcagcgatgac*GAGAGGTTC TAGCCATGTGTG	57.4
Specific α2-R	*cagggcgcagcgatgac*CCATTGTTGG CACATTCCGG	66.6

Table 2.13: Primer sequences and their corresponding annealing temperatures, used for the amplification of *HBA1* and *HBA2* (the common N13 sequences are shown in italics).

This PCR protocol was optimised based on a protocol used by King's College Hospital NHS Trust. Initially a Thermo Scientific Reddymix PCR kit (Thermo Fisher Scientific, UK) was used that combined of all the essential components needed for a PCR including a gel-loading dye, into one single master mix. This allowed the completed PCR sample to be directly loaded onto an agarose gel therefore eliminating many pre- and post- PCR steps.

Due to manufacturing problems encountered with this PCR kit, limiting its supply this assay was adapted using a KAPA 2G Robust PCR Kit (Kapa Biosystems, Africa), as described above for the HPFH and Multiplex-2 assays.

5.6 a Full sequence *HBA1* and *HBA2* PCR set up

Separate PCR reaction mixes for the amplification of the alpha 1 and 2 genes was required, with each master mix containing either the α-1 or α-2 specific primers, along with the common forward primer.

25µL PCR reactions were performed using 0.5mL thin-walled PCR tubes (Peqlab, UK), to which 2.0µL of patient DNA (100-200ng) was added, except in the "no DNA" (blank control) to which 2µL of sterile H_2O was added. Two master mix reactions were made for each of alpha globin gene reactions and for the appropriate number of sample required, each containing: 5µL of 5x KAPA2G GC PCR Buffer (*with* $MgCl_2$), 0.5µL of KAPA 10mM dNTP mix, 0.3µL of KAPA2G Robust DNA polymerase, 1.0µL of (5µM) common alpha forward primer and 1.0µL of either and α-1 or α-2 specific reverse primer and 15.2µL of sterile H_2O.

Each of the PCR reactions were then transferred to a PeqStar 96 Universal Gradient thermal cycler (Peqlab, UK) and either the 'ALPHA-1.Kappa.cyc' or 'ALPHA-2.Kappa.cyc' amplification programme was selected. Each thermal cycler programme consisted of a touchdown PCR protocol, detailed below, in which different starting annealing temperatures were used for the alpha-1 and alpha-2 globin genes.

'ALPHA-1.Kappa'	**'ALPHA-2.Kappa'**
Initial Denaturation: 95°C for 3mins	95°C for 3mins
5 Cycles of:	
95°C for 30 sec (**Denature**)	95°C for 30 sec
62°C for 30 sec (**Anneal**)	66°C for 30 sec
72°C for 1 min (**Extension**)	72°C for 1min
5 Cycles of:	
95°C for 30 sec	95°C for 30 sec
60°C for 30 sec	64°C for 30 sec
72°C for 1min	72°C for 1min
25 Cycles of:	
95°C for 30 sec	95°C for 30 sec
58°C for 30 sec	62°C for 1min
72°C for 1min	72°C for 1min
72°C for 1min (**Final Extension**)	72°C for 1 min
8°C Hold Forever	8°C Hold Forever

Agarose gel electrophoresis was then carried (as detailed in Section 3.3b) in order to check that the amplification had been successful.

The PCR products were then purified by *micro*CLEAN, as detailed earlier in Section (5.1e).

5.6 b Internal quality control

Most genomic DNA samples will generate an α-1 PCR product but for those individuals who are homozygous for the alpha 3.7 kb or 4.2 kb deletion will not produce an α-2 PCR product, therefore prior to alpha sequencing, alpha Gap-PCR analysis was carried out so the alpha deletion status of the individual is known.

5.6 c Cycle Sequencing

The alpha 1 and 2 gene specific PCR products (from step 5.5a) were used as individual templates for the Big Dye Terminator cycle sequencing reaction (as discussed in Section 5.1f). Cycle sequencing of each of the *HBA1* and *HBA2* gene-specific PCR products was performed using four primers; N13F and N13R primers (as previously detailed in Table 2.10) and two internal forward and reverse exon primers. (Sequences detailed below in Table 2.14).

These primers provided a good coverage of each of the alpha globin genes as they provided an overlapping sequence read, their respective product sizes are shown below in Figure 2.10:

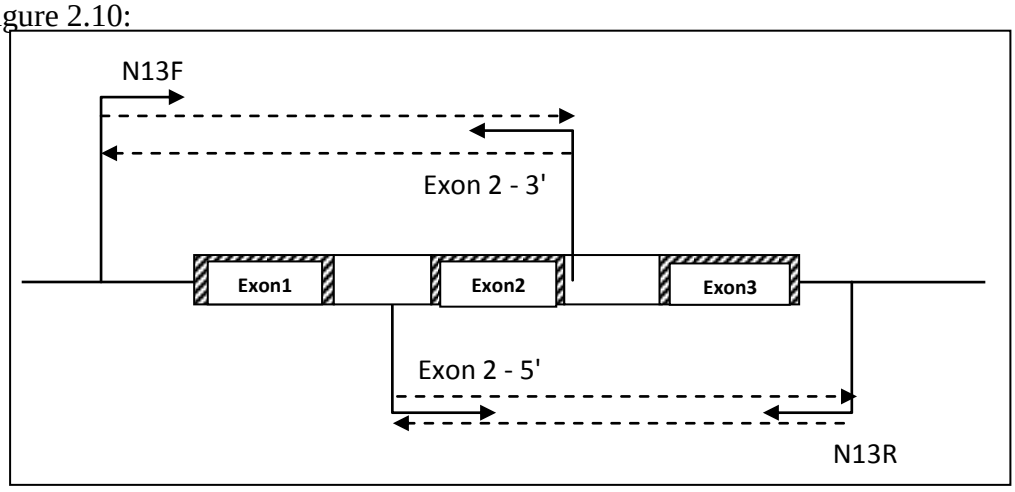


Figure 2.10: Cycle sequencing of *HBA1* and *HBA2*.

Internal exon cycle sequencing primers	Primer sequence 5' to 3'
αEx2-5'F	ACCCCTCACTCTGCTTCTC
αEx2-3'R	CCTCGACCCAGATCGCT

Table 2.14: Details of the internal exon cycle sequencing primers used for cycle sequencing.

See Appendix 5 for *HBA1* and *HBA2* genomic DNA sequence with primer locations

Cycle sequencing reaction set up

The cycle sequencing reactions were carried out directly in a 96-well plate (Thermo Scientific, UK) format. On ice, 2 µL of purified PCR product (template DNA), representing each fragment of the alpha 1 or alpha 2 gene as appropriate, was aliquotted from the individual PCR tubes into each corresponding well of the plate.

For the sequencing of both alpha genes for a single patient, a total of 8 different PCR product and primer combinations are present, but as the master mixes for the α1 and α2 re-sequencing
are the same, only four different sequencing master mixes were required.

Four separate cycle sequencing reaction master mixes were prepared for each of the four different primer (N13F, N13R, αEx2-5' common F, αEx2-3' specific R) cycle sequencing reactions, for the appropriate number of reactions using the 1x reaction mix, to make up a total volume of 13µL, detailed below in Table 2.15:

Reagent	Volume per reaction (µL)
Big Dye Terminator v3.1 (Life Technologies, UK)	1.0
5x Cycle Sequencing Buffer (Life Technologies, UK)	2.5
N13 forward primer (3.3µM)	1.0
N13 reverse primer (3.3µM)	
αEx2-5' common F primer	1.0
αEx2-3' specific R primer	1.0
sH_20	8.5

Table 2.15: Master mix volumes for *HBB* cycle sequencing reaction.

13µL from each of the four master mix reactions were added separately to each appropriate well in the microtitre plate, containing the DNA template. The plate was then sealed with a septum, briefly centrifuged then transferred to a Primus 96 thermal cycler

(MWG Biotech, Germany) and run on the 'Alpha.SEQ.cyc' programme, which consisted of an initial 3-minute denaturation at 96°C, followed by 25 cycles of 96°C denaturation for 15 seconds, 55°C annealing for 10 seconds, and 60°C extension for 2 minutes, then a final hold at 8°C.

The cycle sequencing products were then purified by ethanol precipitation (Section 5.1g) and sequencing analysis was carried out on the ABI Prism 3130*xl* Capillary DNA Sequencer (Life Technologies, UK) (Section 5.2) and then analysed using Mutation Surveyor Software (Sections: 5.3 and 5.4).

See Section 5.8 for details of the *HBA1 and HBA2* reference sequences used and the primer locations.

5.7 Confirmation of results

When a point mutation had been detected in either *HBA1*, *HBA2* or *HBB*, it was confirmed by repeat PCR amplification of the appropriate portion of the gene, using the original patients DNA sample, as described in either Sections 5.1c (*HBB*) or 5.5a (*HBA*), and then sequenced and analysed as detailed in Sections: 5.1d-g, 5.2 to 5.4.

5.8 Online Resources and mutation nomenclature

A list of reported causative alpha and beta thalassaemia mutations are contained within an online repository known as the database of human variants and thalassaemias (http://globin.psu.edu). This locus specific database (LSD) was consulted to determine if a change detected had been previously reported and to assess its pathogenic nature and clinical manifestations. The base and amino acid numbering of each mutation found was determined by referring to this haemoglobinopathy database and published *HBA1*, *HBA2* and *HBB* sequences. Each mutation described at the amino acid level was numbered as +1 from the initiator methionine, as per the LSD. Mutations were also described at the cDNA level according to guidance issued by the Human Genome Variation Society (http://www.hgvs.org/). Note that the LSD does not use HGVS numbering at the protein level. Any unclassified changes that had not been previously reported were assessed by unclassified variant *in silico* analysis tools including Almaut (Interactive Biosoftware, France).

Reference sequences for *HBB, HBG2, HBA1 and HBA2* were obtained from Ensemble database (http://www.ensembl.org/index.html). Their respective GenBank reference

sequences are: *HBB*: NM_000518.4, *HBG2:* NM_000184.2, *HBA1*: NM_000558.3 and *HBA2*: NM_000517.4.

6.0 Multiplex Ligation Probe Amplification (MLPA)

Multiplex Ligation Probe Amplification (MLPA) (MRC-Holland, the Netherlands) is a sensitive laboratory method for the relative quantification of copy number or dosage of many nucleic acid sequences in a single reaction containing, requiring only 200ng of genomic DNA.

It is useful for the detection of rare or previously uncharacterised deletions or duplications affecting the alpha or beta globin gene clusters and unlike other diagnostic assays such as sequencing, it is capable of detecting copy number changes of complete exons.

In this project MLPA was used for the identification of unknown deletions when a causative mutation could not be identified using the assays described above.

MLPA is based on the hybridisation or binding of one MLPA probe (consisting of two oligonucleotides directly flanking the DNA target sequence). The two parts of each probe hybridize to adjacent target sequences and are ligated by a thermostable ligase. All of the ligated probes are subsequently amplified in a single PCR reaction, using only one primer pair. The resulting amplification products of each probe have a unique length of between 130-480 nucleotides which are then separated and analysed by capillary electrophoresis. The relative amounts of probe amplification products, as compared to a control DNA sample, reflect the relative copy number of target sequences. In the presence of a heterozygous gene deletion, fewer MLPA probes will hybridise, resulting in a characteristic reduction in probe height by up to 50%.

Two separate SALSA MLPA kits were used in this project; P140 (version B2) for analysis of the alpha globin gene cluster and P102 (version B1) for analysis of the beta globin gene cluster. Each kit comprises the same reagents needed for ligation and PCR, but each with a unique probemix. The P140-B2 *HBA* SALSA MLPA kit contains 38 individual probes, including one probe that is specific for the detection of the Constant Spring point mutation, with amplification products of between 130 and 409 nucleotides and is designed to detect copy number changes of 24 different sequences within the alpha globin gene

cluster. The P102-B1 *HBB* SALSA MLPA kit contains 43 probes, including one probe that is specific for the sickle-cell point mutation, with amplification products of between 130 and 481 nucleotides and is designed to detect deletions/duplications within the beta globin gene cluster.

In addition, both of these kits contain nine control fragments generating amplification products smaller than 120 nucleotides, including: four DNA Quantity fragments (Q-fragments) at 64, 70, 76 and 82 nucleotides, which are only visible when the DNA sample quality is less than 100ng, three DNA denaturation control fragments (D-fragments) at 88, 92 and 96 nucleotides, which when found at a low signal indicate incomplete denaturation, one X-fragment at 100 nucleotides, specific for the X chromosome and one Y-fragment at 105 nucleotides, specific for the Y chromosome.

Figures 2.11 below show a schematic representation of the location of the probes in the P140-B2 *HBA* and P102-B1 *HBB* MLPA kits, where the numbers above the arrows represent the amplification size of the respective probe in nucleotides:

Figure 2.11a:

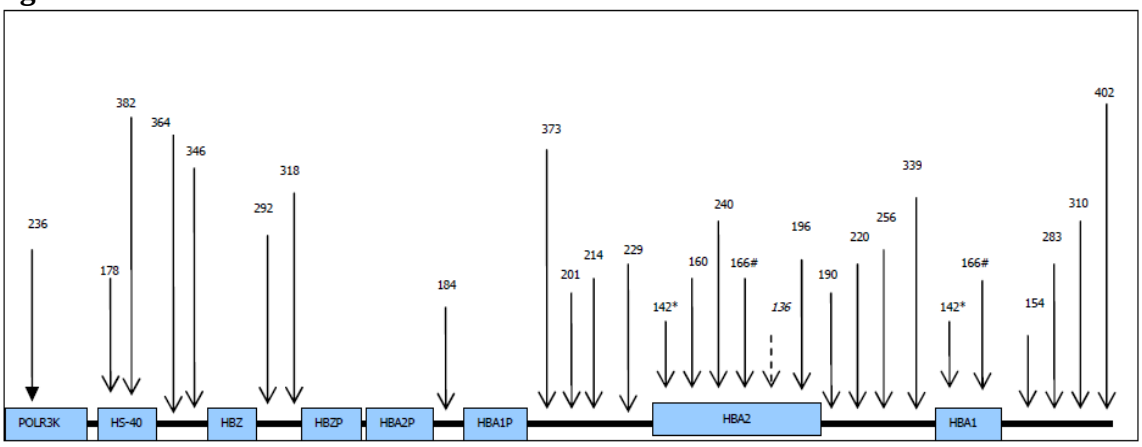


Figure 2.11b:

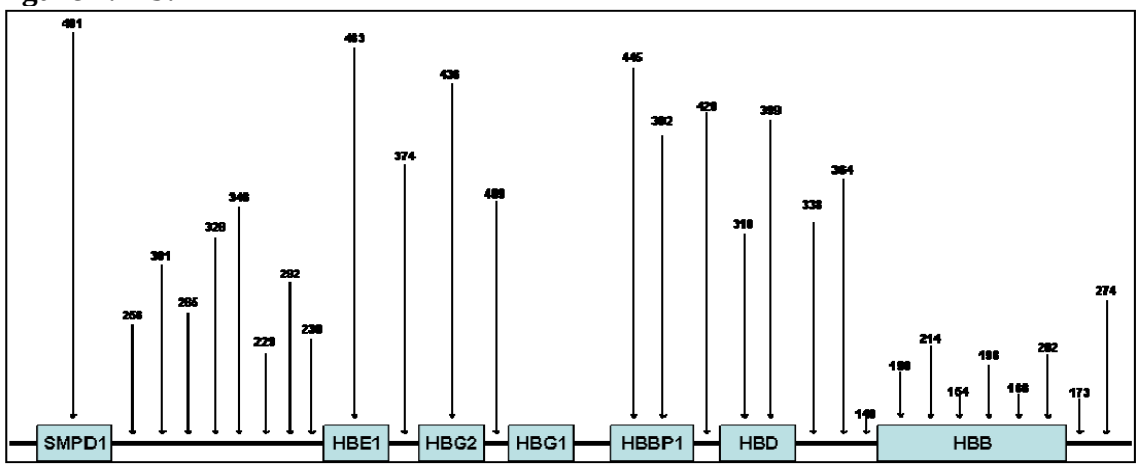


Figure 2.11: Schematic representation of the location of the probes in the (a)P140-B2 *HBA* MLPA kit (Description version 22; 24/12/2011) and the (b) P102-B1 *HBB* MLPA kit (Description version 14; 24/08/2011) (Obtained from MRC-Holland: www.mlpa.com).

See Appendix 6 and 7 for details of SALSA MLPA P140-B2 *HBA* and P102-B1 *HBB* probemix, with their corresponding chromosomal positions.

Figures 2.12 below shows the actual capillary electrophoresis pattern of the amplified probe profile across the alpha and beta globin gene clusters:

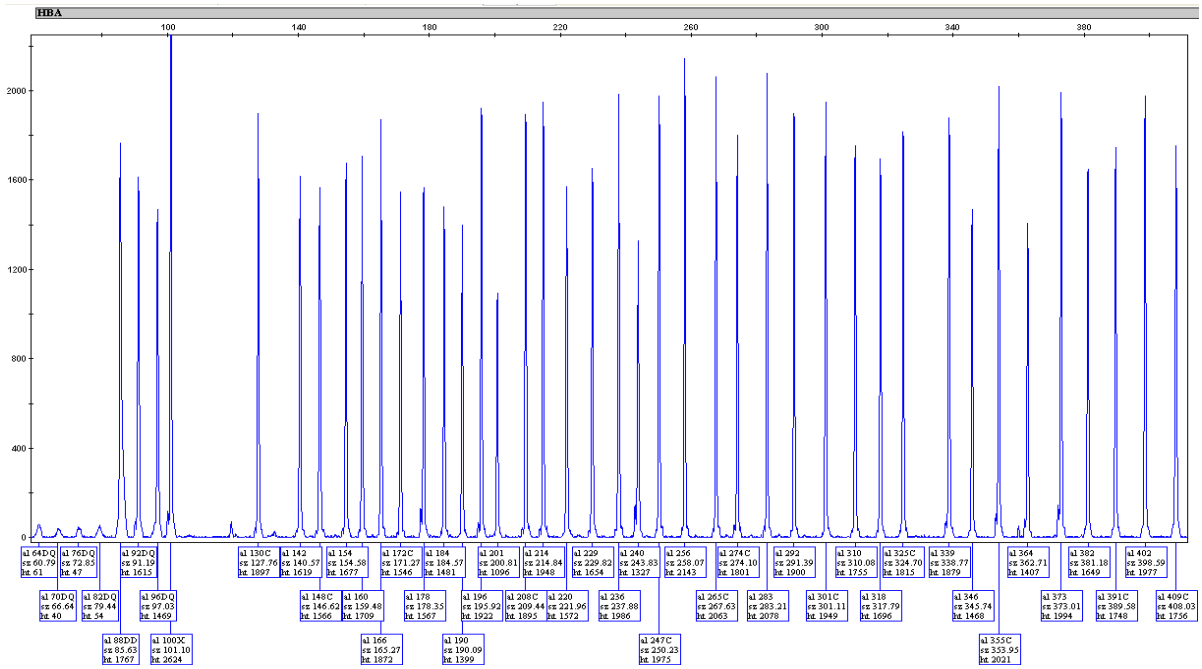


Figure 2.12a: Capillary electrophoresis pattern from a female control DNA sample of analysed with SALSA MLPA kit P140-B2 *HBA* (lot 0508).

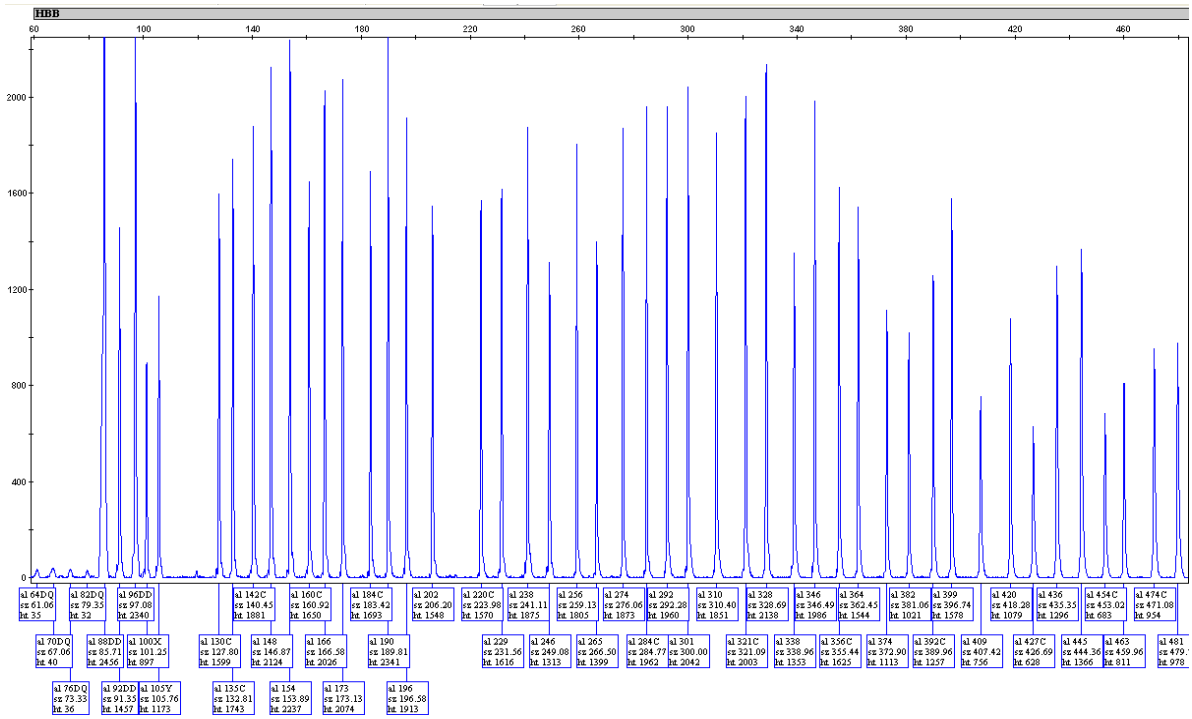


Figure 2.12b: Capillary electrophoresis pattern from a male control DNA sample of analysed with SALSA MLPA kit P102-B1 *HBB* (lot 0508).

6.1 Method

MLPA was performed in a single reaction and can be divided into three main steps:

6.1a DNA Denaturation and hybridisation of the SALSA MLPA probes

1. The DNA was diluted with TE buffer, at least 24 hours beforehand, to a concentration of 40ng/µL.

2. 5µL of diluted DNA was added to appropriately labelled 0.2µL dome cap, thin-walled strip tubes (Alpha laboratories, UK). At least seven 'normal' controls and one blank was included on each run.

3. The sample strips were then loaded onto an Applied Biosystems 2720 thermal cycler (Life Technologies, UK) and denatured at 98°C for 5 minutes then left to cool to 25°C.

4. Once the samples had cooled to 25°C, using a multi-channel pipette, 1.5µL of MLPA buffer and 1.5µL of Probemix (MRC Holland) was added to each of the samples whilst sat on the thermal cycler.

5. The samples were left to hybridise overnight by selecting the 'Hybrid+Ligate' programme on the Applied Biosystems 2720 thermal cycler which consisted of: 1 minute initial incubation at 95°C followed by 18 hours at 60°C.

6.1b Ligation reaction

It was crucial at this stage the samples are not removed from the thermal cycler block

1. A ligase buffer mastermix master of the following was prepared:

Reaction Component	Volume/per reaction (µL)
Ligase Buffer A	3.0
Ligase Buffer B	3.0
Sterile H_2O	25
Ligase-65	1.0

2. 32µL of the above mastermix was transferred to each of the samples on the thermal cycler held at 54°C.

3. The 'Hybrid+Ligate' program was then resumed to the next step, which consisted of: 15 minute incubation at 54°C, followed by 5 minutes at 98°C then final hold at 4°C.

6.1c MLPA PCR reaction

1. A master mix consisting of: 4µL of SALSA PCR buffer and 26µL sterile H_2O was prepared.

2. 30µL of this master mix was aliquotted into a set of new 0.2µL thin-walled, flat-cap strip tubes (Starlab, UK), that were labelled with the same corresponding sample numbers at is in step 2 (Section 6.1a).

3. 10µL of the ligation product (from Section 6.1b) was transferred into the tubes containing the SALSA PCR master mix.

4. The tubes were then loaded onto the thermal cycler and the 'MLPA-PCR' programme, the samples were initially held the 60°C hold step.

5. Meanwhile, a PCR master mix of the following reagents was prepared:

Reaction Component	Volume/per reaction (µL)
SALSA primers	2.0
SALSA enzyme dilution buffer	2.0
Sterile H_2O	5.5
SALSA polymerase	0.5

6. Using a multi channel pipette, 10µL of the above PCR polymerase master mix was added to each tube sitting on the thermal cycler.

7. The 'MLPA PCR' program was resumed, and the thermal cycle conditions consisted of 33 cycles of: 95°C denaturation for 30 seconds, 60°C annealing for 30 seconds, and 72°C

extension for 1 minute, followed by a final 20 minute extension at 72°C, then a hold step at 4°C.

6.2 Separation of amplification products by capillary electrophoresis

1. This step is carried out directly into a 96-well plate format (Thermo Scientific, UK).

2. A master mix of 9µL of Formamide and 0.3µL of ROX-500 size standard was prepared for the appropriate number of samples.

3. 9.3µL of the above master mix was aliquotted into each sample well of the plate.

4. 1µL of the MLPA PCR product, (from Section 6.1c) was added to each corresponding well of the plate, containing the master mix.

5. The plate was sealed with a septum and briefly centrifuged.

6. The plate was then placed on the thermal cycler and denatured for 2 minutes at 94°C.

7. The plate was then placed on ice to 'snap-chill' for two minutes, to ensure the DNA products remain single stranded.

8. The plate was then loaded onto the ABI Prism 3130*xl* Capillary DNA Sequencer (Life Technologies, UK) in order to carry out the fragment analysis, entering the following specific parameters in Life Technologies Genemapper plate editor:

Sample Type: Sample
Size Standard: GS500(-35)
Panel: HBB or HBA panel, as appropriate
Instrument protocol: Fragment Analysis _POP7
Analysis method: HBB or HBA_ANALYSIS

6.3 MLPA Data analysis using Genemapper Software (Life Technologies, UK)

Once the fragment analysis had been completed, the relevant sample files (.fsa) were imported and analysed in Genemapper software (version 4). Provisional analysis of the morphology and designations of the size standard was done, and the MLPA probe data was inspected, ensuring all the appropriate control probes were present and that there was no unexplainable loss in the signal of the probes. The raw result table was then exported in a .txt which was then imported to an Excel spreadsheet, below, for further analysis.

6.4 MLPA Data analysis using manual Excel spreadsheet

Further fragment analysis and visualisation of the probes was performed using the appropriate Excel based interpretation macro (http://www.ngrl.org.uk/Manchester/ publications/ MLPA) for either *HBA* (P140-B2*HBA*-v18) or *HBB* (P102-B1*HBA*-v12). This macro normalised the electrophoresis results by comparing the relative probe signals from all the samples to each other and to the control probes, which was essential for obtaining accurate MLPA data (www.mlpa.com). From this intra assay analysis a 'Sample quality's score was produced, and each sample must pass this score i.e the standard deviation must be less than 0.1. Samples with a quality value exceeding 0.1 were rejected.

Once the reliability of the data has been established a deletion in one copy of a probe target sequence (a deletion in either the alpha or beta globin gene clusters) was apparent by a reduction in relative probe peak height for that probe amplification product of 35-55%.

Figures 2.13 below show the diagrammatic representation of the MLPA probe data as a histogram, produced from the Excel spreadsheet

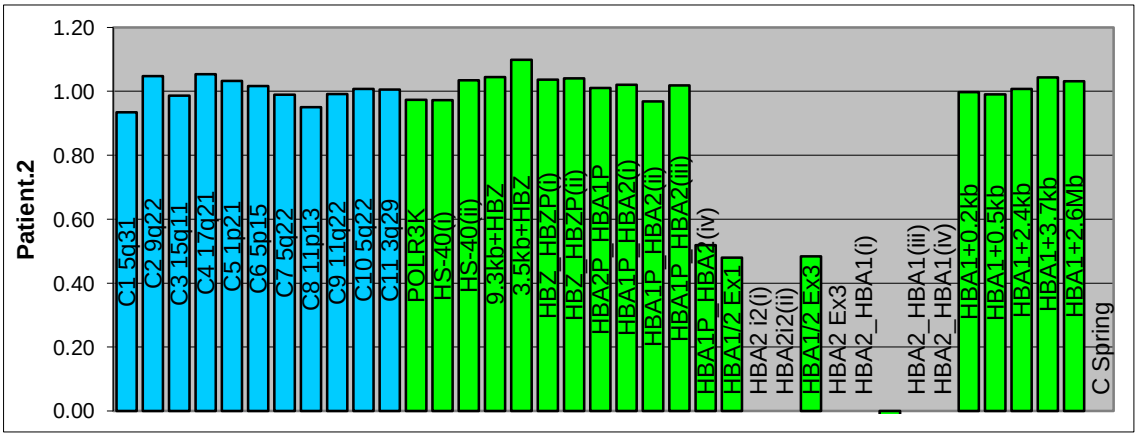


Figure 2.13a: Histogram showing *HBA* MLPA probe data from a patient DNA sample with a homozygous alpha plus deletion, showing probes missing.

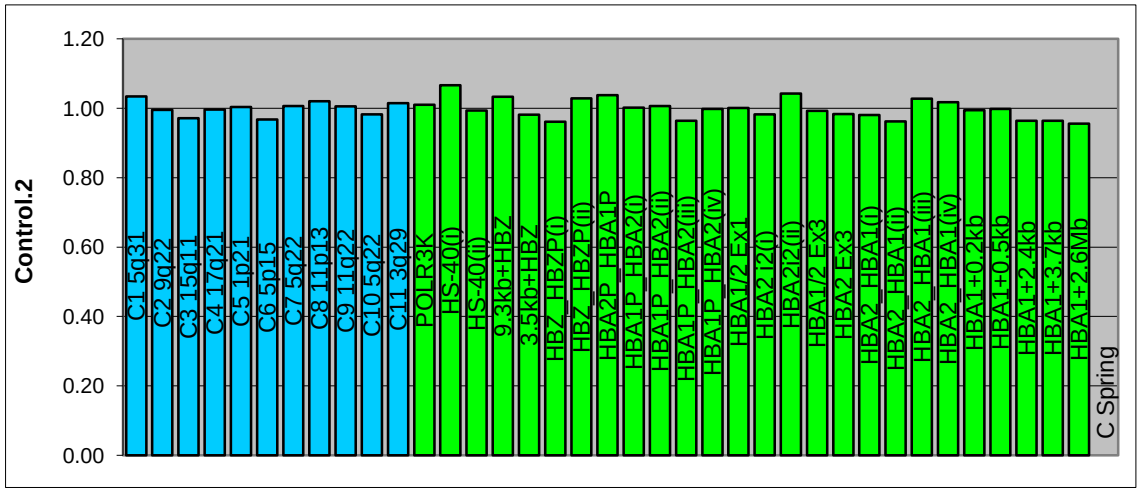


Figure 2.13b: Histogram showing *HBA* MLPA probe data from a control DNA sample.

6.5 Quality control

Mutations or polymorphisms within the sequence detected by a probe may also result in a reduced relative peak area due to failure to bind a probe. Therefore apparent deletions detected by a single probe must always be confirmed by an alternative method. Furthermore, due to the number of alpha globin genes the relevant probes will only show a relative decrease in signal of between 20-25%.

Chapter 3: Results

All patient data discussed in this chapter is derived from the raw information contained within Appendix 8.

3.0 Sample referral

Between July 2008 and December 2011, 328 referrals from the NW were genotyped at Manchester Royal Infirmary as part of the national Antenatal and Newborn Screening Programme. A range of genetic analysis was carried out on each sample, including Gap-PCR, full gene sequencing and MLPA, as described in chapter 2.

3.1 Proportion of all haemoglobinopathy diagnoses

The total number of haemoglobinopathy diagnoses (Table 3.1) comprised of: 34% (*n=113*) alpha thalassaemia, 33% (*n=109*) beta thalassaemia, 17% (*n=55*) were compound heterozygous for an alpha and beta thalassaemia mutation and/or variants and 3% (*n=10*) of patients had a mutation resulting in HPFH or δβ thalassaemia. In the remaining 13% (*n=41*) of referrals either no causative mutation was detected or the genotype detected was not enough on its own to account for the patients' phenotype.

Globin chain/disorder	Number of patients	Percentage of total referrals (%)
Alpha thalassaemia	113	34
Beta thalassaemia	109	33
Compound heterozygous (α and β-chain mutations)	55	17
No causative mutation detected	41	13
HPFH/δβ disorder	10	3

Table 3.1: Breakdown of the total number of haemoglobinopathy diagnoses

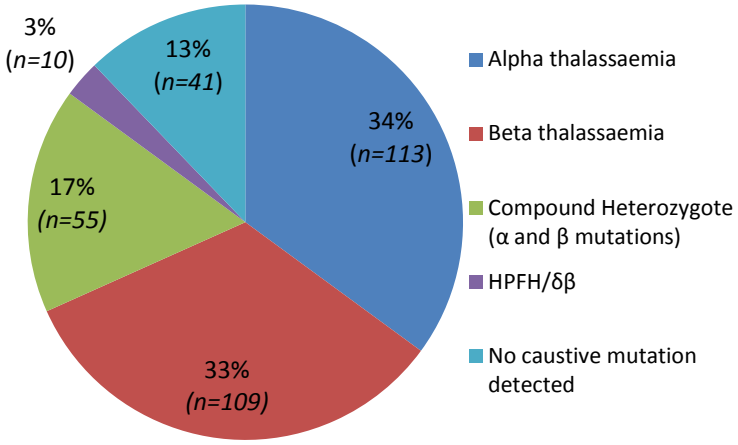


Figure 3.1: Total haemoglobinopathy diagnoses

Figure 3.1 shows that the largest proportion of patient diagnoses was either alpha or beta thalassaemia which comprised 34% (*n=113*) and 33% (*n=109*) of all patient referrals respectively. A total of 17% (*n=55*) of patients had a compound heterozygous genotype i.e. a mutation within both of the beta and alpha globin genes, these cases are discussed further in Section 7.0. A further 3% (*n=10*) of patients had a HPFH or δβ disorder, and in the remaining 13% (*n=41*) of patient cases genetic testing within the Molecular Diagnostics Centre did not detect a causative mutational mechanism, or the genotype detected was not enough on its own to account for the patients phenotype (these are discussed further in Section 11.0).

3.2 Stated ethnicity of all patients referrals

The ethnicity of each patient was recorded either on the Family Origin Questionnaire (FOQ), completed by a healthcare professional or on the appropriate paperwork as part of the routine antenatal genetic referral process.

The stated ethnic group of each patient was recorded according to the categories outlined on the FOQ: 'Mixed', 'White', 'Mediterranean' (MED), 'Asian', 'South East Asian'(SEA), 'Black', 'Arabic', 'Any other' or when this information was not available, 'Not stated'. This information was then collated in order to asses the incidence of haemoglobinopathy disorders within each ethnic group referred from the North West

population, as shown below in Figure 3.2:

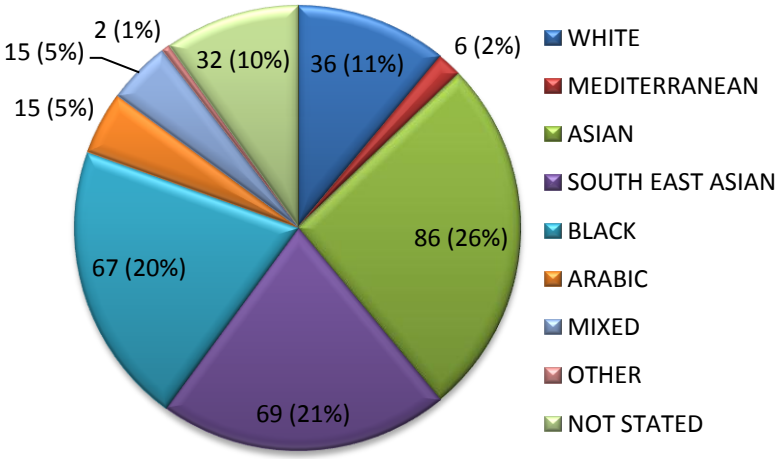


Figure 3.2: Proportion of patient ethnic groups referred

Figure 3.2 demonstrates ethnic diversity of the North West referral population. The incidence of haemoglobinopathies was highest in the Asian (26%) ethnic group, followed by South East Asian (21%) and Black (20%) ethnic groups, with the other ethnic groups collectively comprising 33% of all referrals. The three ethnic groups that formed the largest proportion of patient referrals were further sub-divided into more specific ethnicities so that further analysis could be performed. This is represented graphically below in figures 3.3a to 3.3c.

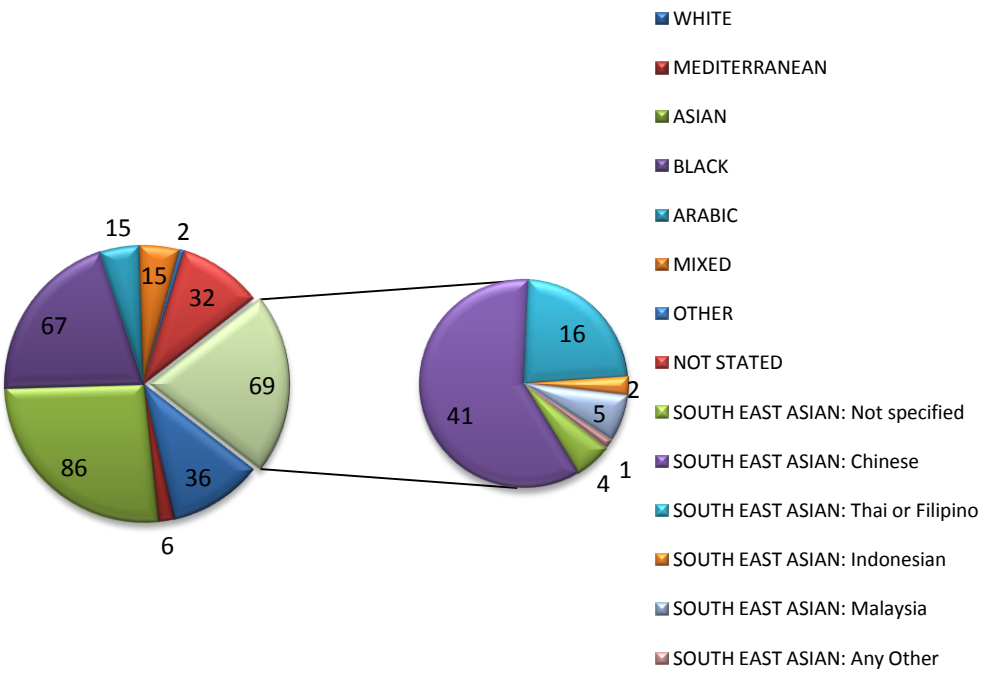


Figure 3.3a: Proportion of South East Asian ethnic subgroups

Figure 3.3a shows that from all the diagnoses made in those patients from the South East Asia (21%, *n=69*) ethnic group, a large proportion comprised those from Chinese (59%, *n=41*) and Thailand or Filipino (23%, *n=16*) ethnic sub-groups.

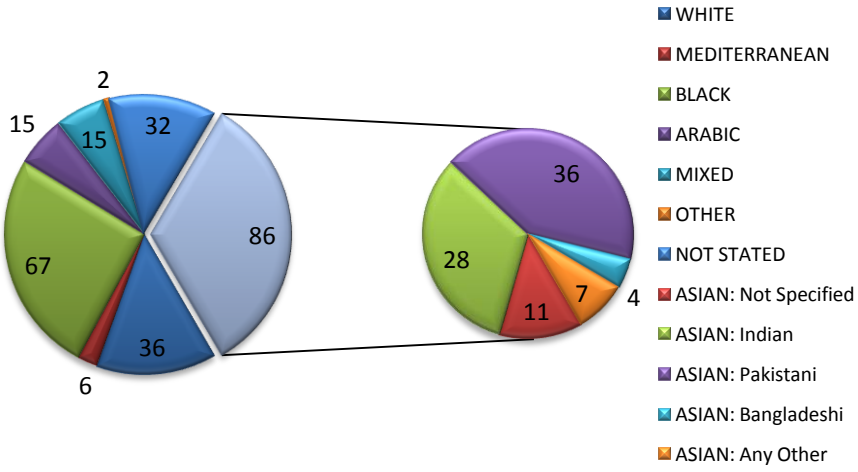


Figure 3.3b: Proportion of Asian ethnic subgroups

Figure 3.3b shows that of the diagnoses made in those patients from the Asian ethnic group (26%, *n=86*), a large proportion comprised those from Pakistani (42%, *n=36*) and Indian (33%, *n=28*) ethnic groups.

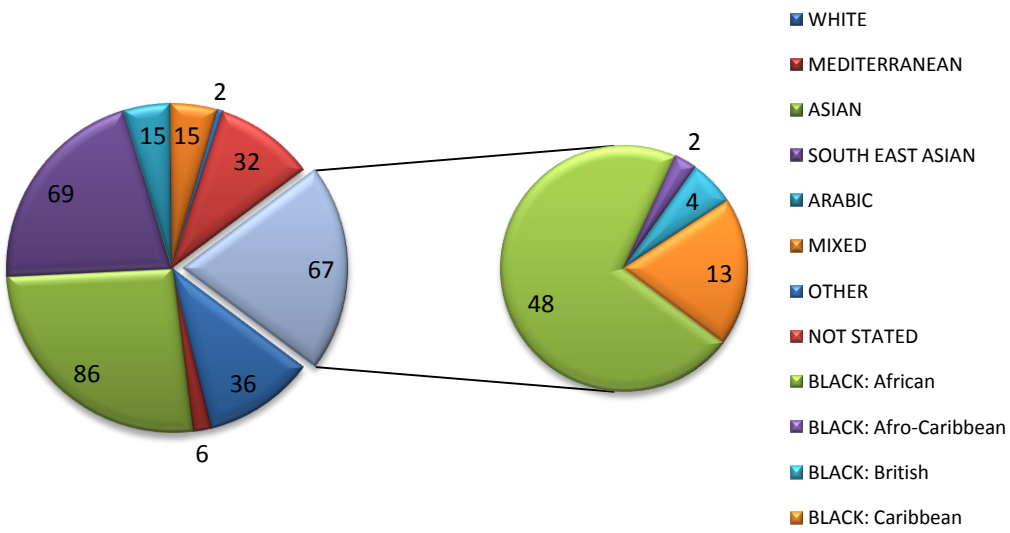


Figure 3.3c: Proportion of Black ethnic subgroups

Figure 3.3c shows that from the diagnoses made in those patients from the Black ethnic group (20%, *n=67*), the largest proportion of these patients were those originating from the African ethnic group (72%, *n=48*).

3.3 Mutation data analysis

In the following sections analysis of all the genotypes encountered with the patient referrals was performed in order report the spectrum and frequency of haemoglobinopathy mutations found within the MDC. The distribution and frequency of mutations detected in each individual was then related to their ethnic group, where known, and assessed in terms of whether or not that mutation would be expected in their ethnic group from previous population studies. This enabled us to assess the effectiveness of our broader genetic screening approach versus ethnic led, targeted mutation screening.

This data is reported below according to the nature of the mutations and the laboratory methods used.

4.0 Gap-PCR analysis for the detection of deletional forms of alpha thalassaemia

A total of 131 patients (40% of the total haemoglobinopathy referrals) had a form of deletional alpha thalassaemia (4 patients were not tested for deletional alpha thalassaemia due to the limited quantity of DNA available).

All patient referrals were then categorised based on the clinical severity of their genotype and then a comparison made to their stated ethnicity. A total of 42% (*n=55*) of all patients identified with a deletional form of alpha thalassaemia had a mild alpha plus genotype (-$\alpha^{3.7}$ or $-\alpha^{4.2}$), whereas 18% (*n=24*) of patients had homozygous alpha plus thalassaemia trait. A total of 32% (*n=42*) of patients had a clinically severe genotype of alpha zero thalassaemia trait, and 4% of patients (*n=5*) had deletional Hb H disease. Furthermore 4% of patients (*n=5*) had the triplicated alpha genotype.

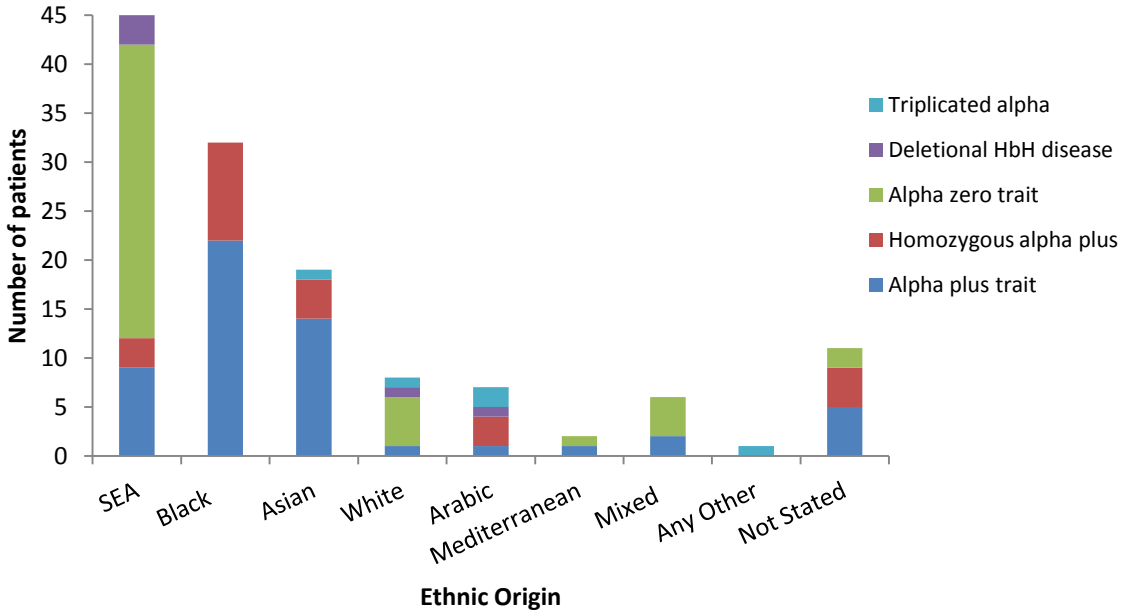


Figure 3.4: Distribution of deletional alpha thalassaemia genotypes between ethnic groups

Figure 3.4 above shows that the majority of severe deletional alpha genotypes were reported mainly in patients from the South East Asian group, where alpha zero thalassaemia and deletional Hb H disease was most prevalent. The mixed ethnic group, although small in number overall (*n=6*) had a high proportion (*n=4*) of alpha zero thalassaemia. There was also a large proportion (*n=5*) of alpha zero thalassaemia reported in the White ethnic group, mainly accounted for by the north west of England originating --BRIT deletion. Furthermore, mild and clinically insignificant genotypes such as alpha plus trait and homozygous alpha plus were prevalent in Black, Asian and Arabic ethnic groups.

The distribution and frequency of all deletional forms of alpha thalassaemia detected were then analysed and their frequency determined within each ethnic group, as shown in Figures 3.4a and 3.4b.

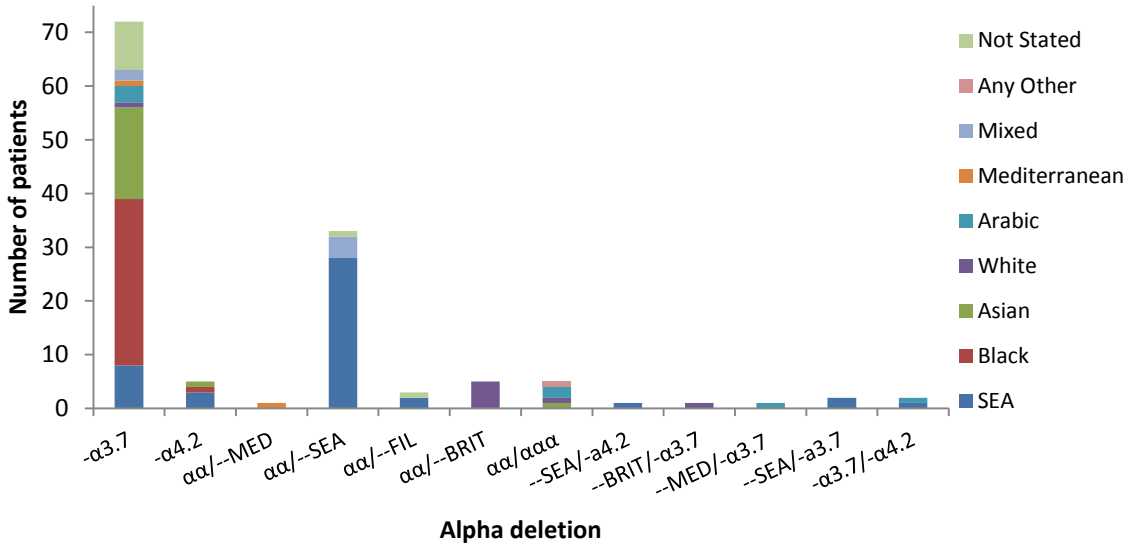


Figure 3.4a: Distribution of deletional alpha thalassaemia genotypes between ethnic groups.

Figure 3.4a shows that heterozygosity for the alpha plus -$\alpha^{3.7}$ deletion and alpha zero --^{SEA} deletion were the most common deletional genotypes within the referral population, comprising 55% (*n=72*) and 25% (*n=33*), respectively of all deletional alpha thalassaemia genotypes.

This data has been further modified in Figure 3.4b to exclude the large number of patients with –$\alpha^{3.7}$ and --SEA deletions in order to show the distribution of other deletional alpha thalassaemia genotypes more clearly:

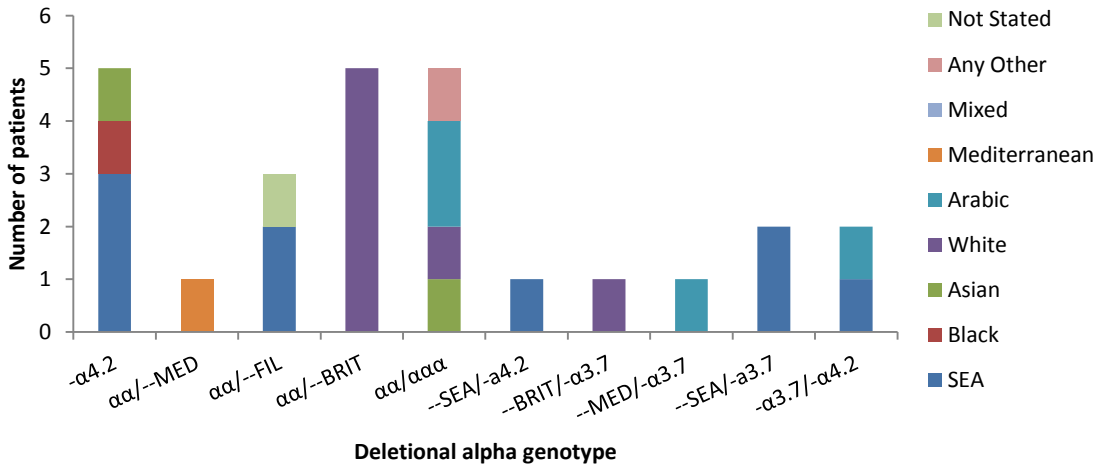


Figure 3.4b: Distribution of deletional alpha thalassaemia mutations between ethnic groups (excluding heterozygous –$\alpha^{3.7}$ and --SEA genotypes).

In proportion to the total number of patients referred within each ethnic group, the largest incidence of deletional alpha thalassaemia is found in those patients from South East Asian, 65% (*n=45*) and Black, 48% (*n=32*) ethnic groups.

The proportion of patients with a deletional alpha thalassaemia genotype was then correlated according to whether their genotype would be expected within their stated ethnicity, shown in Figure 3.5:

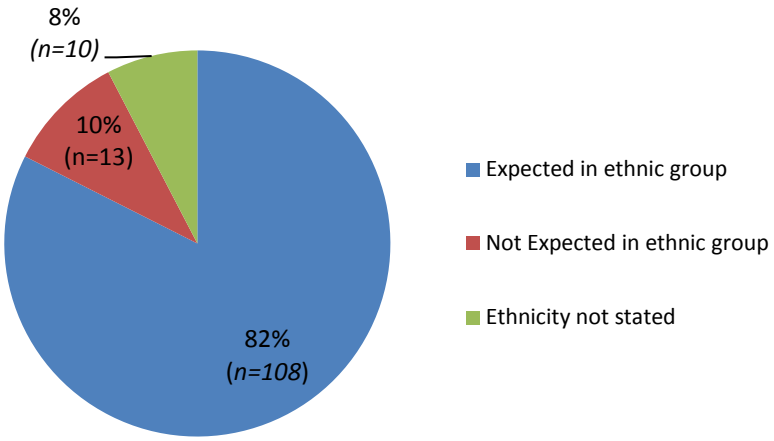


Figure 3.5: Proportion of all deletional alpha thalassaemia patient genotypes according to expected ethnicity.

Figure 3.5 shows that a large proportion of patients with a deletional alpha thalassaemia genotype had a deletion(s) which would be expected in their ethnic group, this shows that there was a strong correlation between the patients' stated ethnicity and their genotype, for example the five reported cases of the alpha zero --BRIT deletion was exclusively found in patients from the 'White' ethnic group, and of the majority (62%, *n=5*) of patients with the alpha plus $-\alpha^{4.2}$ deletion were, as expected were from the SEA ethnic group. Conversely the triplicated alpha genotype was found in a variety of ethnic groups.

5.0 DNA sequencing of *HBA1* and *HBA2* for the detection of point mutations and small deletions associated alpha thalassaemia.

Of the total number of genetic referrals received (*n=328*), 265 patients that had their *HBA1* and *HBA2* genes sequenced, the remaining 63 patients did not have their alpha globin genes sequenced as it was either deemed unnecessary or the causative genotype had already been detected. A causative non-deletional mutation i.e. a point mutation or a small deletion/insertion was detected within either of the alpha globin genes by DNA

sequencing in 18% (*n=59*) of all the patient referrals, including both non-deletional alpha thalassaemia and alpha chain variants. No causative alpha mutation was detected in 58% (*n=190*) of patients, in the remaining 16 patients a candidate change was detected, but the change was not previously unreported on the Hb variant database. Subsequent *in silico* analysis by Alamut software was used to asses the pathogeneicity of these changes, and if it the mutation was subsequently deemed to be of significance i.e. a missense or splice site mutation it was reported alongside the patients' phenotypic data, although in the absence of protein expression studies clinical interpretation could not be offered.

Of the 59 causative mutations detected, 30 were found within the *HBA1* gene and 29 within the *HBA2* gene. The breakdown of the class of mutations detected as shown in Figure 3.6.

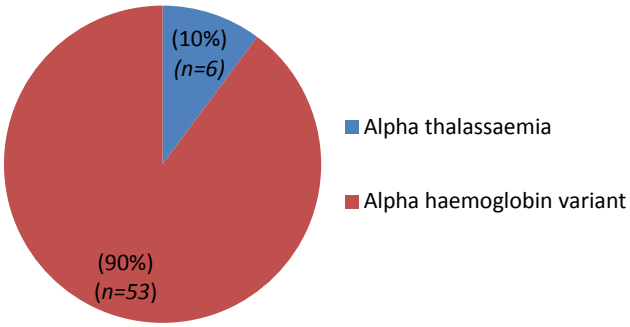


Figure 3.6: Proportion of all non-deletional alpha thalassaemia mutations detected by DNA sequencing of *HBA1* and *HBA2*.

Figure 3.6 shows that out of all of the patient cases in which a causative mutational change was detected within the alpha globin genes, 10% (*n=6*) of patient cases were classified as non-deletional alpha thalassaemia and 90% (*n=53*) were classified as an alpha chain haemoglobin variant. The clinical significance, or otherwise, of these variants is discussed in section 5.2.

5.1 Non-deletional form of alpha thalassaemia

The distribution of the mutations detected that are associated with non-deletional alpha thalassaemia were then correlated according to the patients' stated ethnicity, as shown in Figure 3.7 below:

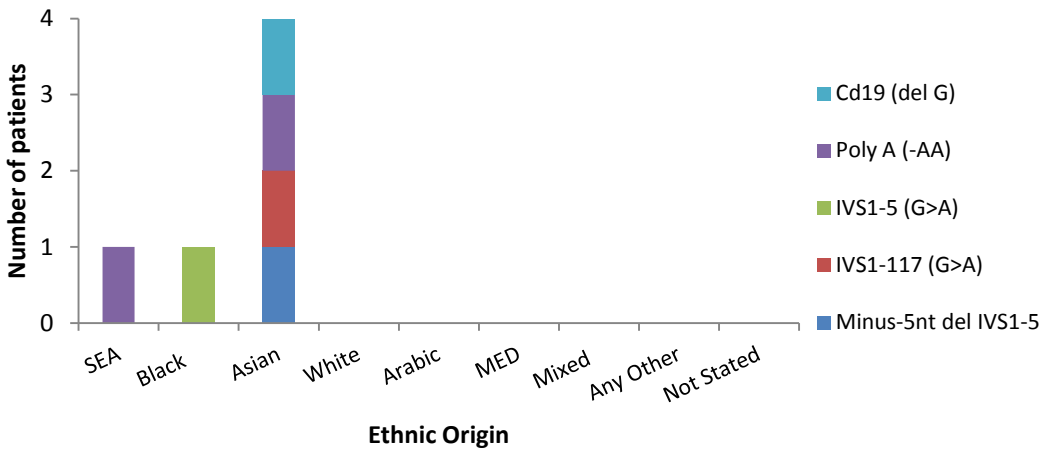


Figure 3.7 Proportion of non-deletional alpha thalassaemia within each ethnic group

Figure 3.7 shows that a total range of 5 different non-deletional mutations causing alpha thalassaemia were found within all patient referrals. The highest incidence of non-deletional alpha thalassaemia was found in those patients from Asian ethnic groups, comprising 67% (*n=4*) of all non-deletional alpha thalassaemia cases, with the remaining two cases reported in one individual from SEA and one from Black ethnicity.

5.2 Alpha chain haemoglobin variants

Thirty three different alpha chain haemoglobin variants were detected within the referral population. Figure 3.8 below shows the total number of individuals within each ethnic group with an alpha chain haemoglobin variant:

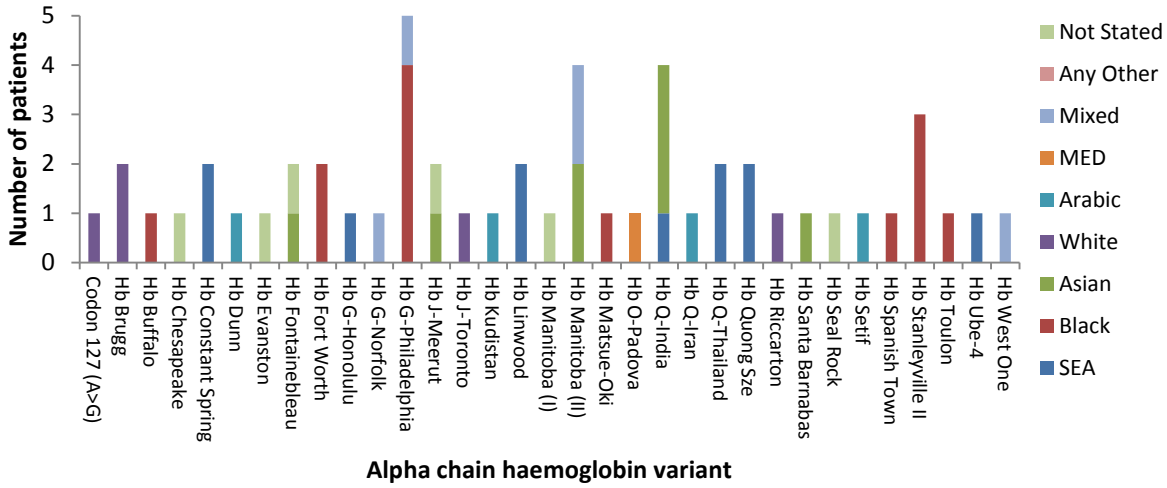


Figure 3.8: Proportion of alpha chain haemoglobin variants detected within each ethnic group

Figure 3.8 shows that 33 different alpha chain haemoglobin variants were detected by DNA sequencing, the most common variant being G-Philadelphia followed by Hb Q-India and Hb Manitoba (II). The largest proportion of variants was found in those patients from Black 25% (*n=13*) and SEA 21% (*n=11*) ethnic groups, followed by 15% (*n=8*) within Asian ethnic groups. A further (10%, *n=5*) cases were reported in White, Mixed and Not Stated ethnic groups. Of this broad spectrum of haemoglobin variants identified, four of these are classified as clinically significant (according to the Hb variant database) because they are associated with a thalassaemic phenotype. These included: Hb Constant Spring, Hb Seal Rock, Hb Quong Sze and Hb Evanston, which were detected in 6 patients, in which 67% (*n=4*) of these cases were found in patients from the SEA ethnic group.

In addition to these 'thalassaemic' variants many other alpha chain variants detected within our referral population could also be considered to be of clinical significance due to the effect of the causative mutation on the quaternary structure of the protein products. These include the following 5 variants that were detected in 6 cases within the referral population: Hb Chesapeake, Hb Dunn, Hb Linwood, Hb Toulon and Hb West-One, all of which are associated with increased oxygen affinity and may cause erthrocytosis. A further 5 haemoglobin variants that were detected within 8 cases (Hb Manitoba (I and II), Hb Riccarton, Hb Setif and Hb Spanish Town) are associated with the formation of unstable haemoglobin tetramers. Furthermore, two haemoglobin variants; Hb G-Philadelphia and Hb Q-Thailand which are classified by the Hb variant database as having a 'normal' clinical presentation in the heterozygous state but are most commonly reported in the compound heterozygous state *in trans* with deletional alpha thalassaemia ($-\alpha^{3.7}$ and $-\alpha^{4.2}$ respectively) resulting in a thalassaemic like phenotype. This clinical presentation was found in all of the seven patient cases with these two haemoglobin variants.

5.3 Population genetics of all causative mutations detected within *HBA1* and *HBA2*.

Of all the patient cases in which a causative point mutation (i.e thalassaemia or haemoglobin variant) was identified, 53% (*n=31*) had an alpha genotype that would be expected in their stated ethnic group, 37% (*n=22*) had an alpha genotype which not be expected within their ethnic group and in 10% (*n=6*) of cases the patient ethnic group was not stated, as shown in figure 3.9 below:

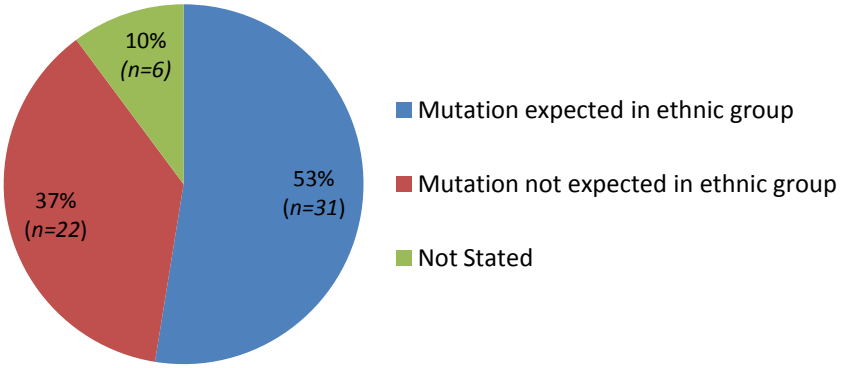


Figure 3.9: Proportion of all alpha genotypes according to expected ethnicity.

Therefore DNA sequencing of the essential regions of the alpha-globin genes indicates that a "first pass" ethnic led mutation screening strategy may miss up to 47% of alpha genotype cases analysed.

6.0 DNA sequencing of *HBB* for the detection of point mutations and small deletions associated beta thalassaemia.

DNA sequencing of the beta-globin gene identified 44 different types of non-deletional mutations i.e. a point mutation or a small deletion/insertion within 168 patients. All together there were a total of 180 beta mutations reported, including heterozygous and compound heterozygous genotypes reported within this group.

Of the total number of genetic referrals received (*n=328*), 308 patients that had their beta globin genes sequenced, the remaining 20 patients did not have their beta globin genes sequenced as it was not necessary. A causative mutation was detected within the beta globin gene by DNA sequencing in 51% (*n=168*) of all patient referrals, and no beta mutation was detected in the remaining 49% (*n=160*) of patients.

Of the 160 cases in which a causative beta mutation was not was detected, a candidate change was detected within 8 patients, but these changes were either a polymorphism of no clinical significance or an unclassified variant in which subsequent *in silico* analysis using Alamut software deemed the change not to be pathogenic in nature.

Out of the 168 patients in whom a causative mutation was detected by DNA sequencing of the beta globin gene, the largest proportion of cases 58% (*n=96*) had a beta chain

haemoglobin variant mutation, 37% (*n=63*) had a beta thalassaemia mutation and remaining 5% (*n=9*) of patients were compound heterozygous for a thalassaemia and a haemoglobin variant mutation, as represented in Figure 3.10.

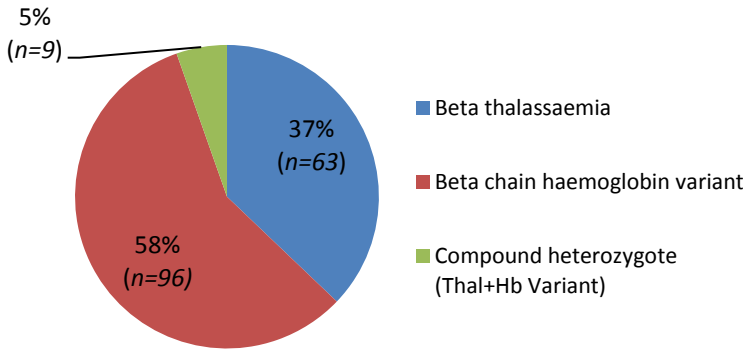


Figure 3.10: Proportion of all causative beta thalassaemia mutations detected by DNA sequencing of *HBB*.

The proportion of beta genotypes (represented above) was then divided according to the patients' ethnicity:

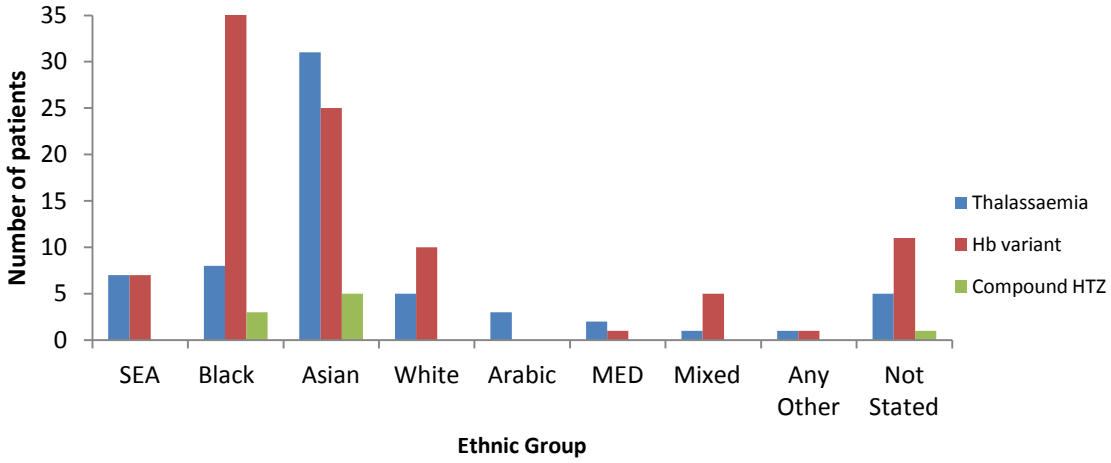


Figure 3.11: Proportion of beta genotypes within each ethnic group.

Figure 3.11 shows that the Asian ethnic group represented the largest proportion of patients with beta thalassaemia mutations, comprising 49% (*n=31*) of all beta thalassaemia genotypes diagnosed. Black ethnic groups had the largest number of haemoglobin variant genotypes comprising 38% (*n=36*) of all cases with this genotype. The SEA ethnic group had equal proportions of thalassaemia and haemoglobin variant mutations, whereas the White ethnic group had a much larger proportion of haemoglobin variant mutations (*n=10*) than thalassaemia mutations (*n=5*), the majority of which (*n=8*) were associated with a change in the protein behaviour, i.e. unstable or associated with an increased

oxygen affinity. The small number of compound heterozygote genotypes was confined to the Asian and Black ethnic groups.

6.1 Mutations detected within *HBB* associated with beta thalassaemia

Further analysis of the mutations associated with beta thalassaemia is represented in Figure 3.12.

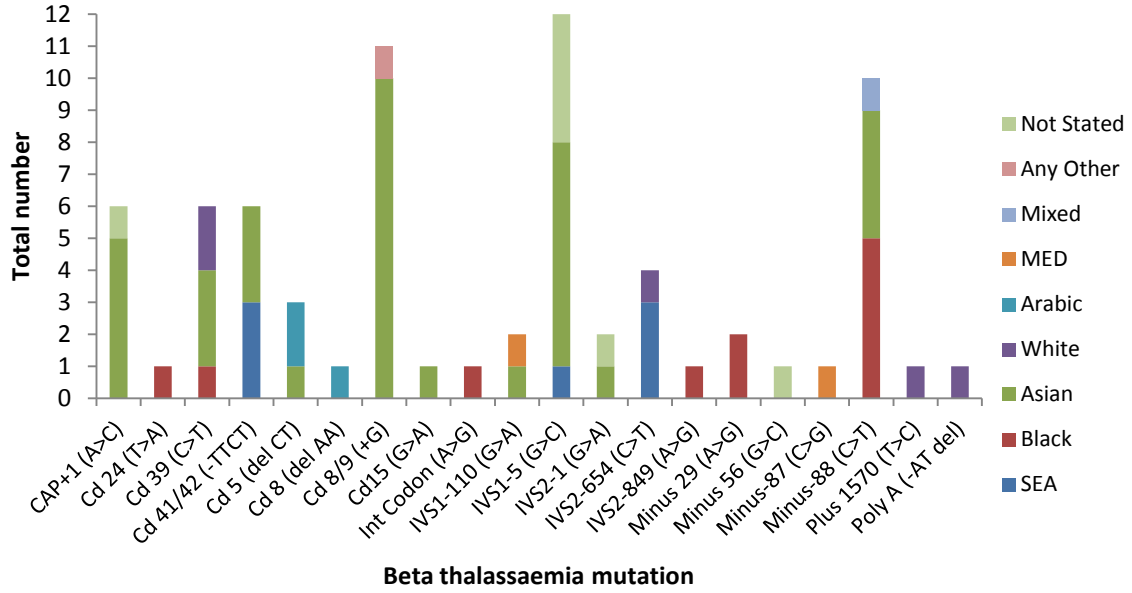


Figure 3.12: Range and distribution of beta thalassaemia mutations within each ethnic group.

Note: Figure 3.12 represents the total number of beta mutations detected and not the total of patient cases with a beta mutation, as one patient was compound heterozygous two beta thalassaemia mutations.

Figure 3.12 shows that 20 different mutations associated with beta thalassaemia were detected within the referral population. The largest number of beta thalassaemia mutations detected were in patients from Asian ethnic groups, comprising 49% (*n=36*) of all beta thalassaemia mutations detected, followed by the Black ethnic group 15% (*n=11*) and patients in SEA ethnic groups, which comprised 10% (*n=7*) of all beta thalassaemia mutations detected.

The total number of beta thalassaemia mutations (*n=73*) found was further categorised based on the clinical severity of the mutation and then compared to the patient ethnic group. Note that the total number of patient cases with beta thalassaemia (*n=72*) is not the same as the total number of beta thalassaemia mutations found (*n=73*) because one patient was compound heterozygous for the CAP^{+1} and IVS1-5 thalassaemia mutations.

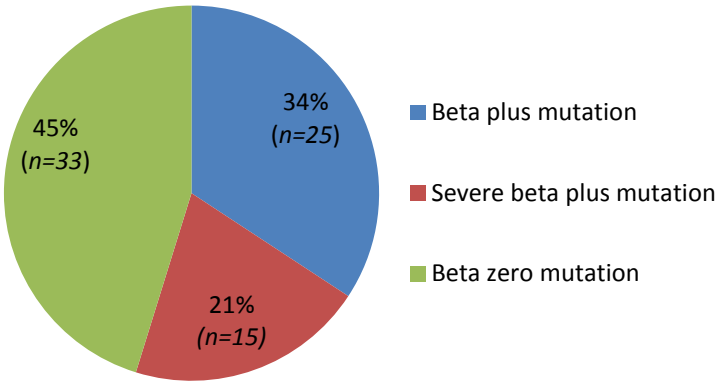


Figure 3.13: Clinical severity of all beta thalassaemia mutations detected.

NOTE: In figure 3.13 one patient was compound heterozygous for a beta plus and a severe beta plus mutation and was included in each category.

Figure 3.13 shows that a total of 34% (*n=25*) of patients had a beta plus thalassaemia mutation, 21% (*n=15*) had a severe beta plus thalassaemia mutation and 45% (*n=33*) had a beta zero thalassaemia mutation.

The total number of clinically significant beta mutations (i.e. beta zero thalassaemia) was highest in those patients from Asian ethnic groups, comprising 26% (*n=19*) of all beta thalassaemia mutations detected. This ethnic group also had the largest number of severe beta plus (10%, *n=7*) and beta plus thalassaemia mutations (14%, *n=10*).

Within the 'White' ethnic group a total of 5 beta thalassaemia mutations were found of which 40% (*n=2*) were clinically significant mutations associated with beta zero thalassaemia. Furthermore, within the 'Arabic' ethnic group a total of 3 thalassaemia mutations were detected, all of which were beta zero thalassaemia mutations.

Screening for beta thalassaemia mutations by DNA sequencing helped to clarify the genotype in young children referred for query beta thalassaemia major, in which trace amounts of Hb A had been detected on HLPC. In many of these cases genetic analysis revealed homozygosity for severe beta plus thalassaemia, often the IVS1-5 (*G>C*) mutation. It also helped to diagnose beta thalassaemia in those patient cases that had been referred with borderline HbA_2 levels, in which case the CAP^{+1} mutation was the most common causative mutation, often found in patient from the Asian ethnic group.

6.2 Mutations detected within *HBB* associated with a haemoglobin variant.

Figures 3.14a and 3.14b show the range of beta mutations detected that are associated with a haemoglobin variant and their distribution within each ethnic group.

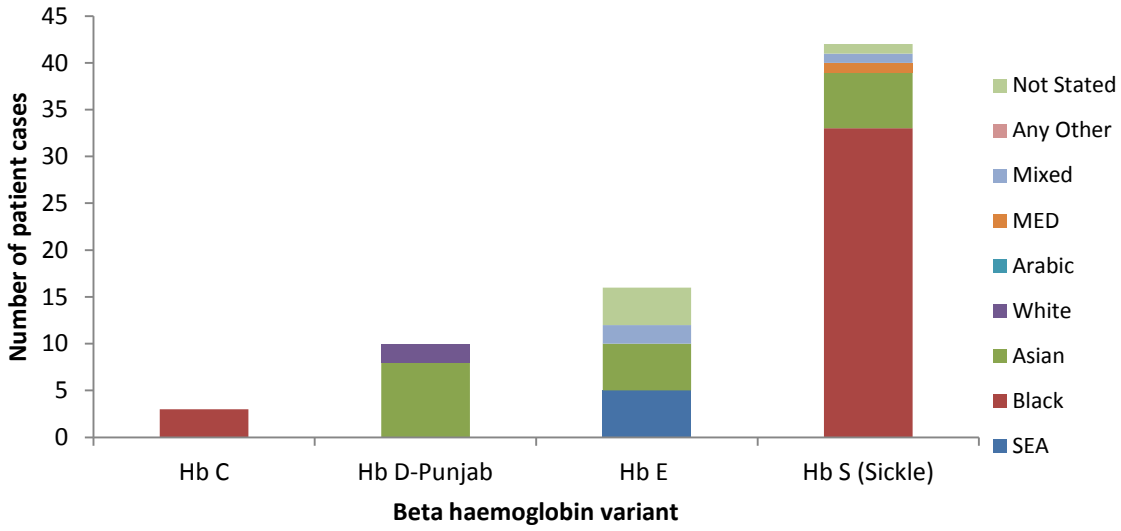


Figure 3.14a: Total number of clinically significant *HBB* variants within each ethnic group

Figure 3.14a represents the total number of clinically significant variants (as outlined by the Antenatal Screening programme) found within the referral population. There was a range of 4 such variants detected in a total of 71 patients. Note there were no genetic referrals received from patients with the Hb O-Arab genotype. The most prevalent clinically significant haemoglobin variant was Hb S which was found in a total of 42 patient cases, comprising 40% (*n=42*) of all patient cases in which a beta chain haemoglobin variant was found. Hb E was found in a variety of ethnic groups.

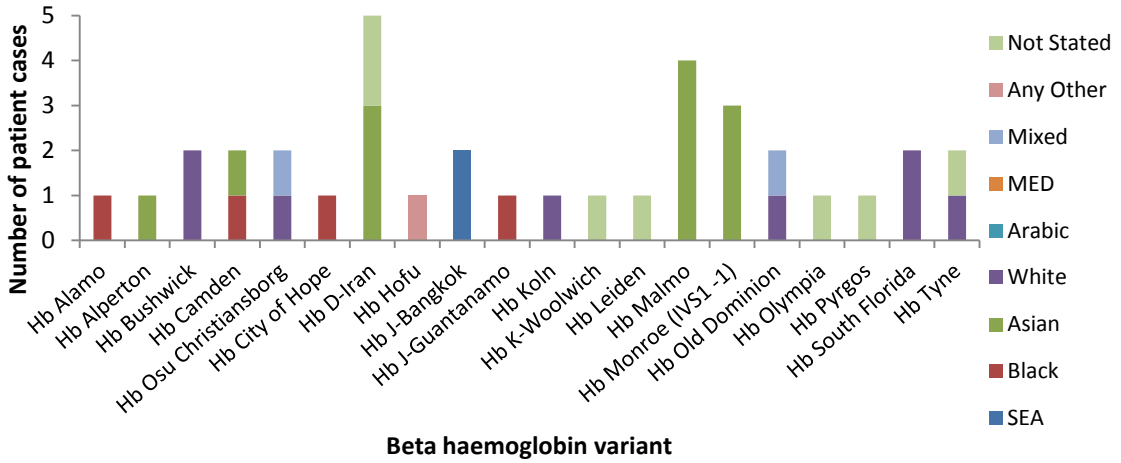


Figure 3.14b: Total number of 'clinically benign' *HBB* variants within each ethnic group

Figure 3.14b represents the total number of cases with a 'clinically benign' *HBB* haemoglobin variant and their distribution within each ethnic group. There were 20 different beta chain haemoglobin variants detected in 36 individuals within a variety of ethnic groups, with those of Asian ethnicity comprising the largest proportion (*n=12*), accounting for 33% of these cases.

The overall incidence of beta haemoglobin variants was highest in individuals from Black ethnic groups comprising 36% (*n=38*) of cases in which a beta chain variant was identified, followed by Asian 22% (*n=23*) and White ethnic groups 10% (*n=10*). The highest incidence of variants reported in patients from Black ethnic groups was accounted for by the large number of cases with sickle haemoglobin (Hb S) which comprised 31% (*n=33*) of all beta haemoglobin variant cases detected and 79% (*n=33*) of all sickle haemoglobin cases.

According to the antenatal screening programme, the 20 beta chain haemoglobin variants shown in Figure 3.14b are considered not to be of clinical significance, but some of these 'clinically benign' variants may actually be associated with a thalassaemic phenotype. Further investigation shows that 10 of these haemoglobin variants, found in 50% (*n=18*) of patients are associated with a completely normal clinical presentation, including Hb Alamo, Hb Alperton, Hb Camden, Hb City of Hope and Hb OSU-Christiansborg. However, the remaining 10 variants are not completely benign due to their altered protein structure, for example Hb Monroe, detected in 8% (*n=3*) of these patients is classified as a thalassaemic haemoglobin variant (according to the Hb variant database) due to its highly unstable nature.

Hb Leiden, Hb Bushwick, Hb Hofu, Hb J-Guantanamo, Hb Köln and Hb Tyne, found in 22% (*n=8*) of these patients all have an unstable protein structure which may result in a severe haemolytic anaemia crisis when induced by oxidative stress. Furthermore Hb Malmö, Hb Old Dominion and Hb Olympia, found in 19% (*n=7*) of these patients are all associated with marked erthrocytosis, which had potential implications for a family referred to ourselves, which is discussed further in Chapter 4. The highest incidence of these types of variants was found in individuals of Asian ethnic origin followed by White ethnic groups.

Therefore, the identification of these variants by DNA sequencing may be of clinical significance, especially in the antenatal context where such haemoglobin variants may be co-inherited with another thalassaemic allele.

6.3 Population genetics of all causative mutations detected within *HBB*.

The total number of cases in which a causative mutation was detected within *HBB* (i.e. thalassaemia or haemoglobin variant) was correlated with patients' stated ethnicity by using previously reported population frequencies of mutations.

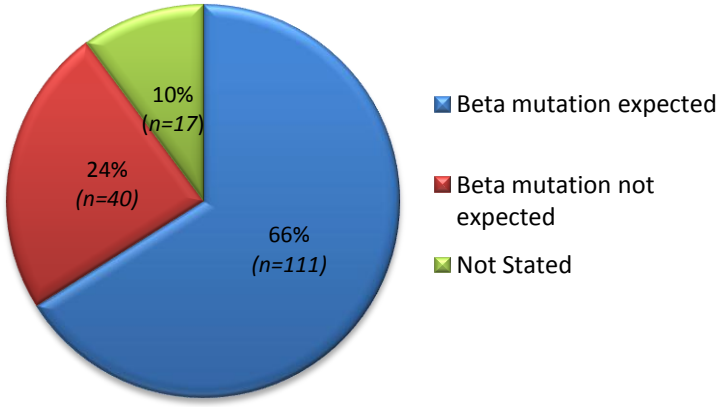


Figure 3.15a: Proportion of all beta genotypes according to expected ethnicity.

Figure 3.15a above, shows that 66% of patients (*n=111*) had a beta genotype that would be expected within their ethnic group and 24% (*n=40*) had a beta genotype that would not normally be predicted to be associated with the patients' stated ethnicities. In the remaining 10% (*n=17*) of cases the ethnicity was not stated.

The total number of cases in which the beta genotype was expected within their ethnic group was further divided into those patients which had mutations commonly found in their ethnic group and those which had mutations rarely found, as shown in Figure 3.15b:

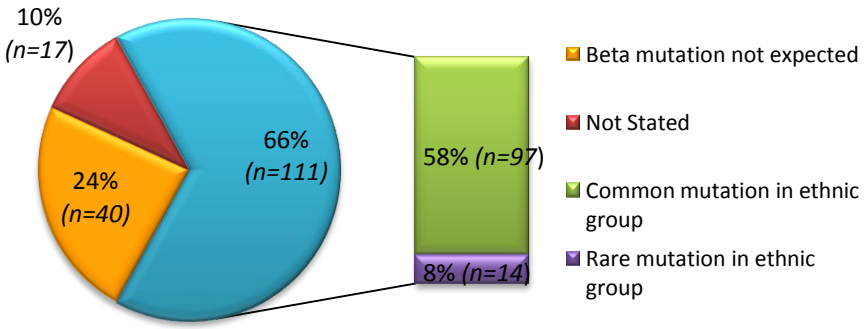


Figure 3.15b: Proportion of rare and common beta mutations in those patient cases with a mutation expected within their ethnic group.

Figure 3.15b shows that out of the 111 cases carrying a beta mutation which was expected within their ethnic group, 87% (*n=97*) had mutations commonly found within the patients' ethnic group and 13% (*n=14*) had mutations that are rarely found within the ethnic group i.e. these cases which would require more extensive targeted mutation screening in order to detect these mutations.

The patients' stated ethnicity was then correlated according to their beta genotype and the expected ethnicity in which their genotype (i.e. the mutation) would normally be expected.

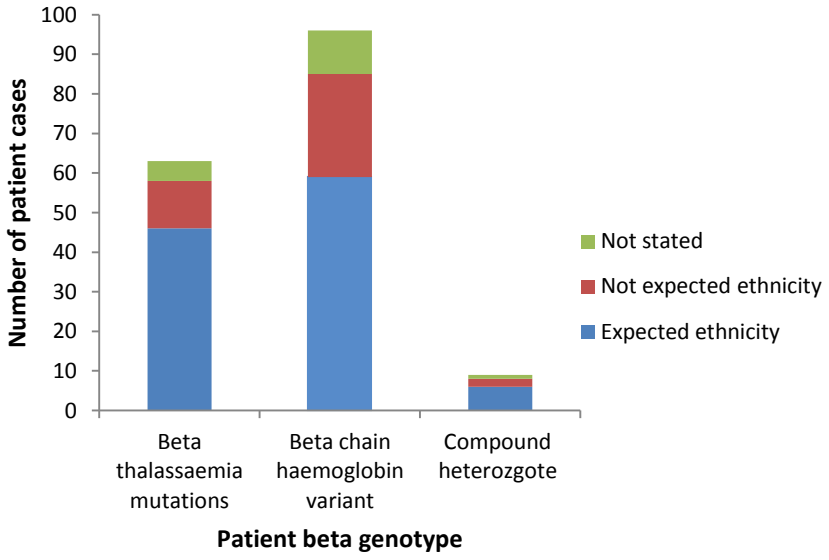


Figure 3.16: Proportion of beta genotypes according to expected ethnicity.

Figure 3.16 shows that the correlation between the patient's ethnicity and the mutation expected within their ethnic group was weakest in those patients with a mutation resulting

in a beta chain haemoglobin variant, in which 27% (*n=26*) of these cases had a haemoglobin variant which would not be expected within their ethnic group.
Further analysis of the correlation between the patients stated ethnicity and the beta chain haemoglobin variant was done. Those patients with a clinically significant haemoglobin variant (as outlined in Figure 3.14a) showed the strongest correlation, in which 75% (*n=53*) of these patients had a variant which would be expected within their ethnic group. This correlation was less noticeable in the 'clinically benign' beta chain haemoglobin variants (as outlined in Figure 3.14b) as only 39% (*n=14*) of these patients had a variant that would be expected within their ethnic group, for example all three reported cases of Hb Monroe, normally expected in black ethnic groups, was reported in only patients from the Asian ethnic group.

Conversely, those patients with a beta thalassaemia genotype showed the strongest correlation between the patients ethnic group and the mutation expected, with only 19% (*n=12*) of these patients possessing a mutation which would not be expected with their ethnic group.

These results show that DNA sequencing of the essential regions of *HBB* indicates that a "first pass" ethnic led, targeted mutation screening strategy may miss up to 34% (*n=57*) of beta mutations in our referral population.

7.0 Compound heterozygous genotypes

Compound heterozygosity for either an alpha or beta thalassaemia mutation or haemoglobin variant was detected in 17% (*n=55*) of all patient referrals. A large number (*n=48*) of these patients were compound heterozygous for a deletional form of alpha thalassaemia and a beta thalassaemia mutation, which is recognised to modulate the clinical severity of the phenotype (as discussed in Chapter 1, section 1.4.2), but it can pose a diagnostic challenge as the red cell indices can be 'normalised'. This was encountered in a few patients within our referral population, including an individual from the Black ethnic group who was referred with a normal FBC, but raised HbA_2 level. Further DNA studies revealed homozygosity for the minus 88 (*C>T*) beta plus mutation and $-\alpha^{3.7}$ deletion. Another patient, from the Mediterranean ethnic group had a normal FBC, but with raised HbA2 and Hb F levels, DNA studies revealed compound heterozygosity for the IVS1-110 (*G>A*) beta plus mutation and $-\alpha^{3.7}$ deletion.

Furthermore, in four patient cases compound heterozygosity for a beta zero thalassaemia mutation and the alpha gene duplication aggravated the globin chain imbalance further, resulting in a beta thalassaemia intermedia phenotype.

A large proportion (14%, *n=47*) of all patient referrals were found to be compound heterozygous for deletional alpha thalassaemia and a haemoglobin variant on either the alpha or beta globin genes. A large number of these cases co-inherited the $-\alpha^{3.7}$ deletion with the sickle (Hb S) beta chain haemoglobin, these comprised 6% (*n=19*) of all patients referrals and 35% (*n=19*) of all compound heterozygous patient cases.

8.0 Haemoglobin variants

As highlighted in earlier sections, a diverse range of alpha and beta chain haemoglobin variants were found within the referral population. DNA sequencing was useful in identifying variants in those patients who had been referred with unusual HPLC results, such as an abnormal Hb A or HbA_2 peak morphology, i.e. a shoulder peak or and split peak morphology. The variants detected within our referral population associated with HPLC anomalies included, Hb Alperton, Hb Tyne, Hb Riccarton, Hb Santa Barnabas, Hb Fort Worth, Hb Seal Rock, Hb Spanish Town. Hb City of Hope, which was completely silent on HPLC, was only detected by DNA sequencing because the patient also had Hb S. Interestingly 6 of these 8 variants may be of clinical significance and therefore it was important that their identity was confirmed by DNA studies.

Additionally, 5 haemoglobin variant referrals were received from IVF patients. The variants detected in these patients included Hb Camden, Hb Hofu, Hb J Bangkok, Hb Q-India and Hb Kurdistan, all of which have a normal clinical presentation.

9.0 DNA sequencing of the 5' region of the *HBG2* (^{G}Gamma) gene for the detection of the *Xmn* I restriction site polymorphism and point mutations associated with non-deletion forms of HPFH.

DNA sequencing of the 5' region of the *HBG2* gene was performed on a total of 37 patients. Of these, homozygosity for the *Xmn* I polymorphism was detected in 5 patient cases, all of which also had a beta thalassaemia mutation. In two of these patient cases, both from Asian ethnic groups, this polymorphism may have helped to ameliorate their beta thalassaemia major genotype to beta thalassaemia intermedia (as discussed Chapter 1, Section 1.4.3). Heterozygosity for the *Xmn* I polymorphism was detected in 20 patient

cases, but these were not considered to be of any physiological importance, as this genotype is reported to have a minimal effect on the phenotype of the patient. Other polymorphisms detected by this screening methodology include those found at minus 309, minus 408 and minus 369 nucleotides 5' to the *HBG2* CAP site, which unlike the *Xmn* I polymorphism do not have any physiological implications. Heterozygosity for the minus 202 (*C>G*) non-deletional HPFH mutation (*HBG2*:c.-255 *C>G*) (Collins, F, S *et al*: 1984 and Huisman, T.H.J *et al*: 1975) was detected in one patient from the Black ethnic group who had been referred for Hb S variant screening with a raised Hb F of 21.7%.

10.0 Deletional forms of beta thalassaemia, including HPFH and δβ disorders detected by Gap-PCR.

A deletional mutation mechanism, associated with either HPFH or δβ thalassaemia was detected in 3% (*n=9*) of all patient diagnoses. The range of deletions found within each ethnic group is shown below in Figure 3.17:

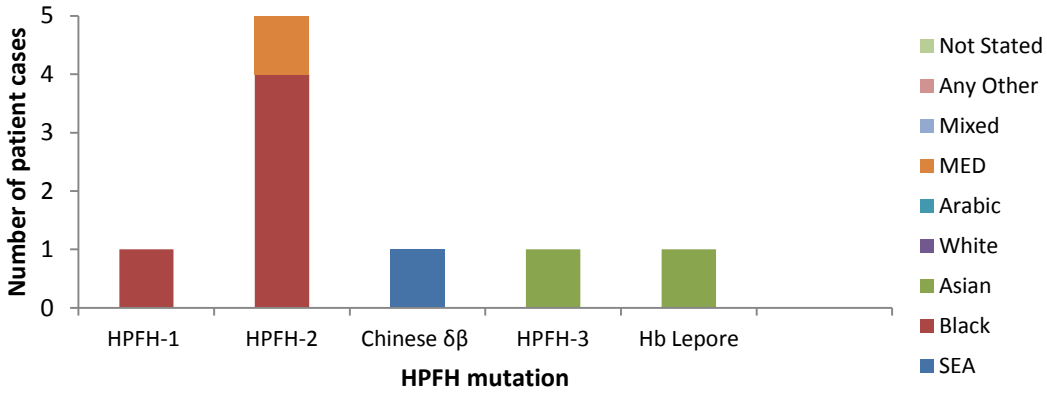


Figure 3.17: Proportion of HPFH and δβ mutations detected within each ethnic group.

HPFH disorders were most prevalent within Black ethnic groups, comprising 67% (*n=6*) of these genotypes detected within the MDC, with the most commonly reported deletion being the HPFH-2 deletion. In four of these patient cases this genotype was co-inherited with the $-\alpha^{3.7}$ deletion.

11.0 Patients in which no causative mutation was detected by the screening methodologies employed within MDC.

In 13% (*n=41*) of all patient referrals to the MDC either no causative mutation was detected using the laboratory screening methods discussed in Chapter 2, or the genotype

detected was not enough on its own to account for the patient's phenotype. The proportion of these cases are represented and discussed further below:

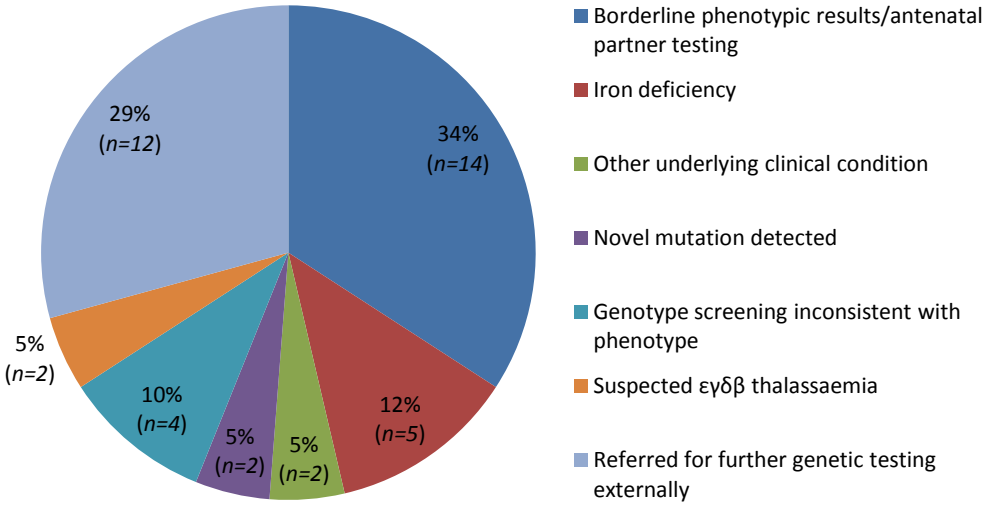


Figure 3.18: Proportion of patients in which no causative mutation was detected by the screening methodologies employed within MDC.

11.1 Borderline phenotypic results/Antenatal partner

Fourteen of these cases had been referred with normal to borderline red cell indices and exclusion of thalassaemia was valuable clinical information, or they were part of an 'at-risk' antenatal couple and clarification of their genotype was useful. No further genetic investigation was performed in these cases. In several cases, repeat phenotypic screening within the MRI of samples referred from an external source revealed normal indices.

11.2 Other underlying clinical conditions

A further 5 patients were diagnosed with iron deficiency and 2 patients had other underlying clinical conditions including a compensation haemolytic anaemia caused by a pyruvate kinase deficiency and myelodysplastic syndrome.

11.3 Novel mutation detected

A novel mutation was detected within the MDC in two patient cases. The first patient was of Northern European descent, referred from an external source for an abnormal haemoglobin variant of 25.9%, with normal red cell indices. DNA sequencing of their *HBA2* detected heterozygosity for a novel mutation within codon 11 (NM_000517.4:c.34A>G). This mutational mechanism that resulted in a haemoglobin

variant was assumed to have a normal clinical presentation, as in a previously described change Hb J-Wenchang-Wuming that affects the same location of the gene (NM_000517.4:c.34A>C).

The second patient was of Middle Eastern descent with a raised HbA_2 of 4.2% and normal red cell indices. DNA sequencing of their *HBA2* globin gene detected heterozygosity for a novel single base change 12 nucleotides into the 3' UTR, non-coding region of the alpha-2 gene (*HBA2*:c.*12G>A). This change was internally reported to be of unlikely clinical significance due to the patient's normal red cell indices.

11.4 Genotype screening inconsistent with phenotype

In 4 patients their normal genotype results were unexplained and were not consistent with their phenotype. Further genetic testing of 3 of these patients by the National Reference Laboratory for Haemoglobinopathies also failed to detect a causative mutational mechanism. The fourth patient, who was male and of unknown ethnicity, had been referred for query alpha thalassaemia because of his significantly reduced red cell indices and was unresponsive in iron therapy, no further action was taken in this case as his Hb level was only 8g/dL and iron deficiency still suspected.

11.5 Suspected εγδβ-thalassaemia cases

Two interesting patient cases are discussed in this section in which a rare form of εγδβ-thalassaemia was suspected after initial extensive genetic screening within the MDC failed to detect a causative mutation. Further analysis of the β-gene cluster by MLPA was subsequently performed once the assay had been established.

The εγδβ-thalassaemia disorders are genetically and phenotypically heterogeneous, caused by a variety of deletional mechanisms that either delete all of the genes within the β-gene cluster or inactivate them as a result of the upstream LCR regions being deleted. This condition presents in the neonatal period with haemolytic anaemia and splenomegaly, with the occasional need for blood transfusions during the first six months of life. Data suggests that heterozygotes for εγδβ-thalassaemia have a phenotype resembling that of a beta thalassaemia carrier, but with more severe microcytosis and hypochromia at birth, and normal levels of haemoglobin F and A_2.(Driscoll, C.M. *et al*: 1989, Rooks, H *et al*: 2005, Furuya, C *et al*: 2008).

The first patient was a young white British female who presented as a neonate with severe anaemia, cardiac failure, hypertrophy and splenomegaly, with a severely microcytic picture (MCV: 54fL, MCH:17.4pg) but normal Hb A and Hb F levels. After extensive analysis on this patient, a repeat sample was requested on which beta MLPA analysis was performed. The result from this analysis confirmed a diagnosis of εγδβ thalassaemia, due to the observation of a large deletion, removing the entire extent of her β-gene cluster. This is was evident from a series of reduced probe heights, as shown in the histogram below in Figure 3.19:

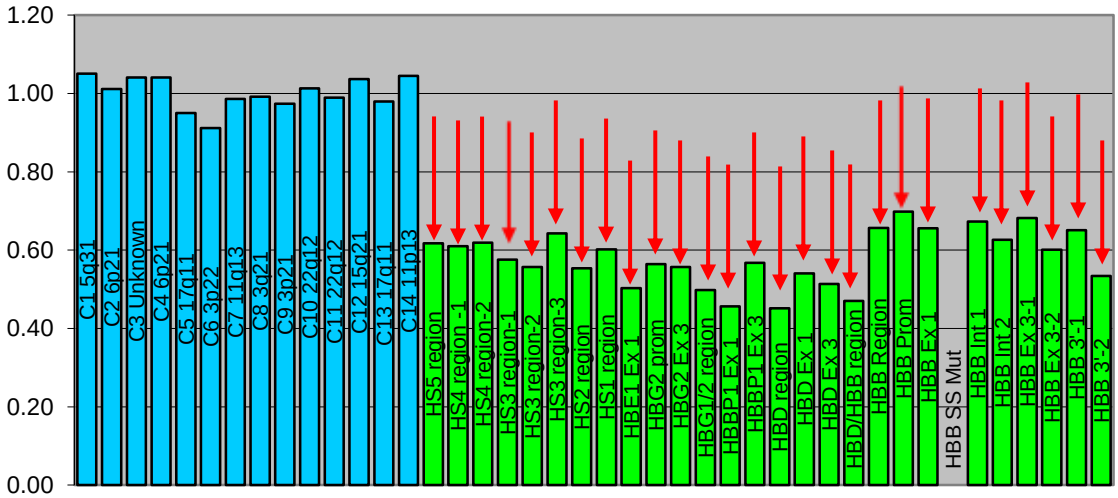


Figure 3.19: Histogram representing the *HBB* MLPA probe data from the patients' DNA sample (the red arrows indicate the probes that showed a reduction in height).

Figure 3.18 show that the probes that showed a relative reduction in probe peak height for that probe amplification product. This indicates a heterozygous deletion that spans across the patient's entire β-globin cluster, consistent with group I εγδβ-thalassaemia. The blue bars represent control probes showing a normal complement of DNA at these locations.

The second patient was a young Caucasian male who presented with a persistently low MCV of 59fl, and an MCH of 18.1 pg, but normal Hb A and Hb F levels. Initial DNA analysis within the MDC detected the triplicated alpha genotype. However, this genotype on its own was not enough to account for the hypochromic indices. *HBB* MLPA analysis was performed as εγδβ-thalassaemia was suspected. His results are shown below in Figure 3.20.

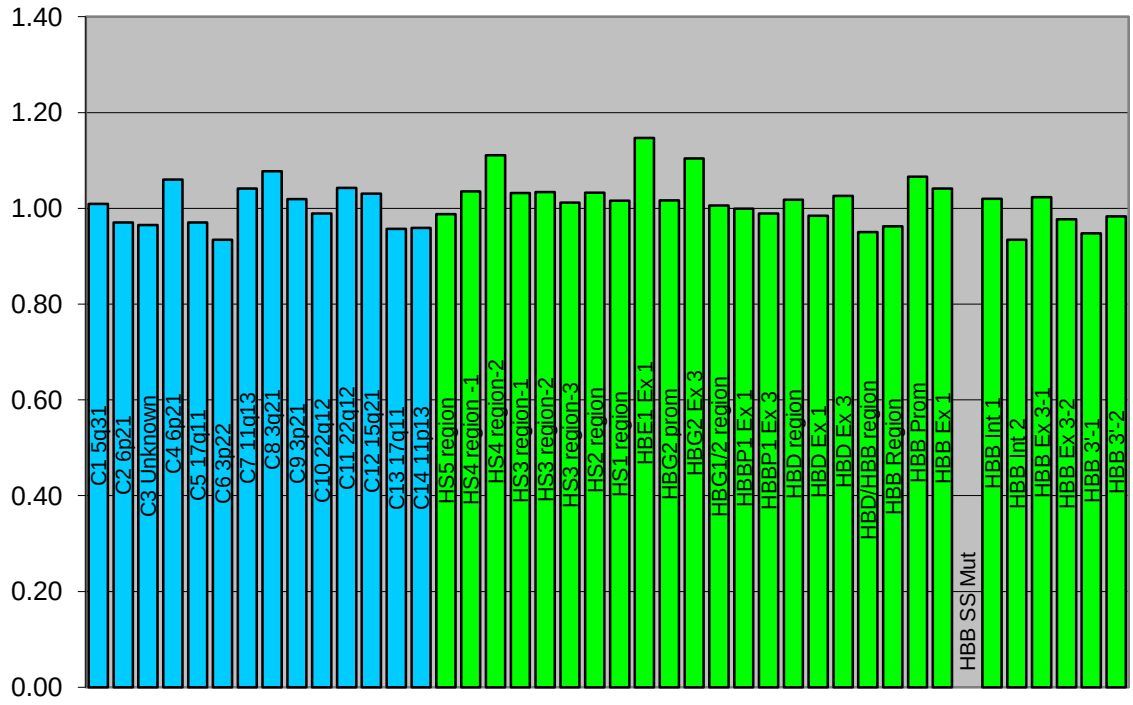


Figure 3.20: Histogram representing the *HBB* MLPA probe data from the patients' DNA sample.

Figure 3.20 shows that the none of the MLPA probes shows a relative reduction in probe peak height. This indicates the absence of a deletion within the β-gene cluster, excluding εγδβ-thalassaemia. This normal result was also confirmed by external genetic testing by King's College Hospital NHS Trust.

11.6 Patients referred to the National Reference Laboratory for Haemoglobinopathies, Oxford for further genetic testing.

In the remaining 12 cases, where failure to detect a mutation using our panel of tests was suspected, further analysis by the National Reference Laboratory for Haemoglobinopathies detected a causative mutational mechanism. These genotypes mainly consisted of deletional forms of beta thalassaemia or novel δβ-deletional mechanisms that could not be confidently characterised by *HBB* MLPA.

These 12 cases are broken down as follows: two patients with the Black 12Kb $\delta\beta^{\circ}$ deletion, one patient with the 909bp Afghan β° deletion, one patient with the Indian 32kb deletion, and one patient with the Indian 619bp β° deletion. Two further patients were reported by Oxford as having a rare, unidentifiable form of deletional HPFH or δβ-thalassaemia. Three other patients were reported as having an unidentified non-deletional

HPFH mutation. One patient was reported by Oxford to have a ‘probable δ-chain variant’, and the remaining patient was suspected to be either normal or to have some form of $\varepsilon\gamma\delta\beta^0$ thalassaemia.

Many of the above patient cases were referred externally for further genetic analysis as *HBB* MLPA was unavailable within the MDC at that time. Such patients were suspected to have an underling deletion if DNA sequencing detected a hemizygous change, for example the Hb S variant in the absence of any heterozygous polymorphisms. Therefore in many of these cases it was not possible to confidently exclude an underlying deletional mechanism such as HPFH or δβ, especially in those patients with a family history of persistently raised Hb F.

Chapter 4.0: Discussion

4.1 Introduction

As discussed in the introductory chapter, the haemoglobinopathies represent the commonest single gene defects in the world and are found at highest frequency in those countries where malaria is prevalent. It has now been recognised that due to the effect of migration they are not confined to specific geographical locations but occur at high frequency in local populations throughout many parts of the world.

Although the haemoglobinopathies are highly heterogeneous at the molecular level with more than 1000 mutant alleles described (Hb Variant database), within each ethnic population group common subsets of mutations are found, consisting of a high frequency of several common mutations with a smaller number of rare mutations.

The impact of immigration from endemic areas to non-malarial countries has significantly increased the range of haemoglobinopathies that need to be characterised by a diagnostic service. Following the implementation of the national NHS Sickle Cell and Thalassaemia (SCT) Screening Programme in England in 2001, the front line element of this process is the offer of phenotypic screening to all antenatal women in high prevalence areas with targeted screening in lower prevalence areas. This screening will identify those individuals who are likely to carry a haemoglobinopathy disorder based on their red cell indices and HLPC results. Unfortunately, due to the molecular heterogeneity of these disorders and the wide range of gene-gene interactions, accurate interpretation of the clinical phenotypic can be difficult. Therefore genetic analysis may be requested for a definitive diagnosis of the main genetic risk combinations as recommended by the antenatal screening algorithm (Appendix 1) in those ethnic groups which are deemed to be at highest risk and permits definitive diagnosis of complex genetic interactions prior to genetic counselling.

This study examines the strategy developed in Manchester which aimed to achieve a high mutation detection rate in the heterogeneous population by developing a comprehensive range of laboratory assays, to ensure that the broad spectrum of genetic defects associated with the haemoglobinopathy disorders could be detected.

4.2 Aims and development of the study

In order to fulfil the study aim of detecting the large spectrum of the mutations likely to be found within the North West (NW) referral population, and to assess the true genetic heterogeneity, a comprehensive mutational screening approach was employed. This included using a variety of PCR based laboratory assays such as multiplex Gap-PCR and direct DNA sequencing of the alpha and beta globin genes, in order to eliminate the need for lengthy targeted mutation screening protocols.

As the study progressed it was clear that development of further assays was needed in order to detect the broad range of different mutational mechanisms encountered within our referral population and in order to confidently asses those patients with complex phenotype and genotype correlations. This included development of alpha and beta gene cluster MLPA analysis for the detection of rarer deletional forms of thalassaemia and the adaption of Gap-PCR and sequencing technologies for the detection of HPFH and δβ thalassaemia disorders.

The mutation frequency data collected within this study was used to assess the genetic heterogeneity of the referral population and the effectiveness of our more comprehensive screening approach could be assessed and evaluated against ethnic led, more targeted, mutation detection approaches.

4.3 Study population

In order to assess the range of genetic defects associated with the haemoglobinopathy disorders, a total of 328 patient referrals from the NW population were genotyped within the Molecular Diagnostics Centre between July 2008 and December 2011. Of these patient diagnoses, 113 patients had a mutational mechanism associated with alpha thalassaemia, 109 patients had a form of beta thalassaemia, 55 were compound heterozygous for an alpha and beta thalassaemia mutations and/or variants and 10 patients exhibited a HPFH or δβ thalassaemia disorder. In the remaining 41 referrals, either no mutation was found or the genotype detected was not enough on its own to account for the patient's phenotype. These are discussed further in Section 4.13.

4.4 Ethnic composition of the North West referral population and incidence of haemoglobinopathies.

A wide range of haemoglobinopathy mutations were found in all ethnic groups, as identified by the Family Origin Questionnaire (FOQ), reflecting the considerable ethnic diversity within the NW referral population (Chapter 3, Figure 3.2). As outlined in Chapter 3, the three ethnic groups which comprised the largest incidence of haemoglobinopathies were the Asian, Black and South East Asian groups. Within the referral population the Asian ethnic group was relatively mixed with different ethnic sub-groups, the largest proportion of patients within the Asian ethnic group were from Pakistani and Indian Asian origin with only a very small proportion of patients of Bangladeshi extraction. Within the SEA ethnic group, the largest proportion were patients of Chinese origin, followed by Thai or Filipino with only a very small number of patients from Indonesian and Malaysian origins. Within the Black ethnic group, nearly all of the patients were of African origin and a smaller proportion from Black Caribbean origin.

The significance of the ethnic composition and the range of mutations found within the North West population will be discussed further in later sections.

Interestingly, 11% of patient referrals were received from those of White British origin, an ethnic group in which haemoglobinopathy disorders are considered rare. Only a small number of patient referrals were received from patients of Arabic and Mixed and Other origins. The smallest proportion (2%) of patient referrals was from patients of Mediterranean origin. Unfortunately the patients' ethnic origin was not stated on the FOQ in 10% of referrals, this was especially common in those patients who had been referred from an external source. The incidence of haemoglobinopathy disorders in relation to their ethnicity could therefore not be assessed. This is surprising as prior information on the ethnic origin of the patient is an important part of the diagnostic strategy, especially in those less heterogeneous populations where patients' stated ethnicity enables a quick identification of the mutations most likely to be present. However, as discussed in earlier sections, this ethnic led strategy becomes less efficient and reliable in those populations such as the North West of England where there is a diverse ethnic mix.

4.5 Summary of all haemoglobinopathy diagnoses according to ethnicity

When the type of haemoglobinopathy disorder was correlated to the patients' ethnicity, the highest incidence of alpha thalassaemia was found in the SEA ethnic groups, beta thalassaemia was found predominantly in patients from Asian origin, and haemoglobin chain variants were found at highest prevalence in the Black and SEA ethnic groups, with a relatively large proportion, from the small number of total referrals from White ethnic groups comprising a haemoglobin variant genotype. Compound heterozygous genotypes were most commonly reported in the Black ethnic groups. These findings from within our referral population correlate with the reported incidence of these genotypes in general literature, although the exact distribution of mutations found within each of these ethnic groups do not, as highlighted in the following sections.

Literature reports that beta thalassaemia is common within the Indian subcontinent and in South East Asia, which loosely correlates with the ethnic composition of beta thalassaemia diagnoses observed in this study. The largest proportion of beta thalassaemia was diagnosed in those individuals of Asian ethnicity, although the prevalence of cases in those patients of Asian Indian origin was no more than in those patients of Asian Pakistani origin.

Variant chain haemoglobins were found at highest prevalence in the Black and SEA ethnic groups. This would be expected as the beta chain variant Hb S is found at highest frequencies within Black African populations and the alpha chain variants Hb Quong Sze, Hb Constant Spring and Hb Q-Thailand are almost exclusively reported in the literature and in our referral population in those patients of SEA origin. Interestingly 39% (*n=14*) of all patient referrals within the 'White' ethnic group comprised haemoglobin chain variants.

Despite the largest proportion of patient referrals received being of Asian origin, only a small proportion of haemoglobin variants were found within this ethnic group. The beta chain variant Hb D-Punjab was almost exclusively found in a small number (*n=8*) of patients within this ethnic group, the other two cases were found in patients of 'White' origin.

Compound heterozygous genotypes were found at highest frequency in patients from the Black ethnic group. This would be expected due to the co-inheritance of the $-\alpha^{3.7}$ deletion with the sickle (Hb S) beta chain haemoglobin commonly reported in this ethnic group.

A relatively large proportion of haemoglobinopathy diagnoses were made in those referrals received from White British patients (*n=36*), an ethnic group in which the haemoglobinopathy disorders are traditionally considered rare. These diagnoses included 5 cases of beta thalassaemia, 7 cases of alpha thalassaemia and 15 cases of haemoglobin chain variants.

Due to the small number of referrals received from the NW population from patients of Arabic, Mixed and Mediterranean origin, only a small proportion of haemoglobinopathy diagnoses were made, therefore, interpretation of the mutational frequencies within these ethnic groups was limited.

4.6 Mutation analysis by Gap-PCR for the detection of deletional forms of alpha thalassaemia

Two independent multiplex Gap-PCR assays (as discussed in Chapter 2) provided an essential screening tool for the rapid detection of a range of common alpha thalassaemia deletions in the heterogeneous referral population found within the North West of England, especially for the diagnosis of compound heterozygous patient cases.

The Liu *et al* multiplex Gap-PCR protocol (discussed in Chapter 2, Section 3.3) failed to provide a true multiplexed assay as the original protocol could not be reliability replicated and all three multiplex PCR reactions required individual optimisation. Subsequently the Chong *et al* based assay (discussed in Chapter 2, Section 3.4) was employed as the primary screening assay for the detection of alpha thalassaemia deletions. A separate assay was set up based on the Eng *et al* protocol for the detection of the --BRIT deletion in selected patients. The Chong *et al* assay and the Eng *et al* assay were optimised using the same PCR kit system, therefore for ease of use in a diagnostic setting they were incorporated into the same SOP document.

Deletional alpha thalassaemia was found in a variety of ethnicities, with the highest frequency reported in patients from SEA and Black ethnic groups. This genotypes showed

the strongest correlation between patients' stated ethnicity and deletion found. The alpha plus $\alpha^{-3.7}$ and alpha zero $--^{SEA}$ deletions were found at the highest prevalence in our referral population. This can be explained not only by the large proportion of referrals from Black and SEA ethnic groups but also the high incidence of the --SEA deletion in Chinese populations and the $\alpha^{-3.7}$ deletion in Black African populations, especially as its co-inheritance with sickle cell haemoglobin offers a selective advantage against malaria. The alpha zero --BRIT deletion was almost exclusively found in patients of White British origin and the --MED deletion exclusively in Mediterranean patients. Conversely, all five cases of the triplicated alpha genotype were in a variety of ethnic groups, reflecting the heterogeneity of the alpha thalassaemia disorders. This supports the findings of a study by Fisher *et al* (2003) which highlights the possible implications of not detecting underlying triplicated alpha thalassaemia when only a targeted mutational screening approach is used in complex ethnic populations. Four of these cases were found in the compound heterozygous state with a beta thalassaemia allele, whilst the other case was found in a young White British male who had been referred with anaemia with a persistently low MCV. This genotype alone was not sufficient to explain his phenotype and further *HBB* MLPA analysis was performed, the findings of which are discussed further in Section 4.12b below.

Deletional Hb H disease was detected in a total of 5 patients, 3 of these were found in patients of SEA origin and was commonly caused by the genotypes: --SEA/-$\alpha^{3.7}$ and --SEA/-$\alpha^{4.2}$, as would be expected in this ethnic group. The other two cases were found in an Arabic patient with the genotype --MED/-$\alpha^{3.7}$ and a White British patient with the genotype --BRIT/-$\alpha^{3.7}$.

4.6.1 Summary of Gap-PCR analysis for the detection of deletional alpha thalassaemia.

Data on different deletional alpha thalassaemia genotypes obtained by Gap-PCR analysis on 324 patient referrals shows that the genotypes found correlated well with the patients stated ethnicity (as shown in Chapter 3, Figure 3.5) and supports reported findings in the general literature.

The highest prevalence of alpha plus thalassaemia was found in individuals of Black-African and Asian-Indian origin, caused by the -$\alpha^{3.7}$ deletion. When a two gene deletion

was found in these ethnic groups, the genotype was always caused by a single gene deletion on each chromosome i.e. homozygosity for the alpha plus $-\alpha^{3.7}$ deletion ($-\alpha^{3.7}/-\alpha^{3.7}$). As expected there were no cases of alpha zero thalassaemia found in the either the Black or Asian ethnic groups, as the literature reports this is extremely uncommon. Deletional alpha plus thalassaemia was also reported in three patients of SEA origin in which the $-\alpha^{4.2}$ deletion was common, although this deletion was also reported in one patient of Asian origin and one of Black origin.

As expected, alpha zero thalassaemia was found in those ethnic groups which are traditionally identified to be at highest risk, such as SEA, in which two genes are deleted from each chromosome by either of the following genotypes: --SEA,--FIL and --MED. Unexpectedly, there was one case of the compound heterozygous genotype $-\alpha^{3.7}/-\alpha^{4.2}$ in a SEA patient, who had been referred for query alpha zero thalassaemia. One further similar case has also been reported since the completion of this study. This demonstrates the importance of Gap-PCR analysis in order to distinguish the genotypes in those MED or SEA individuals where both α^{+} and α^{0} thalassaemia are common.

Interestingly, there was a large proportion (22%, *n=8*) of alpha thalassaemia found in the White ethnic group. In which five of these cases with alpha zero thalassaemia had the --BRIT deletion. One further case had Hb H disease caused by the co-inheritance of this --BRIT deletion with the $-\alpha^{3.7}$ deletion. The reported prevalence of these disorders within White ethnic groups serves to highlight the importance of incorporating this Gap-PCR assay into our spectrum of diagnostic tests, in order to detect this British deletion which appears to originate in the North-West of England.

The mutational data presented above reflects the overall high frequency of carriers for alpha thalassaemia found throughout the world. This is evident within this study by the high incidence of deletional alpha thalassaemia reported in a wide variety of ethnic groups. The $-\alpha^{20.5}$ deletion was not detected in any of the patient referrals, this may be due to the small proportion referrals received from Mediterranean patients.

4.7 Mutation screening by DNA sequencing of *HBA1* and *HBA2* for the detection of non-deletional forms of alpha thalassaemia

Traditionally, screening for non-deletion forms of alpha thalassaemia using PCR based sequencing techniques was not routinely performed in diagnostics laboratories. This is partly because non-deletional forms of alpha thalassaemia are relatively uncommon and there are inherent problems associated with the amplification of the *HBA1* and *HBA2* genes due to their sequence homology and the GC-richness of their gene sequences. A robust and reliable sequencing protocol was developed within the MDC which allowed routine DNA sequencing of the alpha globin genes in a proportion of patient referrals.

In this study, heterozygosity for five different types of non-deletional alpha thalassaemia mutations was found in a small proportion (*n=6*) of patients, the majority of which were reported in the alpha-2 globin gene. Three of these mutations; IVS-I donor site (-5nts), IVS1-5 (G>A) and IVS1-117 (G>A) result in aberrantly spliced RNA which partially explains their alpha thalassaemia phenotype presentation. In the other mutation; Cd19 (del G), the single nucleotide deletion results in a frame shift introducing a premature stop codon at codon 48. Premature stop codons are often associated with an increased degradation of the mutant RNA transcript (nonsense mediated RNA decay) which again would explain the alpha thalassaemia phenotype associated with this mutation. In the Poly A (-AA) mutation the two nucleotide deletion within polyadenylation site results in an extended RNA transcript due to inefficient transcription termination and subsequently results in a thalassaemic phenotype. It has been reported that these mutational mechanisms found within the polyadenylation site on the *HBA2,* which result in an extended transcript may also interfere with and downregulate the expression of *HBA1* on the same chromosome (Harteveld, C.L et al: 1994 and Weatherall and Clegg: 2001).

Of the six patients in which a non-deletional form of alpha thalassaemia was found, the highest incidence (*n=4*) was found those patients of Asian ethnicity, with the two remaining cases reported in the SEA and Black ethnic groups.

Literature reports that non-deletional forms of alpha thalassaemia are traditionally found at highest incidence within Middle Eastern populations, especially in Saudi Arabia, but this was not reflected from the distribution of mutations found within our referral population, despite the relatively large proportion (*n=15*) of Arabic patient referrals received.

Furthermore the unexpected high incidence of these mutations found in Asian patients may just be a reflection of the large number of referrals received from this ethnic group. These findings are still significant, as none of these mutations would have normally been expected in this ethnic group. Such mutations may have not been detected if an 'ethnic-led' screening approach was used, especially as in four of these patient cases their phenotype did not suggest the presence of an underlying non-deletional α-thalassaemia mutation.

4.7.1 Summary of DNA sequencing of *HBA1* and *HBA2* for the detection of non-deletional forms of alpha thalassaemia

Despite the low frequency of non-deletional forms of alpha thalassaemia found within our referral population, the findings from this study highlight the advantages of employing a comprehensive screening programme that includes DNA sequencing of the α-globin genes. DNA sequencing proved especially valuable in this group of patients as many presented with complex phenotypes, therefore it was not possible accurately assess their alpha thalassaemia status, as a result these mutations may be easily overlooked by phenotypic screening alone.

Subsequently, the true population frequency of non-deletional forms of alpha thalassaemia are likely to be underestimated, especially those affecting the *HBA1* gene, as it is widely recognised that those associated with the *HBA2* lead to a more marked reduction of α-globin production. None of these patients within this had co-inherited a deletional form of alpha thalassaemia but in those ethnic groups where alpha zero thalassaemia is common it is important to identify these compound heterozygote genotypes, especially those mutations within the *HBA2* Poly A signal as they belong to a group of non-deletional mutations that can interact to produce severe Hb H disease or even Hb H hydrops fetalis, in which prenatal diagnosis is indicated.

4.8 Mutation screening by DNA sequencing of *HBB* for the detection of point mutations associated with beta thalassaemia.

Direct DNA sequencing of the beta globin gene provided an essential and reliable screening assay for the heterogeneous patient population found within the North West of England. A single-pass detection analysis could be performed in order to detect point mutations associated with beta thalassaemia and haemoglobin variants. After the initial

primer design had been done in order to amplify the three sections of the essentials regions of *HBB*, no further assay optimisation was needed.

The beta thalassaemia disorders are extremely heterogeneous at the genetic level, with more than 200 causative mutations described in a wide range of ethnic groups. Within this study a total of 308 patients had their beta globin genes sequenced. Within the North West of England referral population a total of 20 different mutations causing beta thalassaemia were found in 72 patient cases (note this figure does not include those mutations resulting in a haemoglobin variant). The most common mutation found was IVS I-5 (*G>C*) (16%), followed by codon 8/9 (+G) (15%), Minus 88 (*C>T*) (14%). The three mutations, Codon 41/42 (−TTCT), Codon39 (*C>T*) and CAP+1 (*A>C*) all comprised 8%.

The severity of beta thalassaemia mutations found was analysed. The largest proportion of patients (45%, *n=33*) had a mutation that would be associated with a clinically significant beta zero thalassaemia genotype. A further 34% (*n=25*) of patients had a beta mutation that was associated with mild beta plus thalassaemia, and the smallest proportion of patients (21%, *n=15*) had a mutation classified as severe beta plus thalassaemia.

The highest prevalence of beta thalassaemia was found in patients of Asian origin, comprising nearly half (49%, *n=36*) of all beta thalassaemia mutations detected. Within this ethnic group the mutations found at highest frequency was the frame shift mutation Codon 8/9 (+G) and the splice site mutation IVS1-5 (*G>C*), both of which are reported in Asian Indian ethnic groups. These two mutations collectively comprised nearly half of all beta thalassaemia cases within this ethnic group, which would explain the highest frequency of β^0 and severe β^+ mutations found in patients of Asian ethnicity. Conversely, these figures may just reflect the large number of Asian patient referrals received and the smaller number of referrals received from other ethnic groups, as the Asian ethnic group also had the highest frequency of mild β^+ mutations.

Homozygosity for a further four beta thalassaemia mutations were found in six patients of Asian origin, the codon 39 (*C>T*), IVS1-110 (*G>A*), codon 5 (del CT) and IVS2-1 (*G>A*), all of which are normally reported in Mediterranean regions. Additionally, DNA sequencing detected the CAP+1 mutation in five patients of Asian ethnicity, which would have required more extensive targeted mutation screening in order to detect this mutation

if the ethnic led approach was employed as it is rarely found in Asian ethnicities. Furthermore, detection of this mutation by DNA sequencing was also helpful as many of the patients with this genotype were referred with borderline HbA_2 levels.

Furthermore, the minus 88 (*C>T*) promoter mutation was detected in 4 patients of Asian Pakistani origin, three of whom were homozygotes. This would be expected within the Asian ethnic group, but in those individuals belonging to the Indian sub-group rather than the Pakistani sub-group. The larger proportion of individuals within the North West population from the Pakistani Asian ethnic sub-group, rather than the Indian subgroup did not seem to affect the range of beta thalassaemia mutations found.

The second highest prevalence of beta thalassaemia was found in the Black ethnic group, comprising 15% (*n=11*) of all beta thalassaemia mutations detected. The minus 88 (*C>T*) promoter mutation was found at highest prevalence (*n=5*) within this ethnic group. This would be expected as published beta thalassaemia distribution maps (such as in Chapter 1, Figure 1.7) show that Africa has an extremely high prevalence of mild beta thalassaemia promoter mutations -29 (*A>G*) and -88 (*C>T*) which are associated with particularly high levels of Hb F (see Chapter 1, Section 1.4.2). This accounts for the large proportion of relatively mild β-thalassaemia cases found in Black African patients (Weatherall, D.J and Clegg, J.B:2001).

It has also been reported that compound heterozygosity for mild β-thalassaemia and Hb S is relatively common in this ethnic group, resulting in a less severe phenotype than an interaction with β^0-thalassaemia. Despite the high frequency of Hb S (*n=42*) found in this study (highlighted in Chapter 3, Section 6.2), only a small number of cases (*n=2*) of compound heterozygosity for mild β-thalassaemia/Hb S was found.

The detection of the -88 (*C>T*) beta thalassaemia mutation and other beta thalassaemia mutations which are rare in the indigenous population of the UK, highlights the effects of immigration to a non-thalassaemia endemic country and the subsequent racial heterogeneity found.

β^0 thalassaemia was found in two individuals of Black ethnicity, caused by the frame shift mutation codon 39 (*C>T*) and the initiation codon (*A>G*) mutation respectively, both of

which would not normally be expected to be found within this ethnic group. The initiation codon (*A>G*) mutation is reported to be associated with particularly high HbA_2 levels (Bain, B: 2006) which was consistent with that found (6.3%) in the affected patient in this study.

A beta thalassaemia mutation was found in 10% (*n=7*) of patients from the SEA ethnic group. The mutations codon 41/42 (-TTCT) and IVS2-654 (C>T) were both found in three patient cases and the IVS1-5 (*G>C*) mutation was found in the remaining patient. All three of these mutations are associated with a severe phenotype and would normally be expected in patients belonging to this ethnic group. None of these 7 affected patients had their phenotype ameliorated by the co-inheritance of deletional alpha thalassaemia.

Despite the large proportion (21%) of SEA patients received for genetic testing within the North West referral population, a higher frequency of beta thalassaemia diagnoses would have been expected to be observed, as literature reports that beta thalassaemia is also most commonly found around the Mediterranean and South East Asia, as well as India and Africa. The very small number of referrals received from patients of Mediterranean origin account for the low frequency observed of some of the beta thalassaemia alleles such as the IVS1-110 (*G>A*) and IVS1-(-1) *G>C*).

A further three cases of beta thalassaemia were found in patients of Arabic origin, caused by two different frame shift mutations, in which a small deletion within *HBB* results in a premature stop codon. Two patients were heterozygous for codon 5 (del CT), and their phenotype was further aggravated by the co-inheritance of the alpha gene duplication, resulting in beta thalassaemia intermedia. One patient was homozygous for codon 8 (del AA) resulting in beta thalassaemia major. None of the beta thalassaemia mutations reported within this ethnic group would have been expected, as they are normally reported in those individuals of Mediterranean origin, therefore ‘ethnic-led’ mutation screening within this ethnic group would have been less effective.

Significantly, there were a total of 5 cases of beta thalassaemia reported in patients of White ethnic origin, an ethnic group in which beta thalassaemia is considered rare. None of the four mutations detected in these individuals had been reported in White British populations. Of these 5 cases, 2 patients were heterozygous for a clinically significant

mutation, codon 39 (*C>T*) which is associated with β^0 thalassaemia. This mutation creates a stop codon at codon 39, resulting in premature termination of translation. Another patient was heterozygous for the severe β^+ mutation IVSII-654 (*C>T*). This mutation creates a new donor splice site specificity, leading to the accumulation of mis-spliced mRNA which then translates into a highly unstable β–chain variant with a dominant negative effect. This has been suggested to be the cause of the unusually severe disease seen in patients heterozygous for this mutation. The affected patient's phenotype was consistent with these findings, as he had marked hypochromia and microcytosis, with a particularly high HbA_2 level of 3.8%.

The remaining two patients were heterozygous for mutations found near or within the Poly A signal of the beta globin gene; the +1570 (from the CAP site) *T>C* nucleotide substitution, which is 12 nucleotides 5' to the poly A site and the Poly A (-AT del) mutation. Both of these mutations result in a mild beta plus thalassaemia phenotype as it has been reported that an intact Poly A signal is not essential for correct cleavage of the mRNA transcript. Interestingly, the patient with the +1570 T>C mutation presented with a normal HbA_2 level of 3.1%, even though this mutation is normally associated with a HbA_2 level of between 4.8 and 6.4%.

4.8.1 Summary of mutation screening by DNA sequencing of *HBB* for the detection of point mutations associated with beta thalassaemia.

The data above reflects the heterogeneity of the beta thalassaemia disorders found within a variety of ethnic groups within the North West referral population and supports the findings of a study by Henderson, S *et al* (2009), in which the haemoglobinopathy mutations found in the UK can be divided into two groups; those very rare ones that are specific to individuals of white British origin, such as those with a dominant β-thalassaemia phenotype, or more often those that are commonly observed in specific ethnic groups from the countries in which thalassaemia is prevalent (i.e. Mediterranean, Asian Indian, Southeast Asian, and Africa).

It is the effect of immigration from thalassaemia endemic regions which poses a challenge for the molecular screening of β-thalassaemia mutations within the UK. Studies have shown that the UK was found to have the largest number of distinct mutations (*n=68*) compared to other countries (Henderson, S *et al*: 2009). It is therefore not possible to

perform a quick ethnic-led screen for the detection of the common mutations in this heterogeneous population. The current study indicates that an ethnic-led, targeted mutation screening strategy may miss up to 34% (*n=57*) of beta mutations in our referral population. This suggests that diagnosis of beta thalassaemia in heterogeneous populations is best performed using methods such as direct DNA sequencing in order to detect the diverse panel of mutations likely to be encountered.

4.9 Mutation analysis by Gap-PCR for the detection of deletional forms of HPFH and δβ disorders

The Gap-PCR protocol used, adapted from Craig, R.A *et al*: 1994 (as detailed in Chapter 2, Section 4) provided a reliable and robust method for the molecular detection of the deletions associated with either the HPFH or δβ-thalassaemia disorders. The KAPA2G DNA polymerase PCR kit worked reliably, but this laboratory assay required extensive optimisation from the original protocol. It could not be multiplexed as each of the nine individual, deletion specific, PCR reactions required different concentrations of $MgCl_2$ and different annealing and extension times. All mutations detected were confirmed either by external genetic analysis or MLPA, although due to lack of positive control samples the following deletions could not validated: Spanish $(\delta\beta)^o$, Sicilian $(\delta\beta)^{o,}$ Asian-Indian $^G\gamma(^A\gamma\delta\beta)^0$ Inversion/Deletion A and B and Turkish $(\delta\beta)^o$ Inversion/Deletion A and B. This did not hamper any patient diagnoses as the HPFH and δβ disorders are relatively easy to identify by phenotypic screening due to their relative increase in Hb F levels and absence of any additional haematological changes in HPFH heterozygotes. The findings in this study correlate with general literature that reports these deletions tend to show geographical and racial clustering. Therefore where the ethnicity of the patient was known, it was used as a guide in choosing the appropriate Gap-PCR assay(s).

Only a small proportion (3%, *n=9*) of all patient diagnoses were heterozygous for either a HPFH or δβ deletion. The HPFH-2 deletion was the most commonly reported deletion as it was found in 5 patient cases. The remaining 4 patient cases had either the HPFH-1, HPFH-3, Hb Lepore or Chinese δβ deletion. Seven of the nine cases had a deletion that would be expected within the patients' ethnic group, including HPFH-1 and 2 deletions in black ethnic groups, HPFH-3 in an Indian patient and Chinese δβ in a patient of Chinese origin. The remaining two patients had a deletion that would not normally be expected within their ethnic group, including a HPFH-2 deletion in a patient of Mediterranean

origin and Hb Lepore in a patient of Bangladeshi origin, although in this latter case it is difficult to accurately correlate the patient's ethnicity to the genotype as there are three different molecular varieties of Hb Lepore (Hb Lepore Washington/ Boston, Baltimore and Hollandia) that are indistinguishable by phenotype, of which only Hb Lepore Hollandia has been reported in Bangladeshi populations.

None of the above nine cases were co-inherited with a beta thalassaemia mutation, but 3 patients were compound heterozygous for the HPFH-2 deletion and the $-\alpha^{3.7}$ deletion and either the Hb S or Hb C variant, which is in agreement with general literature that suggests these combinations tend to be restricted to either the HPFH-1 or HPFH-2 deletions.

The ability to distinguish between HPFH and δβ-thalassaemia rapidly using Gap-PCR methodology proved useful in the antenatal context in three patient cases in which the partner was found to be either heterozygous for β-thalassaemia or the β^S variant. Furthermore, detection of HPFH deletions by this methodology helped to diagnose and differentiate those severe genotypes associated with sickle cell disease from those mild genotypes associated with HPFH deletions.

4.9.1 DNA sequencing of the 5' region of *HBG2* (^{G}Gamma) for the detection of the *Xmn* I restriction site polymorphism and point mutations associated with non-deletional forms of HPFH.

DNA sequencing of the 5' region of the *HBG2* ($^G\gamma$) gene was originally set up for the detection of the *Xmn* I polymorphism in patients in which it may serve to ameliorate their genotype i.e. those with beta thalassaemia intermedia, major or the β^S variant, as it is thought to only have a minimal effect on the synthesis of Hb F in normal individuals (See Chapter 1, Section 1.4.3). This assay also detected the promoter point mutations within *HBG*, found between minus 175 and minus 202 nucleotides (relative to the CAP site) that are associated with non-deletional forms of HPFH (as described in Chapter 1, Section 1.4.3).

Heterozygosity for the *Xmn* I polymorphism was detected in a total of 20 patient cases (as discussed in Chapter 3, Section 8) all with a wide variety of associated genotypes. Six of these patients had the β^S variant, including one case that had sickle cell disease in the presence of a particularly raised haemoglobin F level of 20.8%, suggesting a HPFH

mutational mechanism was involved. DNA sequencing revealed that this individual was homozygous for the *Xmn* I polymorphism, therefore contributing to her high Hb F levels. The remaining cases are not discussed further due to the reported limited effect of heterozygosity for this polymorphism on thalassaemia genotypes.

Homozygosity for the *Xmn* I polymorphism was detected in 5 patient cases with a beta thalassaemia mutation. Two cases were found in unrelated patients of Asian origin who were homozygous for the β^+ IVS1-5 (*G>C*) and β^0 IVS2-1 (*G>A*) mutations respectively. Interestingly, a study by Ho *et al* (1998) showed that there was a strong association of this polymorphism within Asian Indian populations with the β^0 IVS1-1 (*G>T)* mutation. The presence of the polymorphism in these two cases may have helped to ameliorate their beta thalassaemia major genotype to beta thalassaemia intermedia. Due to the limited number of patient samples analysed within this study, together with the limited clinical information available on some patients, it was not possible to truly asses the modulating effect of this polymorphism in our referral population, which was further compounded by the diverse range of beta genotypes found.

Heterozygosity for the minus 202 (*C>G*) HPFH point mutation (*HBG2*:c.-255 *C>G*) was found in a male Black African patient, who had been referred with a Hb S level of 47% and a Hb F level of 21.7%. This mutation has been previously described in Black patients resulting in a HPFH phenotype (Collins, F, S *et al*: 1984 and Huisman, T.H.J *et al*: 1975, Hb Variant database). As the literature supports, this compound heterozygous genotype does not result in a marked anaemia as the patient only had slightly reduced MCV and MCH values.

4.9.2 Summary of results obtained by DNA sequencing of the 5' region of *HBG2* for the detection of *Xmn* I restriction site polymorphism and point mutations associated with non-deletional forms of HPFH.

Findings from the phenotypic analysis of individuals within this study support the view that the regulation of Hb F production is a complex subject area, and the variation of Hb F levels observed between patients is yet to be fully understood (Chapter 1, Section 1.4.3). The degree to which Hb F was elevated varied widely between patients, even between those members of the same family or those with the same mutation. Some of this variation may be explained by the different haplotypes of the chromosome carrying the allele, such

as the β^S variant (as described in Chapter 1, Section 1.4.4), other factors linked to the β-globin cluster are known to affect gamma gene expression such as the *Xmn* I $^G\gamma$ polymorphism and non-deletional point mutations within *HBG2*. Therefore detection of these sequence changes may be useful in a research context as they could serve as a prognostic factor in those patients with sickle cell anaemia or β-thalassaemia or for therapeutic purposes in order to predict a patients' response to hydroxycarbamide therapy.

It is also worth mentioning that in this study only those sequence changes associated with non-deletional forms of $^G\gamma$ HPFH and not $^A\gamma$ HPFH were analysed, as the same mutations have been reported in the same position 5' within these two genes, due to their sequence homology. Other genetic factors known to segregate independently of the β-globin cluster could also account for some of the variability of Hb F observed (as mentioned in Chapter 1, Section 1.4.3).

4.10 Mutation screening by DNA sequencing of *HBB, HBA1* and *HBA2* for the detection of variant haemoglobins.

DNA sequencing of the alpha and beta globin gene detected a wide range of haemoglobin variants within the study population, including those variants commonly associated with disease such as haemoglobins S, C and E, but there are also rarer haemoglobin variants that do not always result in a clinically significant haemoglobinopathy.

The suspected presence of a structural haemoglobin variant was provided by either HPLC or IEF phenotypic screening, clinical or family history or from peripheral blood film examination with evidence of haemolysis. In this study, characterisation of variant haemoglobins by DNA sequencing was essential for distinguishing between apparently identical variant phenotypes and also to confirm the genetic identity of variable phenotypes, including the differentiation of relatively benign genetic interactions from those that result in disease, for example differentiating Hb E homozygosity from Hb E/β thalassaemia.

4.10.1 Clinically significant variants detected by DNA testing that result in disease

The variants discussed in this section include those variant haemoglobins, detected by DNA sequencing of *HBB,* of clinical relevance as outlined by the Antenatal and Newborn screening programme including haemoglobins S, C, D-Punjab, E, Lepore and O-Arab.

There are no haemoglobin variants reported within the alpha globin genes that must be detected in this context.

Primary identification of these variants was based solely on the phenotypic screening by HPLC and in some circumstances secondary screening by IEF, further genetic testing was requested in those patient cases where a definitive diagnosis was needed. Haemoglobins S, E, D-Punjab and C were found in 22% (*n=71*) of all patient referrals. No patient diagnoses were made with the genotype Hb O-Arab, even though it would have been expected within our referral population group as it has been reported in a variety of ethnic groups, although at low frequencies.

4.10.1a β^S variant

The Hb S variant was found at highest frequency (*n=42*) within our referral population. It was mainly detected in those patients of Black ethnicity (79%, *n=33*) with a further six cases found in the Asian Indian ethnic group. The two remaining cases were each found in the Mediterranean and Mixed ethnic groups. The ethnic group was not stated in the remaining sickle case. The distribution of the Hb S variant observed in this study would be expected as is most commonly reported throughout African, Mediterranean and Indian regions. An even higher frequency of this variant may have been diagnosed if the North-West of England referral population comprised a larger proportion of patients from Mediterranean regions.

Extensive DNA analysis on 13 patients with the β^S variant helped to differentiate those with a clinically significant phenotype, caused by the interaction with other thalassaemic genotypes e.g. β-thalassaemia, from those with a mild phenotype caused by an interaction of an ameliorating genotype e.g. HPFH or alpha thalassaemia. This proved especially valuable in the 12 antenatal cases referred. DNA analysis of these 13 patients revealed that 4 were homozygous for Hb S resulting in sickle cell anaemia and 6 patients were compound heterozygous for sickle-cell and β-thalassaemia. Of these, three patients had a mild β^+ mutation, one patient had a severe β^+ mutation and two had a β^0 mutation. The remaining three patients had the compound heterozygous genotype Hb S/HPFH, in which their raised levels of Hb F may have served to ameliorate their phenotype.

There was a remarkable degree of clinical heterogeneity (obtained from the reason for referral information supplied) observed between all sickle patients, which is in agreement with general literature that highlights the highly variable clinical course of sickle cell disease. There is often very limited information to explain this observation. As already highlighted above in Section 4.5, a large number of sickle patients (*n=20*) had a deletional form of α-thalassaemia, which may serve to ameliorate their phenotype, although some studies still debate the significance of this interaction, suggesting that the overall effect of the co-inheritance of α-thalassaemia with Hb S is minimal (Steinberg *et al* 1984 and Vyas *et al* 1988).

The clinical heterogeneity observed, especially between different ethnic groups, may also be partly explained due to the heterogeneity of the β-thalassaemia mutations co-inherited with this variant. It is well recognised that Africans with the β^S variant have a mild clinical course due to the elevated levels of Hb F associated with the co-inheritance of mild beta promoter mutations. Interestingly, in this study there were two African sickle-cell trait individuals that presented with a particularly severe phenotype. One patient reportedly suffered from severe pain crisis and an additional sickling haemoglobin variant was suspected, but DNA sequencing eliminated this. The other patient had been receiving multiple transfusions even though they only had sickle cell trait and the IVS1-5 (*G>A*) alpha mutation. Furthermore, an Asian patient who was also receiving regular blood transfusions was found to have the genotype β^S/IVS1-110 (*G>A*), which conflicts with case reports of this interaction in Mediterranean populations which suggests that it is only associated with a mild clinical course.

DNA studies also helped to clarify those patients who presented with complex phenotypic results whereby additional bands were observed on HPLC, such as when the β^S variant is co-inherited with the $\alpha^{-G\text{-}Philadelphia}$ variant, which was encountered in 3 cases. Another patient of African origin who was referred for characterisation of an additional variant found co-eluting with haemoglobin HbA_2, was identified to be homozygous for the *HBA2* gene variant Hb Fort Worth. There is limited clinical information in this interaction, but it has been reported to show thalassaemic properties and may result in microcytosis. DNA studies also helped to eliminate underlying beta thalassaemia in this case as it was not possible to accurately quantify the amount of HbA_2.

DNA sequencing also detected the beta gene variant Hb City of Hope in a Black African patient who had been referred for sickle cell and query underling iron deficiency/ alpha thalassaemia. This additional variant was 'silent' on HPLC as it co-migrates with haemoglobin A. Again, there was limited data on the clinical effects of this interaction.

There were no cases of sickle cell disease caused by the interaction of haemoglobin S with haemoglobins C, D-Punjab or O-Arab.

4.10.1b β^E variant

Hb E, which is associated with a thalassaemic phenotype, was found in only a small proportion (5%, *n=16*) of all diagnoses, but it was found in the largest variety of ethnic groups, including 5 in SEA patients, five Asian patients and two in patients of mixed ethnic origin (one Asian and White European and one White British and Thai). The ethnic group was not stated in the remaining 4 cases.

DNA studies on this group of patients was especially valuable in order to provide a definitive diagnosis, as there is huge clinical variability in the range of conditions associated with Hb E disorders (as highlighted in Chapter 1, Section 1.4.4). Compound heterozygosity for β^E and either α^0-thalassaemia or β^0-thalassaemia is not uncommon due to the high frequency of these alleles found within Thai, South East Asian and Indian populations.

DNA analysis is a vital tool in order to differentiate between the different Hb E disorders as Hb E co-migrates with HbA_2 on HPLC. It is difficult to distinguish by phenotypic by screening alone homozygous E patients (a asymptomatic condition) which is of no clinical significance, from those compound heterozygous cases of Hb E/β-thalassaemia (resulting in a more severe clinical picture). Therefore it is not surprising that the largest proportion of these patient referrals (*n=10*) were for clarification of their ambiguous phenotypic screen where underlying beta thalassaemia was suspected.

HBB DNA sequencing of these 16 patients revealed that nine had homozygosity for Hb E. Compound heterozygosity for Hb E and a β-thalassaemia allele was only found in one patient of Asian origin, who had the rare β-chain variant Hb Monroe, which is also associated with a thalassaemic phenotype. This interaction is not only rare in general but

particularly within the patient's ethnic group as Hb Monroe is usually reported within Black ethnic groups. Due to the limited amount of literature on this interaction, its clinical consequence is unclear.

The remaining six patients were referred from high risk ethnic groups in which underlying α-thalassaemia was suspected due to their low percentage of Hb E. Of these, two patients had the alpha zero --SEA deletion, one of whom was of mixed ethnicity, three patients had the alpha plus $-\alpha^{3.7}$ deletion, including a homozygous case found in a patient of SEA origin, and one patient, of Malaysian origin had the Poly A (-AA) non-deletional alpha thalassaemia mutation which is normally reported in Indian populations.

Even though there were no cases identified in this study, Hb E is commonly found in the same population groups in which α^0 thalassaemia and alpha chain haemoglobin variants occur, for example Hb Constant Spring in SEA. These compound heterozygous states can result in particularly severe forms of Hb H disease and are therefore important to identify by employing a comprehensive DNA testing strategy.

4.10.1c $\beta^{\text{D-Punjab}}$ variant

The Hb D-Punjab variant was found in 3% (*n=10*) of all diagnoses. As expected it was reported at highest frequency (*n=8*) in patients of Asian Pakistani ethnicity, as generally reported. It does not commonly occur in those of Asian Indian ethnicity. The two other cases were in patients of White British origin which is in agreement with reports that it exists at low prevalence amongst Caucasian populations within England.

Again, DNA sequencing provided a valuable tool for definitive diagnosis in this group of patient, as phenotypic screening by HPLC was unable to identify the exact variant within the 'D/G window'. In this antenatal context it was important to distinguish Hb D-Punjab from other alpha or beta chain variants with lesser clinical significance.

DNA sequencing revealed homozygosity for Hb D-Punjab in two patient cases, which would only result in a mild clinical phenotype, and heterozygosity in 6 patient cases. Compound heterozygosity for Hb D-Punjab and beta thalassaemia was detected in two patients, one with the IVS1-5 (*G>C*) mutation and one with the codon 15 (*G>A*) mutation,

which was also co-inherited with a non-deletional alpha mutation (Poly A -AA). These results confirm general findings that this variant tends to be co-inherited with β^0 thalassaemia mutations.

4.10.1d β^C variant

All 3 cases of this variant were found in patients from the Black ethnic group. One patient had a common compound heterozygous genotype Hb C/β^+ thalassaemia caused by the -88 (*C>T*) mutation. The other two cases showed compound heterozygosity with the HPFH-2 deletion. This interaction is commonly reported in patients with Hb C, as both of these genotypes usually occur in African populations.

Despite the large number of patient referrals received from Black ethnic groups, only a small number of patients were diagnosed with the Hb C genotype, therefore meaningful conclusions cannot be drawn. However, the findings above support general literature reports stating that this group of disorders tend to present with mild clinical manifestations mainly in African populations, due to the association with mild beta thalassaemia promoter mutations such as -88 (*C>T*) and HPFH deletions.

4.10.1 Summary of clinically significant beta chain variants detected by DNA testing that result in disease.

There was a strong correlation observed between those patients with a clinically significant haemoglobin variant and their stated ethnicity, which would be expected due to their high frequency found throughout the world. Even though phenotypic screening by HPLC and IEF alone are sufficient to identify these haemoglobin variants, DNA sequencing provides a vital tool in differentiating their homozygous states (e.g. Hb EE) from their compound heterozygous states with a β-thalassaemia mutation (e.g. Hb E/β^+ or β^0-thalassaemia), all of which result in very different clinical disorders, the latter of which is associated with a more severe clinical phenotype. Characterisation of the exact genotype, by DNA sequencing is vital in order to accurately predict the genetic risk to offspring in those ‘at risk’ antenatal couples. Furthermore this may serve to relieve any unnecessary worry by differentiating those clinically significant disorders from those benign disorders.

4.10.2 Rare variants detected by DNA sequencing of the alpha and beta globin genes

There are a further group of rarer haemoglobin structural variants that may either be associated with a form of thalassaemia or may be clinically 'silent'. The identification of these haemoglobin variants is often presumptive as it is based on their electrophoretic mobility on HPLC, family history or ethnicity or by other clinical characteristics presented by the individual. This section evaluates the merits of identifying the wide range of variants found within our study at the molecular level by DNA sequencing, not only in relation to their clinical significance, but also because of their common co-inheritance with the thalassaemia disorders within the same populations.

The Antenatal and Newborn screening programme does not generally indicate DNA testing on those patients with haemoglobin variants, apart from those clinically significant ones mentioned in Section 4.10.1. However, BCSH genetic guidelines for the diagnosis of significant haemoglobinopathies (Ryan, K *et al*: 2010) indicate that haemoglobinopathy investigations may also encompass those clinical disorders in which there is an unexplained anaemia, splenomegaly, haemolysis, polycythaemia or cyanosis, with normal oxygen saturation. These disorders may be caused by the synthesis of an abnormal haemoglobin variant in which the causative mutation causes destabilisation of the haemoglobin tetramer. The varying properties of an abnormal haemoglobin variant determine the nature of the actual clinical manifestations and may encompass those variants that are associated with either an unstable protein structure or altered oxygen affinity, i.e. increased or decreased, or those associated with a thalassaemic phenotype.

Therefore the significance of the detection of these haemoglobin variants by DNA sequencing within this study are discussed in the following sections below, in terms of their clinical manifestations and the globin chain in which they are found.

In this study a total of 82 patients were referred for DNA analysis for a suspected haemoglobin variant that had either been detected by HPLC or for the reasons mentioned above (therefore the variants discussed in this section exclude those clinically significant variants discussed in Section 4.10.1 that can be provisionally identified by phenotypic screening).

In a large proportion (71%, *n=58*) of referrals in which a haemoglobin variant was suspected, the variant haemoglobin could not be positively identified by phenotypic screening alone and further genetic analysis by DNA sequencing was needed in order to provide a definitive diagnosis.

In a smaller proportion (23%, *n=19*) of cases, phenotypic screening by HPLC was sufficient to identify the variant, although in eight of these patient cases this provisional identification of the variant was aided by the family history i.e. the same haemoglobin variant had been previously reported in another family member.

In a further two patient cases, DNA testing revealed an alternative mutation mechanism i.e. a mutation resulting in a haemoglobin variant was not found. One patient, who had been referred for query α-thalassaemia and a possible variant of 1.9% co-eluting with haemoglobin F was diagnosed with β^+ thalassaemia and homozygous α^+ thalassaemia, further external genetic testing reported a 'probable δ-chain variant'. One patient who was referred in order to exclude a haemoglobin M variant due to her clinical presentation of methaemoglobinaemia was diagnosed with β^+ thalassaemia.

One further patient was diagnosed with a novel mutational mechanism, resulting in a haemoglobin variant within codon 11 of *HBA2* (*HBA2:p*.Lys11Glu). No genetic abnormality was detected in the remaining two patients (as included in Chapter 3, Figure 3.18). One of these patients, who had been referred for explanation of his compensative haemolytic anaemia, was found to have an alternative clinical disorder (pyruvate kinase deficiency). The other patient, who had been referred with a 'flat HbA_2 peak' detected on HPLC, was suspected to have a HbA_2 variant, but no further DNA analysis was requested as these variants are reported not be of any clinical significance.

Furthermore, DNA sequencing on a further eight patient referrals identified a haemoglobin variant that was completely silent on HPLC, and included the *HBB* variant Hb City of Hope (discussed below in Section 4.10.2a) and the *HBA* variants: Hb Evanston, Hb Fontainebleau, Hb Constant Spring, Hb Quong Sze and Hb Seal Rock (discussed below in Section 4.10.2b).

4.10.2a: Beta chain (*HBB*) haemoglobin variants detected by DNA sequencing.

A larger proportion of beta chain haemoglobin variants were found by DNA sequencing than alpha chain variants. Additionally, a significantly larger proportion of patients had a haemoglobin variant genotype (*n=96*) than had a beta thalassaemia genotype (*n=63*).

Excluding those variant haemoglobins already mentioned in Section 4.10.1, a further 20 different beta chain haemoglobin variants were detected in 36 individuals within a variety of ethnic groups. Further investigation of these 20 variants within published literature demonstrated that 10 of them, found in 50% (*n=18*) of patients, were associated with a completely normal clinical presentation, i.e. the variant globin chains are still capable of forming viable haemoglobin tetramers and the variant haemoglobin molecule is synthesised at a normal rate when compared to Hb A. These variants include: Hb Alamo, Hb Alperton, Hb Camden, Hb City of Hope, Hb D-Iran, Hb OSU-Christiansborg, Hb J-Bangkok, Hb K-Woolwich, Hb Pyros and Hb South Florida. The data is inconclusive with the latter variant as there have been no functional studies reported. The mutational mechanism in this variant affects the first codon of *HBB* resulting an elongated β-globin chain suggesting it could potentially affect the quaternary protein structure, but both patient cases in which this variant was found presented with red cell indices within their normal range.

A further 10 haemoglobin variants were detected in the remaining 50% (*n=18*) of patients which were considered to be of clinical significance, either because of their inability to form stable haemoglobin tetramers, or because the tetramers they form are synthesised at a reduced rate resulting in a thalassaemic phenotype caused by the imbalanced α:β globin chain ratio. These are discussed below according to their clinical classification:

Unstable haemoglobin variants

Unstable haemoglobin variants are caused by the inability of the mutated globin chains to combine to form stable, viable haemoglobin tetramers, even though they are synthesised at a normal rate (as highlighted in Chapter 1, Section 1.4.4). Unlike the thalassaemia disorders, these haemoglobin variants tend to follow an autosomal dominant pattern of inheritance and present as a dominant form β-thalassaemia. They result in severe haemolytic anaemia caused by the precipitation of the excess of variant globin chains

within red cell precursors. Therefore affected individuals are almost exclusively heterozygotes.

The largest number of characterised unstable haemoglobins (see Hb Variant database) affect β-globin chains. This probably reflects the more pronounced clinical effects that such β-globin chain variants would have on half of the total haemoglobin formed compared to those found on the α-globin chains which only affect a quarter the total alpha chains formed and are likely to be overlooked and therefore underreported.

For the purposes of this study, Hb Monroe (IVS-I (-1) *G>C*) was included in the discussion of this group of variants, even though it is commonly considered as thalassaemic variant. Hb Monroe was detected in three patients, two of which were found in an Asian Antenatal couple. This highly unstable variant is reported to result in transfusion-dependent β-thalassaemia major when inherited in the homozygous state. In this study there was only one case of compound heterozygosity in which Hb Monroe was co-inherited with Hb E. Furthermore, two patients who were heterozygous for Hb Monroe may have had their phenotype ameliorated by the coinheritance of α^{+} thalassaemia.

In the heterozygous state, the severity of disease caused by unstable haemoglobin variants is variable. They may be associated with reticulocytosis and mild to moderate haemolytic anaemia as a result of the unbalanced globin synthesis caused by the alteration in the relative amounts of α and β globin mRNA (Rieder, R.F and James, G.W: 1976). These clinical symptoms may become more pronounced when exacerbated under conditions of oxidative stress, such as particular drugs or viral infections. At the severe end these unstable variants can mimic symptomatic forms of β-thalassaemia, which is well documented in the case of Hb Leiden (Lie-Injo *et al* 1977), in which compound heterozygosity with β^{0}-thalassaemia in a Chinese patient had severe haemolytic anaemia.

A further six different unstable variants were detected in a total of eight patients referrals, they include Hb Bushwick, Hb Hofu, Hb J-Guantanamo, Hb Tyne, Hb Köln and Hb Leiden. Hb Köln and Hb Leiden are also associated with an increased oxygen affinity, which usually leads to polycythaemia, but here their dominant clinical feature is haemolysis due to their highly unstable nature. They are therefore primarily categorised as unstable, rather than high-affinity, variants.

In this group of eight patient referrals received, reticulocytosis was not a prominent feature and their MCV values were not significantly increased, which supports general literature reports that these variants often result as a well-compensated mild haemolytic anaemia and do not always result in a thalassaemic picture (Hoyer, J.D *et al*: 1999). Therefore, despite the highly unstable nature of these variants, phenotypic screening may not always highlight these properties, and therefore their potential clinical significance. Furthermore, these unstable variants tend to present at very low percentages on HPLC, for example, the Hb Köln variant was found at only 7.5%. These highly unstable variants may be easily missed or mistaken for a benign alpha chain variant.

Interestingly, the two patients in whom the Hb Köln and Hb Leiden variants were detected showed no evidence of an expected polycythaemia as their haemoglobin levels were 11.0 g/dL in the female patient with Hb Leiden and 14.4 g/dL in the male patient with Hb Köln, although this latter patient case did show evidence of haemolysis with an MCV of 113 fL.

High-affinity haemoglobin variants

There was a further subgroup of rare but clinically significant haemoglobin variants detected within our referral population, in which the underlying mutational mechanism results in haemoglobin variants with increased oxygen affinity when compared to the normal haemoglobin A molecule. They are of clinical interest as they tend to present with haemoglobin levels above the normal range and usually lead to polycythaemia (Bain, B: 2006). In reality the actual haemoglobin levels can be highly variable between patients, as encountered in this study, as the degree of shift of the oxygen dissociation curve may influence the final haemoglobin levels.

A total of three high-affinity haemoglobin variants: Hb Old Dominion, Hb Olympia and Hb Malmö were detected in 7 patient cases referred for the detection of an abnormal variant with signs of polycythaemia, as they all had Hb levels above or within the upper range of normal levels.

Heterozygosity for Hb Malmö was found in four Asian Pakistani family members. Initially the parents were referred as an antenatal couple with an unknown ‘haemoglobin J variant’ of 47.6% and 45.9%, respectively, detected on HPLC, with normal red cell indices but

with evidence of mild polycythaemia as their respective haemoglobin levels were 15.1 and 16.4 g/dL. The findings from subsequent DNA analysis highlighted the potential implication for the next generation in this family, given the possibility of their offspring inheriting homozygosity for this variant. Again, these high-affinity variants are usually reported in the heterozygous state, there is very limited data on the consequence of a homozygous genotype. Available literate suggests that, even though polycythaemia rarely requires treatment, *in vivo* these variants may be associated with fetal hypoxia (Weatherall, D.J and Clegg, J.B: 2001) due to the impaired oxygen transfer to the fetus. However, none of these findings have been reported *in vitro*. DNA analysis of the two male offspring found that they were both heterozygous for Hb Malmö, with haemoglobin levels of 15.7 and 14.6 g/dL at 7 and 8 years of age respectively.

Correlation of rare beta chain haemoglobin variants with stated ethnicity:

The highest frequency of rare beta chain haemoglobin variants was found in individuals of Asian (*n=12*) origin, followed by those of White (*n=8*) ethnicity. When the patients stated ethnicity was compared to their beta genotype, as represented in Chapter 3, Figure 3.16, it shows that a large proportion (61%, *n=59*) of patients with a haemoglobin variant genotype had a variant haemoglobin that would be expected within their ethnic group. This strong correlation can be accounted for by the large number of sickle haemoglobins reported in patients of Black ethnicity. Therefore, when the clinically significant variants are omitted, this correlation is much weaker, only 13% (*n=12*) of all patients with a haemoglobin variant genotype had a rare variant that would be expected within their ethnic group. For example, only four out of the 12 Asian patients had a rare variant that would be normally expected within this ethnic group. Furthermore, all three cases of the Hb Monroe variant were found in this ethnic group, which is normally expected in Black ethnic groups. Additionally, only half of those patients of white ethnicity had a rare variant that would normally be expected within this ethnic group.

Summary of rare *HBB* haemoglobin variants detected by DNA sequencing:

The identification of these *HBB* variants by DNA sequencing was important in some cases as there are a subset that cannot be separated by HPLC and in some cases may even mask underlying β-thalassaemia trait as they elute within the HbA_2 window. In this study they included: Hb D-Iran, Hb Spanish Town and Hb OSU-Christiansborg. Furthermore, DNA analysis was an essential tool for the differentiation of those variants that elute within the

'D-window' on HPLC, such as Hb G-Norfolk, Hb G-Philadelphia and Hb D-Punjab, the latter of which must be identified as part of the antenatal screening guidelines.

The frequency of rare beta chain haemoglobin variants found within individuals of Asian ethnicity is significant as this is an ethnic group in which consanguinity is not uncommon, giving rise to the possibility of homozygosity for these variants. Depending on the nature of these variants they may have clinical consequences, such as in the Hb Malmö scenario outlined above. Additionally, only DNA sequencing was capable of detecting the high affinity variant Hb Köln, as it elutes within the Hb A window on HPLC. Again, homozygosity for this variant may result in a severe haemolytic anaemia.

The data presented above contrasts with discussion around DNA screening guidelines that suggest the detection of abnormal haemoglobin variants by DNA sequencing creates 'noise' for screening programmes and that there is little clinical value in their detection and identification. Within our referral population, 50% of the rare beta chain variants detected have an altered protein structure and may be considered to be of clinical significance, especially in the antenatal context due to their possible interactions with β-thalassaemia. This, coupled with the poor correlation observed between patients' ethnicity and the variant found, indicates that our screening methodologies employed are of clinical benefit. The detection of these rare *HBB* variants may help to inform clinical management and monitoring of a patient, although their identification would not form an essential part of a PND screening service.

4.10.2b Alpha chain haemoglobin variants detected by DNA sequencing of *HBA1* and *HBA2*.

Direct DNA sequencing of *HBA1* and *HBA2* detected 33 different alpha chain variant haemoglobins in 16% (*n=53*) of all patient referrals (Chapter 3, Figure 3.8). Surprisingly, nearly an equal number of haemoglobin variant cases were detected within *HBA1* (*n=30*) and *HBA2* (*n=29*). This is unexpected because general literature tends to only report those variants within *HBA2* which are deemed to be of greater clinical significance. The clinical significance of each of these variants detected is discussed further in this section.

The variants found at highest frequency within our referral population were all found within *HBA1* and include Hb G-Philadelphia (*n=5*), Hb Manitoba (II) (*n=4*), Hb Q-India

(*n=4*) followed by Hb Stanleyville II (*n=3*). The remaining 29 variants were found at low frequency in a variety of cases and ethnic groups.

The relatively high frequency of Hb G-Philadelphia and Hb Stanleyville II, all of which were found in the Black or mixed Black ethnic groups within our referral population, is expected due to the overall large proportion of Black patient referrals received and the common co-inheritance of Hb G-Philadelphia in *cis* with $-\alpha^{3.7}$ or with β^{S}. Hb Manitoba II, normally reported in British and Italian populations, was found in two individuals of Mixed ethnicity (Indian and French), and unexpectedly in two individuals of Asian ethnicity. Hb Q-India was unexpectedly found in one SEA patient but was also found in three Asian individuals, which is more consistent with published reports.

Despite literature reports suggesting that a large proportion of haemoglobin variants found within the alpha globin genes are of no clinical significance, a surprisingly large proportion of alpha variants detected within our referral population were associated with some form of clinical disease. From the wide range (*n=33*) of alpha chain variants detected within this study it would be expected that only a very small proportion of these variants would be of any clinical significance, but in fact only 45% (*n=15*) of the different variants detected, found in within 40% (*n=21*) of cases were associated with a clinical abnormality.

Furthermore, even though two of these variants (Hb G-Philadelphia and Hb Q-Thailand) are reported on the Hb variant database as having normal functional protein properties and therefore a normal clinical presentation, they are most commonly found in the compound heterozygous state with an alpha thalassaemia deletion ($-\alpha^{3.7}$ and $-\alpha^{4.2}$ respectively), as found in all seven patient cases within this study. When Hb G-Philadelphia is found in *trans* with $-\alpha^{3.7}$ deletion, this interaction can increase the relative proportion of the variant, from the usual 30-35% to around 45%, and therefore reinforce any associated clinical phenotype. It was not possible to determine the mode of inheritance of this variant in the five reported cases, due to the co-inheritance of Hb S in three cases and homozygosity for $-\alpha^{3.7}$ in two cases, but all of these genetic interactions resulted in a distinct microcytosis and hypochromia.

Within this study, heterozygosity for Hb Q-Thailand was found in *trans* with the $-\alpha^{4.2}$ deletion in two individual antenatal Chinese patients, as reflected by the high quantity of the variant detected (30.8 and 30%, respectively), again resulting in mild microcytosis and hypochromia. Unfortunately there was no phenotypic information available on their respective partners.

In one further case, the alpha chain variant (*HBA2*:pLys127Glu) was found in a White Irish patient referred for an unknown variant of 18.8% detected on HPLC, however, it was not possible to assess the clinical significance of this variant due to limited case reports.

The remaining 14 clinically significant alpha chain variants detected within 38% (*n=20*) of these patients are discussed below. They are all associated with a change in protein structure and are associated with either a thalassaemic phenotype or haemolytic anaemia, caused by the resultant unstable protein structure of the variant or polycythaemia caused by the increased oxygen affinity of the variant.

Alpha chain haemoglobin variants causing a thalassaemic phenotype:

Four different variants (Hb Constant Spring, Hb Seal Rock, Hb Quong Sze and Hb Evanston) found in six patients within our referral population are classified, according to the Hb variant database, as a non-deletional form of alpha thalassaemia as heterozygotes for these variants have an α-thalassaemia phenotype with a chronic haemolytic anaemia (see Chapter 1, Section 1.4.4). This phenotype is mostly due to the alteration of the tertiary structure of the haemoglobin molecule in which the resultant highly unstable haemoglobin tetramer precipitates within the red cells.

Definitive diagnosis of this class of variants by mutation detection, in this case DNA sequencing, in our patient cohort was vital as this is the only current technique, apart from mass spectrometry, in which the presence of such variants can be reliably demonstrated. Detection of these highly unstable alpha chain haemoglobin variants by HPLC can be unreliable. Rather than forming discrete bands at around 25%, as usually found in carriers for alpha chain variants (Chapter 1, Section 1.4.4), they are rapidly degraded and may go undetected, for example heterozygotes for the Hb Constant Spring variant have ~1% of Hb Constant Spring in their red blood cells (Bain, B (2006). Of all six cases in which this

class of variant was found, phenotypic screening by HPLC failed to detect the variant haemoglobin, and only by performing DNA sequencing analysis were they detected.

Compound heterozygosity for these unstable variants with a deletional form of α^+ or α^0 thalassaemia, commonly found within the SEA population groups, tend to result in particularly severe forms of Hb H disease compared to those caused by deletional compound heterozygous genotypes (Chapter 1, Section 1.4.4) as they are associated with a dominant mode of inheritance. Homozygosity for the most severe variants may even be incompatible with life (Weatherall, D.J and Clegg, J.B: 2001). DNA studies on one SEA patient, who had been referred for possible Hb H disease, with an MCV of 74 fL and an MCH of 18.0 pg and a fast band detected on HPLC, confirmed that they did have Hb H disease caused by the compound heterozygous genotype --SEA/$\alpha^{QS}\alpha$. Unexpectedly, this patient had a less marked hypochromic and microcytic anaemia in comparison to those five patients with deletional forms of Hb H diseased in this study.

DNA analysis on another female patient, of unknown ethnicity, who had been referred with microcytic indices (MCV 77 fL and MCH 25.2 pg) and a possible HbA_2 variant of 1.8%, showed that she was compound heterozygous for Hb Seal Rock and $-\alpha^{3.7}$ deletion, resulting in a mild form of Hb H disease.

These genetic interactions are important to diagnose within the antenatal context, as the two cases discussed here highlight. Both reports of the Hb Constant Spring variant in this study had been referred as part of an 'at risk' antenatal couple. One individual with this variant had been referred as part of a Chinese antenatal couple, with possible alpha thalassaemia (MCV 78 fL and MCH 26.5 pg). DNA studies on their partner revealed heterozygosity for the α^0 --SEA deletion. Offspring of this couple therefore have a 1 in 4 risk of inheriting Hb H disease caused by the interaction of these genotypes. The second patient with Hb Constant Spring had been referred as part of a Thai antenatal couple for possible alpha thalassaemia (MCV 77.8 fL and MCH 25.7 pg), with no variant detected by HPLC. DNA studies on their partner, who had been referred with suspected Hb E with a raised HbA_2 level of 21.7% (MCV 67.1 fL and MCH 21.5 pg), revealed heterozygosity for the α^0 --SEA deletion and heterozygosity for the β^E variant. Offspring of this couple have a 1 in 4 chance of inheriting the genotype: --/$\alpha^{CS}\alpha$, $\beta\beta^E$ which is associated with a

phenotype similar to Hb H disease. The literature reports that such genetic interactions are common in Thai and SEA populations.

Hb Evanston was detected in one patient of unknown ethnicity who had been referred for query β or δβ-thalassaemia (MCV 63.2 fL and MCH 21.2 pg) with a raised HbA_2 of 4.2% and Hb F of 18.0%. Extensive DNA analysis of this patent also revealed heterozygosity for severe β^+ thalassaemia, caused by the mutation IVS1-5 (*G>C)*. Interestingly, Hb Evanston is found on the *HBA1* gene and is not only associated with a thalassaemic phenotype but also with increased oxygen affinity. It is rapidly destroyed and is not detectable on HPLC. There are limited case reports of this interaction, but unlike deletional forms of alpha thalassaemia, this alpha genotype is unlikely to ameliorate the beta genotype. Further DNA sequencing of the $^{G}\gamma$ gene showed this patient was heterozygous for the *Xmn* I polymorphism, which might account for their high Hb F level, but this effect tends to be only observed under conditions of erythropoietic stress, which the patient's raised RBC count (6.80 $x10^{12}$/L) indicates. Further genetic analysis by *HBB* MLPA was not performed as this patient was unlikely to carry a deletional type of HPFH or δβ because their Hb A was too high (14.4 g/dL) for this genotype combination. This case reflects the importance of extensive DNA analysis in clarifying the exact genotype in those patients presenting with complex phenotypic screening results.

Unstable haemoglobin variants

A further group of three different unstable alpha chain variants: Hb Setif, Hb Spanish Town and Hb Manitoba types I and II, which are classified together as they both lead to the same protein defect, were detected in seven cases from our referral population. These variants alter the tertiary structure of the haemoglobin molecule and can result in a chronic haemolytic anaemia. Unexpectedly this clinical picture was only observed in those individuals with the Hb Manitoba (Type II) variant, found within *HBA1*, in which all of these patients had borderline to reduced red cell indices.

The individual with Hb Manitoba (Type I), found on *HBA2*, presented with a higher percentage (15.7%) of variant than those with the Type II variant. This is consistent with literature reports that state variants within *HBA2* tend to present at a higher proportion (see Chapter 1, Section 1.4.4) than those within *HBA1* due to their differential transcription rates. Unexpectedly, unlike those patients with the Type I variant, this patient presented

with normal red cell indices. It may be expected that this variant type would be associated with a more pronounced clinical picture due to the higher proportion of protein production from the *HBA2* gene.

Interestingly another unstable variant which was detected in one patient by DNA sequencing and was not detected on HPLC was Hb Riccarton. This variant is unlike those unstable variants discussed above in that it belongs to a group that are so highly unstable they undergo very rapid post synthetic degradation and are unable to form haemoglobin tetramers. This means there is no associated loss of normal β chains as they remain in excess within the red cells and the clinical presentation is normal. This correlated with the phenotype of the affected White British male as the Riccarton variant was not detected on HPLC and their red cell indices were normal. DNA sequencing was important in this case as this patient had been referred as part of an 'at risk' antenatal couple in which his Chinese origin partner had a raised Hb A_2 level of 3.6%, although DNA analysis failed to detect an underlying mutation in this case.

Again, as with the unstable beta chain variants, it is important that these unstable abnormal haemoglobins are diagnosed by DNA sequencing as the resultant unstable proteins cannot be detected by HPLC screening and they may be associated with an alpha thalassaemia phenotype.

High affinity haemoglobin variants

A further five alpha chain variants with an increased oxygen affinity were detected in six cases (Hb Chesapeake, Hb Dunn, Hb Linwood, Hb Toulon and Hb West-One), in a variety of ethnic groups. These variants are associated with erythrocytosis, interestingly, two of which (Hb Toulon and Hb Chesapeake) were found within *HBA1*.

Hb Chesapeake, was detected in one patient of unknown ethnicity who had been referred for DNA studies for the identification of an unknown haemoglobin variant of 22.5%. This variant is also reported to be mildly unstable, as well as being associated with mild erythrocytosis; this was evident within this patients phenotype as she had normal red cell indices and a RBC of 4.95 ($x10^{12}$/L).

The two cases of Hb Linwood were found in a SEA family, in which the antenatal couple had been referred for genetic analysis from the phenotypic screening results obtained. One partner with Hb Linwood was referred for an unknown variant detected at 27.3%. Their partner was referred for possible alpha thalassaemia (MCV 59 fL and MCH 17.8 pg). DNA analysis revealed heterozygosity for the α^0 --SEA deletion. There was limited data on the possible clinical phenotype associated with this potential interaction. Subsequent DNA analysis of their child showed that he had inherited heterozygosity for Hb Linwood. This variant is reported to have an increased oxygen affinity and is associated with reticulocytosis, therefore offspring of this couple are likely to have marked erythrocytosis. Both individuals in this family with the Hb Linwood variant had Hb levels at the high end of normal range (14.5% and 14.8% respectively).

Summary of rare haemoglobin variants detected by DNA sequencing of *HBA1* and *HBA2*:

This study detected an unexpectedly high proportion of alpha chain variants within our referral population, a surprisingly large number of which were not completely clinically benign as they result in an altered protein structure. This is contradictory to literature reports that suggests the majority of alpha chain variants do not result in a clinical disorder and therefore do not warrant identification by DNA studies in the antenatal context.

The pathophysiology of some of the variants identified in this study was similar to that found in the thalassaemia disorders, and because they are frequently co-inherited together with different forms of thalassaemia, DNA sequencing provided a vital tool for their identification, especially with those that cannot be detected by phenotypic screening alone due their highly unstable nature. Their identification proved especially valuable in those patients who presented with a microcytic and hypochromic phenotype where Gap-PCR had failed to detect an alpha deletional mutation mechanism.

Even though in some DNA diagnostic laboratories it is not always possible to identify alpha chain variants due to methodology restraints, and the antenatal screening programme states that it is not always of potential clinical significance to do so, the cases discussed above highlight the wide range of clinical conditions associated with these variants, including those which may result in particularly severe forms of Hb H disease. There

remain some variants which are important to identify by DNA sequencing in the antenatal context, and potentially even for PND purposes in isolated cases.

4.11 Summary of mutation screening by DNA sequencing of *HBA1* and *HBA2* for the detection of all point mutations associated with alpha thalassaemia and haemoglobin variants.

Routine DNA sequencing of *HBA1* and *HBA2* provided an important diagnostic screening tool, as a wide range of non-deletional alpha thalassaemia mutations and haemoglobin variants were detected within a variety of ethnic groups from the North West of England referral population, the clinical significance, or otherwise, of which has been discussed above. Out of all of these patient cases, in which a point mutation was found within either of the alpha globin genes, only 53% (*n=31*) had an alpha genotype that would be expected in their stated ethnic group and 37% (*n=22*) had an alpha genotype not normally associated with their ethnic group. In the remaining 10% (*n=6*) of cases the patients' ethnicity was not stated (See Chapter 3, Figure 3.9). Therefore DNA sequencing of the essential regions of the alpha-globin genes indicates that a "first pass" ethnic led mutation screening strategy may miss up to 47% of alpha genotypes analysed.

A wide range of *HBA* variants were detected within this study by DNA sequencing, not all of which result in a clinically significant disease. It is widely reported in literature that a proportion of these variants, especially when co-inherited with a deletional form of α-thalassaemia result in severe forms of Hb H hydrops fetalis, a publication by Viprakasit, V *et al* (2001) highlights a case of Hb H hydrops fetalis syndrome caused by the interaction of two common α-thalassaemia genotypes $--^{MED}/\alpha^{TSaudi}\alpha$. *HBA* variants found within this study that could potentially result in a similar severe phenotype include Hb Constant Spring, Hb Quong Sze, Hb Seal Rock and Poly A (del AA). This contrasts with the antenatal screening programme, which states that these variants do not warrant further genetic analysis within the antenatal or PND context because they do not cause severe disease in the fetus. Furthermore, many of these variants are silent on HPLC due to their highly unstable nature and can therefore only be identified, in the heterozygous state, by performing DNA analysis. Therefore broader mutation detection by DNA sequencing of *HBA* could be used to identify such genetic interactions and enable counselling for potentially severe forms of Hb H disease.

4.12 MLPA analysis for the detection and confirmation of deletional forms of alpha and beta thalassaemia.

Once established and validated, MLPA proved to be a useful assay for the detection of rare or previously uncharacterised deletions and duplications affecting either the alpha or beta globin gene clusters (Chapter 2, Section 6.0). In this study MLPA was applied to those individuals in whom alternative methodologies had failed to detect a causative mutation.

Initial intensive optimisation of both the *HBA* (P140-B2) and *HBB* (P102-B1) MLPA kits was required. This assay was validated using known positive and negative control samples in which the sample had either been genotyped within the MDC using alternative methodologies or externally by MLPA analysis. Once these assays had been validated strict quality control criteria relating to data obtained from analysed samples was always adhered to (Chapter 2, Section 6.4), in which samples with a standard deviation score exceeding 0.1 were rejected. MLPA was also used in this study to confirm and validate those HPFH and δβ-thalassaemia samples that were genotyped using the Craig, R.A *et al* (1994) Gap-PCR protocol. Furthermore, when the interpretation of the expected probe heights was unavailable for certain deletions from MRC Holland, the characteristic reduced probe height pattern observed in known positive controls was recorded to aid future deletion characterisation, such as Hb Lepore and Chinese δβ-thalassaemia.

4.12a *HBA* MLPA

HBA MLPA analysis was mainly used for the confirmation of a normal alpha genotype and to exclude rarer forms of deletional α-thalassaemia in those patients in which alternative methodologies did not found a causative mutation and their phenotype did not correlate. This approach was very useful in order to comprehensively genotype two 'at risk' antenatal couple cases, discussed below, and to accurately predict the risk to their offspring for genetic counselling purposes.

The first couple, of Filipino origin, had both been referred for query alpha thalassaemia. Gap-PCR analysis on the partner revealed heterozygosity for the South East Asian α^0-thalassaemia deletion, which was consistent with his phenotype (MCV of 66.3 fL and MCH of 21.0 pg). In contrast, the normal results obtained by extensive DNA analysis of

the female was insufficient to explain her phenotype (MCV of 72.2 fL and MCH of 23.1 pg). Further *HBA* MLPA analysis revealed a normal probe pattern, therefore excluding a rare form of deletional alpha thalassaemia. Given the preliminary nature of our MLPA assay development, confirmatory analysis by the National Reference Laboratory for Haemoglobinopathies was requested. This confirmed our findings and it was concluded that iron deficiency was the underlying cause of her microcytosis and hypochromia. This couple were therefore not at risk of having a child affected with Hb Bart's Hydrops Fetalis syndrome.

The second couple, both of Chinese origin, had been referred as the partner had query alpha thalassaemia with an MCV of 71 fL and MCH of 22.4 pg, in which subsequent Gap-PCR analysis revealed heterozygosity for the South East Asian α^0-thalassaemia deletion. The female was referred for query β-thalassaemia, with a HbA_2 of 5.5%, an MCV of 64 fL and a MCH of 20.1pg. It was therefore of clinical importance to exclude underlying alpha thalassaemia in this patient, which *HBA* MLPA helped to do. DNA sequencing of her β-globin genes confirmed a diagnosis of severe β^+ thalassaemia trait, caused by the IVSII-654 (*C>T*) mutation.

HBA MLPA analysis on selected patients within our referral population did not detect any additional deletional mechanisms that had not already been identified by existing methodologies such as Gap-PCR, including any rare or novel deletions. This reflects the effectiveness of a multiplex Gap-PCR protocol for the detection of the common deletional causes of alpha thalassaemia. *HBA* MLPA analysis performed on a young Caucasian male patient subsequent to the completion of this study, detected a deletion within the regulatory region of his α-globin gene cluster (HS-40) which would result in a marked decrease in the transcription rate of the α-globin genes on that chromosome. These findings correlated with the patient's microcytic, hypochromic phenotype (MCV: 53 fL, MCH: 18.2 pg) as his triplicated alpha genotype alone, did not (as discussed in Section 4.6 above and Chapter 3, Section: 9.5).

4.12b *HBB* MLPA

Even though deletional forms of β-thalassaemia are reported to comprise a relatively rare group of disorders, as found in this study as only two cases were reported, these types of

beta thalassaemia account for nearly 20% of all β-thalassaemia alleles within the Punjab region of India (Mikula, M *et al*: 2011), with the 619 bp deletion reported to be the second most common cause of β-thalassaemia in Gujarat, India (Colah *et al*: 2010) (Chapter 1, Section 1.4.2). Their identification in areas that have a diverse immigrant population such as the North West of England is important, as unlike the common point mutations which can be associated with either a β^0 or β^+ thalassaemia phenotype, large deletional forms always result in β^0 thalassaemia, therefore their identification is important for those antenatal couples considering prenatal diagnosis. In addition, fine scale mutation analysis by DNA sequencing is likely to miss beta gene deletions as the presence of the intact copy will mask their existence on sequence traces.

Within our referral population only two deletional forms of beta thalassaemia were found in two cases. Even though these deletions were detected by the National Reference Laboratory, due to the unavailability of *HBB* MLPA within the MDC at the time (Chapter 3, Section 9.6), these cases were easily highlighted by their characteristic phenotype, as deletional forms of β-thalassaemia tend to show significantly higher levels of haemoglobin HbA_2 and occasionally F. Therefore, external testing was requested in these cases when the severity of the phenotype was discrepant from the genotype identified with the MDC.

One case of the Afghan 909 bp β° deletion (which removes the 5' end of *HBB*) was found in a male patient of unknown ethnicity who had been referred for β-thalassaemia with a very high proportion of HbA_2 (7.4%), Hb F (9.5%), an MCV of 71 fl and an MCH of 21.5 pg. This phenotypic presentation is expected, as these 5' *HBB* deletions tend to show a characteristic increase in haemoglobin F as the 5' regulatory region of *HBB* is deleted.

The one case of the 619 bp β° deletion (which removes the 3' part of *HBB*, including part of intron-2, exon 3 and a segment of DNA 3' to the termination codon) was found in a female patient of Asian Indian origin, which is expected as it is normally found in Indian or Asian populations. She had been referred as part of an antenatal couple for possible β-thalassaemia with a significantly raised HbA_2 level of 5.3%, MCV of 64 fl and an MCH of 20.3 pg. Her Hb F level of 1.7% was not significantly raised as this level could be explained by pregnancy. This patient's phenotype was consistent with case reports of this deletion, in which 3' deletions affecting *HBB* do not show an increase in haemoglobin F because the 5' regulatory region is not deleted. This couple were not at risk of having an

affected child with a clinically significant haemoglobinopathy as the partner only had an alpha plus ($-\alpha^{3.7}$) deletion.

A further three cases of deletional forms of thalassaemia found within our referral population associated with the δβ-thalassaemia disorders, which remove both the δ- and β-genes. Again these deletional mechanisms were initially detected by *HBB* MLPA performed by Oxford (Chapter 3, Section 11.6), but were highlighted within the MDC due to their characteristic raised haemoglobin F levels.

The Indian 32 kb, $^{G}\gamma^{A}\gamma(\delta\beta)^{0}$ deletion was found in an antenatal female patient of Asian Indian ethnicity, referred for query δβ-thalassaemia or HPFH, with a raised Hb F level of 19.2% and a normal Hb A_2 of 2.3%, she also had microcytic and hypochromic indices with an MCV of 77 fl and an MCH of 25.7 pg. This patient's phenotype was consistent with other reports concerning this mutation and subsequent screening of her partner revealed a normal phenotype.

The two cases of the Black 12 kb $\delta\beta^{\circ}$ deletion were found in two female siblings referred for possible δβ-thalassaemia (as in their mother). Both had marked microcytic and hypochromic indices and elevated haemoglobin F levels of 15.2% and 15.8% respectively, and haemoglobin A_2 levels within the normal range. Their phenotype was consistent with case reports of this deletion with the co-inheritance of α-thalassaemia.

A further 6 cases of deletional forms of thalassaemia were detected by *HBB* MLPA analysis performed at Oxford, including rare, uncharacterised HPFH and δβ-thalassaemia deletions. This demonstrates that MLPA can be a useful tool for the detection of deletions in the α or β-globin clusters genes with no prior knowledge of breakpoints of the deletions, providing a useful DNA screening tool in appropriate cases. Once *HBB* MLPA had been established and validated within the MDC laboratory, it was useful in those cases where DNA sequencing was uninformative. MLPA analysis helped to differentiate those patients with the Hb SS genotype from Hb S/HPFH or Hb S/δβ-thalassaemia. This was especially helpful in patients with the Arab-Indian sickle haplotype (Chapter 1, Section 1.4.4) or those with only homozygous polymorphisms detected within the β-globin gene by DNA sequencing, possibly indicating the presence of a whole or partial deletion of one allele.

HBB MLPA was also used within the MDC for the analysis of two suspected cases of εγδβ thalassaemia (Chapter 3, Section 11.5), in which initial extensive genetic screening by existing methodologies had failed to detect a causative mutation. *HBB* MLPA analysis helped to exclude εγδβ-thalassaemia in a young Caucasian male patient case, in whom only the triplicated alpha genotype had been found (but as highlighted in Section 4.12a, subsequent *HBA* MLPA analysis detected an additional deletional mechanism). The second patient, who was a white British female, showed a characteristic reduced probe pattern (See Chapter 3, Section 11.5 and Figure 3.19) which was indicative of group I εγδβ-thalassaemia (which removes the entire β-globin cluster), therefore explaining this patient's clinical presentation of severe neonatal anaemia.

Even though the εγδβ-thalassaemia disorders are rare, their detection and characterisation by *HBB* MLPA is useful as it can be used as an elimination tool for the proportion of patients who present with unexplained hypochromia and microcytosis and normal Hb A_2 levels in which an α-thalassaemia genotype was not found. It would also be impractical to detect this group of disorders by Gap-PCR methodologies since their molecular heterogeneity is broad, as they can be caused by at least 15 known deletional mechanisms (Rooks, H *et al*: 2005).

Furthermore, MLPA is a useful tool for the diagnosis of those thalassaemic conditions which may be caused by deletions within the upstream regulatory elements of either the α-globin gene cluster (HS-40) or the β-globin gene cluster (β-LCR), as other methodologies may indicate that the globin genes are intact, but would fail to indicate as to whether or not they are actively transcribing and synthesising globin chains on that chromosome (as demonstrated from the male Caucasian patient case discussed in Section 4.12a with the HS-40 deletion and the white British female discussed earlier in this Section with εγδβ-thalassaemia).

4.12.1 Summary of *HBA* and *HBB* MLPA analysis

MLPA is a useful screening tool for the detection of deletions within *HBA* and *HBB*, although it does not allow the user to map the exact boundaries of a deletion. Identifying a deletional mechanism by MLPA analysis is important in the antenatal context for the accurate genetic counselling of couples, especially for those conditions in which prenatal

diagnosis may be indicated. It is important that a haemoglobinopathy DNA diagnostic service employs such gene dosage based methodologies that can detect changes in the relative quantity of copy numbers, as alternative fine scale assay methodologies such as DNA sequencing cannot always detect those complex genetic abnormalities associated with the interaction of α and β-thalassaemia.

4.13 Patients in which no causative mutation was detected by the DNA screening methodologies employed within MDC.

In a proportion of patient referrals (13%, *n=41*), no causative mutation was detected using the laboratory screening methods outlined in Chapter 2, or the genotype detected within the MDC was insufficient on its own to account for the patients' phenotype. (Chapter 3, Section 11.0). As already outlined (Chapter 3, Figure 3.18), the largest proportion (34%, *n=14*) of these patients were those who had been referred as part of an antenatal couple in which extensive DNA testing on both partners was beneficial in accurately assessing the risk to the offspring, especially in those presenting with a borderline HbA_2 level in which the other partner had β-thalassaemia.

DNA analysis for the exclusion of a thalassaemia disorder was requested in a proportion of patients and young children where it was deemed to be of clinical interest i.e. in those where it was not possible to exclude underlying iron deficiency, especially in those not responding to iron therapy, or those in which ZPP levels were unreliable (as they can often be falsely elevated in some cases of thalassaemia), or in order to avoid further invasive clinical testing.

This group of patients also included those urgent antenatal cases in which one partner had borderline red cell indices and there was insufficient time to trial a course of iron therapy. Of this group of patients 12% (*n=5*) were diagnosed with iron deficiency either from the exclusion diagnosis by DNA or by their low serum ferritin levels. The two other cases, with suspected εγδβ-thalassaemia, which comprised 5% these patient cases, have already been discussed above (Section 4.12b). This deletional mechanism was not initially detected due to the unavailability of *HBB* MLPA at that time.

The remaining 11 patient cases have been previously highlighted (Chapter 3, Section 11.0) and included those with other underlying clinical conditions, a novel mutation, or their phenotype remained unexplained despite confirmatory genetic testing by the National Reference Laboratory for Haemoglobinopathies.

4.14 Patients referred to the National Reference Laboratory for Haemoglobinopathies for further genetic testing - indicating further service development within the MDC.

Within this study there were a total of 12 cases in which no mutation was detected using the panel of tests employed within the MDC at the time, where subsequent DNA analysis by the National Reference Laboratory for Haemoglobinopathies detected a causative mutational mechanism (Chapter 3, Section 11.6).

Of these 12 patient cases, five consisted of four different deletional genotypes, which have already been discussed above in Section 4.12b, and included deletional β-thalassaemia or novel δβ-deletional mechanisms that could not be detected or confidently characterised within the MDC due to the unavailability of *HBB* MLPA at that time. Our current testing regime would now be able to detect these deletional mechanisms since the establishment of *HBB* MLPA. A further five cases included rare or unidentified forms of HPFH or δβ-thalassaemia which were also detected by *HBB* MLPA (highlighted in Chapter 3, Section 11.6). Again, these genotypes can now be detected using our current testing regime. In one further case, in whom the National Reference Laboratory reported to be either normal or to have some form of $\varepsilon\gamma\delta\beta^0$ thalassaemia could be diagnosed in the same manner as those cases discussed above in Section 4.12b.

In the remaining one patient case, who was reported by the National Reference Laboratory as having a probable δ-chain variant, further DNA analysis or assay development would not be worthwhile, as δ-chain variants are not clinically significant.

4.14.1 Future service developments

As well as MLPA, which is relatively cumbersome and does not map deletions with a high degree of accuracy, detection of the deletional forms of β-thalassaemia can be achieved by using a Gap-PCR based methodology (provided that the deletional breakpoints are known), which could then be used alongside current methodologies to act as an independent confirmatory technique. A future service development for the MDC diagnostic service would be the development of a Gap-PCR method for the detection of the following β^0 thalassaemia deletions which all occur at significant frequencies to warrant their identification: Indian 619 bp, Black 1393 bp, Afghan 909 bp, Filipino ~45 kb and the Indian 32 kb deletion.

In order to provide a pre natal diagnostic service, where speed of diagnosis is very important, it would be prudent to consider the development of a panel of rapid assays which could be used as independent confirmatory techniques. An obvious candidate for this would be rapid testing for the Hb S mutation and a panel of common beta zero mutations found in our local referral population. Techniques such as ARMS-PCR may be suitable for this purpose.

Furthermore, the development of mass spectrometry would be useful for the positive identification of haemoglobin variants. A large proportion of variants can be identified by this technique as mass spectrometry can determine if the variant is on either the α- or β-globin chain and can also estimate the proportion of the variant. Further analysis may also enable prediction of the possible amino acid substitution based on the observed change in mass of the variant. This technique may eliminate the need for DNA sequencing in those cases which cannot be positively identified by HPLC and would also be useful for the detection of fusion gene variants such as Hb Kenya. The disadvantage of this technique is the initial cost, as the apparatus is very expensive and the data interpretation of the results requires considerable experience. Whilst it would be useful in an antenatal context, it would not, in the conventional approach, replace genetic PND investigation.

4.15 Summary and conclusions of mutation detection strategy employed versus ethnic led mutation screening for the North West referral population

Within most populations, a large proportion of carriers for alpha and/or beta thalassaemia will have a common subset of mutations which are strongly associated with their ethnic group, followed by a smaller proportion exhibiting rarer mutations. Therefore, the strategy for identifying α- and β-thalassaemia mutations within many diagnostic laboratories is based on this knowledge of the spectrum of expected common, and less common but still relatively frequent, mutations expected to be found within the ethnic group of the individual. This 'ethnic-led' PCR based screening approach should therefore detect a large proportion of mutations, eliminating the need to perform DNA sequence analysis on every case.

Recently, a number of literature reports have highlighted the increasing racial heterogeneity of the UK population, in which the spectrum of haemoglobinopathy mutations found has significantly increased (see Chapter 1, Section 1.3). This large

molecular heterogeneity poses challenges for diagnostic laboratories which must be able identify this wide range of mutations which will be present within ethnically diverse populations, as well as rare mutations which can present in any population. The referral population served by the Molecular Diagnostics Centre at Manchester Royal Infirmary represents a diverse ethnic mix. Therefore, in order to achieve a high mutation detection rate in such a heterogeneous population, a comprehensive mutation screening approach is required. This diagnostic screening approach was employed within the MDC by using multiplex Gap-PCR, DNA sequencing and MLPA methodologies in order to detect the wide range of mutations likely to be present within the North-West of England.

Results obtained from DNA sequencing of *HBB* identified 44 different types of non-deletional mutations, which comprised of 20 that cause β-thalassaemia and 24 that are associated with variant haemoglobins. Furthermore, DNA sequencing of *HBA1* and *HBA2* also identified 38 different types of non-deletional mutations, which comprised a small proportion of five that result in a non-deletional form of α-thalassaemia and 33 that are associated with variant haemoglobins. The variety of alpha and beta chain haemoglobin variants detected in this study is an important observation since our referral population is comprised mainly of individuals of Asian ethnicity, an ethnic group in which consanguinity is not uncommon. Subsequently this may serve to increase the frequency at which homozygosity for certain haemoglobin variants occurs, even though these homozygous genotypes are considered to be rare in the general population. These may have potential implications for any future offspring in which some circumstances may result in clinically significant disorders, such as the Hb Malmö case discussed in Section 4.10.2a.

DNA sequencing proved a useful assay for the detection of a proportion of haemoglobin variants that could not be detected or positively identified by HPLC analysis, including those *HBB* variants that tend to co-elute within the Hb A or Hb A_2 windows or those *HBA* variants that are so unstable are undetectable at the protein level.

A broad spectrum of mutations were detected within our referral population by DNA sequencing, resulting in a wide range of genetic interactions between different alpha and beta thalassaemia associated alleles, reflecting the complex molecular heterogeneity of these disorders.

A compound heterozygous genotype (i.e. a mutation in both the α and β-globin chains) was found in 17% (*n=55*) of all patient diagnoses made, and included cases where the α:β globin chain ratio was imbalanced by their co-inheritance, including cases in which β-thalassaemia was aggravated by the co-inheritance of triplicated alpha thalassaemia. In contrast, there were cases in which the individual's phenotype was ameliorated by the co-inheritance of deletional alpha thalassaemia, particularly α^0 thalassaemia.

Such gene-gene interactions between different alleles often made it difficult to accurately interpret patients' clinical phenotype, such as in those cases with β-thalassaemia that had been masked by co-inheritance with α^0 thalassaemia, resulting in the normalisation of their red cell indices. DNA sequencing in these cases helped to clarify their genotypes, which was especially valuable in those antenatal couples where both partners had thalassaemic red-cell changes but with normal HbA_2 levels. This was important in order to assess their risk of having a child affected with Hb Bart's hydrops syndrome.

The frequency and distribution of the haemoglobinopathy mutations detected within our referral population, by employing a broad mutation detection approach, was then correlated with patients' stated ethnicity using previously published mutation frequencies. The strength of this correlation varied between the different genotypes: there was a strong correlation observed between patients' stated ethnicity and those with a deletional form of alpha thalassaemia (detected by Gap-PCR) in which a large proportion (82%, *n=108*) of patients had a deletion which would be expected within their ethnic group. This correlation was less pronounced in those patients that had a mutation within *HBA1* or *HBA2* (including variant haemoglobins) detected by DNA sequencing, as 53% (*n=31*) of patients in this group had an alpha genotype that would be expected in their ethnic group.
There was a stronger correlation in those patients that had a non-deletional mutation within *HBB* (including variant haemoglobins) detected by DNA sequencing, in which 66% (*n=111*) of patients had a beta genotype that would be expected within their ethnic group. When this group of patients were analysed further, the correlation between patients' ethnicity and beta genotype found was less pronounced in those with a haemoglobin variant genotype than those with a β-thalassaemia genotype, in which 61% (*n=59*) and 73% (*n=46*) respectively had a mutation which would normally be expected within their ethnic group.

In conclusion, the results obtained from DNA sequencing of the essential regions of *HBA1* and *HBA2* indicates that a "first pass" ethnic led targeted mutation screening strategy may miss up to 47% of alpha point mutation genotype cases analysed, whilst a similar ethnic led strategy for *HBB* may miss up to 34% (*n=57*) of all beta genotype cases analysed. This, together with the large spectrum of alpha point mutations detected and the significant number of complex alpha and beta genotype interactions found, indicates that a comprehensive mutational screening approach, using multiplex Gap-PCR, DNA sequencing and MLPA is appropriate to detect the mutation spectrum found within the heterogeneous populations of the North-West of England, or in any other ethnically diverse populations.

References

Aguilar-Martinez, P. and Gulbis, B.(2010). Molecular Genetics of the α-thalassaemias. *Encyclopaedia of Life Sciences*. John Wiley & Sons Ltd, Chichester.

Akinsheye, I., Alsultan, A., Solovieff, N., Ngo, D., Baldwin, C.T., Sebastiani, P., Chui, D.H.K. and Steinberg, M.H. (2011). Fetal haemoglobin in sickle cell anaemia. *Blood*, **118,** 19-27.

Arnold, S., Tan, C.,Thuan, C., Quah, P.S., Poh, S.L. and Chong, S. (2001). A rapid and reliable 7- deletion multiplex polymerase chain reaction assay for α-thalassaemia. *Blood*, **98**, 250-251.

Bain, B. (2006). Haemoglobinopathy Diagnosis. Second Edition. Blackwell Publishing. Oxford.

Bain, B. (2011). Haemoglobinopathy diagnosis: Algorithms, lessons and pitfalls. *Blood Reviews*, **25**, 205-213.

Baysal, E. and Huisman, T.H.J. (1994). Detection of common deletional α-thalassaemia-2 determinants by PCR. *American Journal of Haematology*, **46**, 208-213.

Bowden, D.K., Vickers, M.A. and Higgs, D.R. (1992). A PCR-based strategy to detect the common severe determinants of α-thalassaemia. *British Journal of Haematology*, **81**, 104-108.

Chong, S., Boehm, C.D., Higgs, D.R. and Cutting, G.R. (2000). Single-tube multiplex-PCR screen for common deletional determinants of α-thalassaemia. *Blood*, **95,** 360-362.

Colah, R., Gorakshakar, A., Phanasgaonkar, S., D'Souza, E., Nadkarni, A., Surve, R., Sawant, P., Master, D., Patel, R., Ghosh, K. and Mohanty, D. (2010). Epidemiology of β-thalassaemia in Western India: mapping the frequencies and mutations in sub-regions of Maharashtra and Gujarat. *British Journal of Haematology*, **149**, 739-747.

Collins, F, S., Stoeckert, C.J., Serjeant, G.R., Forget, B.G. and Weissman, S,M. (1984). $^{G}\gamma\beta^{+}$ Hereditary persistence of fetal haemoglobin: Cosmid cloning and identification of a specific mutation 5' to the $^{G}\gamma$ gene. *Proceedings of the National Academy of Sciences of the USA*, **81**, 4894-4898.

Craig, J.E., Barnetson, R.A., Prior, J., Raven, J.L. and Thein, S.L. (1994). Rapid detection of deletions causing δβ thalassaemia and hereditary persistence of fetal haemoglobin by enzymatic amplification. *Blood*, **83**, 1673-1682.

Craig, J.E., Rochette,J., Fisher, C.A., Weatherall, D.J., Marc, S., Lathrop, G.M., Demenais, F. and Thein, S. (1996). Dissecting the loci controlling fetal haemoglobin production on chromosomes 11p and 6q by the regressive approach. *Nature Genetics*, **12**, 58-64.

Dode, C., Krishnamoorthy, R., Lamb, J. and Rochette, J. (1992). Rapid analysis of $-\alpha^{3.7}$ thalassaemia and $\alpha\alpha\alpha^{anti3.7}$ triplication by enzymatic amplification analysis. *British Journal of Haematology*, **82**, 105-111.

Donze, D., Townes, T.M. and Bieker, J.J. (1995). Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. *Journal of Biological Chemistry,* **270,** 1955-1959.

Driscoll, C.M.,Dobkin, C.S. and Alter, B.P. (1989). γδβ-Thalassaemia due to a *de novo* mutation deleting the 5' β-globin gene activation-region hypersensitive sites. *Proceedings of the National Academy of Sciences of the USA*, **86**, 7470-7474.

Embury,S.H., Miller, J.A., Dozy, A.M., Kan,Y.W., Chan, V. and Todd, D. (1980). Two different molecular organizations account for the single alpha- globin gene of the alpha-thalassemia-2 genotype. *Journal of Clinical Investigation*, **66**, 1319-25.

Eng, B., Greenlay, B. and Waye, J.S. (2009). Characterisation of the British a^0-thalassaemia deletion: evidence of a founder effect in Newfoundland, Canada. *British Journal of Haematology,* **147,** 150–156.

Eng, L.I., Baer, A., Lewis, A.N. and Welch, Q.B. (1973). Haemoglobin Constant Spring (Slow-moving Hemoglobin X Components) and Haemoglobin E in Malayan Aborigines. *American Journal of Human Genetics*, **25**, 382-387.

Furuya, C.,Yamashiro,Y., Hattori, Y., Hino, M., Nishioka,H., Shimizu,Y., Okano, K. and Horibe, K. (2008). A novel εγδβ thalassaemia of 1.4Mb deletion found in a Japanese patient. *American Journal of Haematology*, **83**, 84-86.

Galarneau, G., Palmer, C.D., Sankaran, V.G., Orkin, S.H., Hirschhorn, J.N. and Lettre, G. (2010). Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. *Nature Genetics,* **42,** 1049–1051.

Harteveld, C. and Higgs, D.R. (2010). α-thalassaemia. *Orphanet Journal of Rare Diseases*, **5**, 13.

Harteveld, C.L., Voskamp, A., Phylipsen, M., Akkermans, N., den Dunnen, J.T., White, S. and Giordano, P. (2005). Nine unknown rearrangements in 16p13.3 and 11p15.4 causing alpha- and beta-thalassaemia characterised by high resolution multiplex ligation-dependent probe amplification. *Journal of Medical Genetics*, **42**, 922-931.

Harteveld, C.L., Losekoot, M., Haak, H., Heister, G.A., Giordano, P.C. and Bernini, L.F. (1994). A novel polyadenylation signal mutation in the alpha 2-globin gene causing alpha thalassaemia. British Journal of Haematology, **87**, 139-143.

Higgs, D.R., Ayyub, H., Clegg, J.B., Hill, A.V.S., Nicholls, R.D., Teal, H., Wainscoat, J.S. and Weatherall, D.J. (1985). α-thalassaemia in British People. *British Medical Journal*, **290**, 1303–1306.

Higgs, D.R. and Bowden, D.K. (2001) Clinical and laboratory features of the α-thalassaemia syndromes. In: Disorders of Haemoglobin: Genetics, Pathophysiology, and Clinical Management. Cambridge University Press, Cambridge.

Ho, P.J., Hall, G.W., Luo, L.Y., Weatherall, D.J. and Thein, S.L. (1998). Beta-thalassaemia intermedia: is it possible consistently to predict phenotype from genotype? *British Journal of Haematology*, **100**, 70-78.

Hoyer, J.D., Wick, M.J., Snow, K., Finkelstein,W. and Fairbanks, V.F. (1999). DNA sequence analysis of Hb Leiden (β6[A3] or β7[A4]Glu>0). *Hemoglobin*, **23**, 299-301.

Huisman, T.H.J., Miller, A. and Schroeder, W.A. (1975). A $^{G}\gamma$ type of the Hereditary Persistence of Fetal Haemoglobin with β-chain production *in cis*. *American Journal of Human Genetics*, **27**, 765-777.

Kan.Y.W., Todd, D. and Dozy, A.M. (1974). Haemoglobin Constant Spring synthesis in red cell precursors. *British Journal of Haematology*, **28**, 103-107.

Lie-Injo, L.E., Herrera, A.R. and Kahn, Y.W. (1981). Two types of triplicated *a*-globin loci in humans. *Nucleic Acids Research*, **9**, 3707-3717.

Lie-Injo, L.E., Ganesan, J., Randhawa, Z.I., Peterson, D. and Kane, J.P. (1977). Hb Leiden-β^0 thalassaemia in a Chinese patient with severe haemolytic anaemia. *American Journal of Haematology*, **2**, 335-342.

Perutz, M. F. (1976). Structure and mechanism of haemoglobin. *British Journal Bulletin*, **32**, 195-208.

Rieder, R.F. and James, G.W. (1976). Translation of human globin mRNA: globin synthesis in cells containing Hb Leiden. *Blood*, **47**, 489-494.

Rooks, H., Bergounioux, J., Game, L., Close, J.P., Osborne, C., Best, S., Senior, T., Height, S., Thompson, R., Hadzic, N., Fraser, P., Bolton-Maggs, P. and Thein, S.L. (2005). Heterogeneity of the εγδβ thalassaemias: characterisation of three novel English deletions. *British Journal of Haematology*, **128**, 722-729.

Ryan, K., Bain, B.J., Worthington, D., James, J., Plews, D., Mason, A., Roper, S., Rees, D.C., De la Salle, B. and Streetly, A. (2010). Significant haemoglobinopathies: guidelines for screening and diagnosis. *British Journal of Haematology*, **149**, 35-49.

Sampietro, M., Thein, S.L., Contreras, M. and Pazmany, L. (1992). Variation of HbF and F-cell number with the G-gamma Xmn I (C-T) polymorphism in normal individuals. *Blood*, **79**, 832-833.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Sciences USA*, **74**, 5463-5467.

Smith, L.M., Sanders, J.Z. and Kaiser, R.J. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature*, **321**, 674-679.

Steinberg, M.H., Forget, B.G., Higgs, D.R. and Weatherall, D.J. (2009). Disorders of Haemoglobin: Genetics, Pathophysiology, and Clinical Management. Second Edition. Cambridge University Press. New York.

Thein, S.L., Menzel, S., Peng, X., Best, S., Jiang, J., Close, J., Silver, N., Gerovasilli, A., Ping, C., Yamaguchi, M., Wahlberg, K., Ulug, P., Spector, T.D., Garner, C., Matsuda, F., Farrall, M. and Lathrop, M. (2007). Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal haemoglobin levels in adults. *Proceedings of the National Academy of Sciences of the USA*, **104**, 11346-11351.

Thein, S.L. and Menzel, S. (2009). Discovering the genetics underlying foetal haemoglobin production in adults. *British Journal of Haematology*, **145,** 455-467.

Thomas, J.J., Kutlar, A., Scott, D.F. and Lanclos, K.D. (1998). Inhibition of Gene Expression by the Gγ 5' Flanking Region of the Bantu β^s Chromosome. *American Journal of Haematology*, **59**, 51–56.

Trent, R.J., Higgs, D.R., Clegg, J.B. and Weatherall, D.J. (1981). A new triplicated α-globin gene rearrangement in man. *British Journal of Haematology*, **49**, 149-152.

Tubsuwan, A., Thongperm, M., Jearawiriyapaisarn, N., Winichagoon, P., Fucharoen, S. and Svasti, S. (2011). Molecular analysis of globin gene expression in different thalassaemia disorders: individual variation of β^E pre-mRNA splicing determine disease severity. *British Journal of Haematology*, **154**, 635-643.

Viprakasit, V., Green, S., Height, S., Ayyub, H. and Higgs, D.R. (2002). Hb H hydrops fetalis syndrome associated with the interaction of two common determinants of α-thalassaemia ($--^{MED}/\alpha^{TSaudi}\alpha$). *British Journal of Haematology*, **117**, 759-762.

Weatherall, D.J. and Clegg, J.B. (2001). The Thalassaemia Syndromes. Fourth Edition. Blackwell Science publishing. Oxford

Weatherall, D.J. (2001). Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. *Nature Reviews Genetics* **2**, 245-255.

Weatherall, D.J. (2011). The Challenge of haemoglobinopathies in resource-poor countries. *British Journal of Haematology*, **154**, 736-744.

Wilber, A., Nienhuis, A.W. and Persons, D.A. (2011). Transcriptional regulation of fetal to adult haemoglobin switching: new therapeutic opportunities. *Blood*, **117**, 3945-3953.

WEBSITES:

Database of Human Haemoglobin Variants and Thalassaemias: http://globin.bx.psu.edu/hbvar

Ensemble database: http://www.ensembl.org/index.html

Human Genome Variation Society: http://www.hgvs.org/

Location of the alpha and beta gene clusters: http://www.ncbi.nlm.nih.gov/books/NBK10023/

MLPA excel interpretation macros: www.ngrl.org.uk/Manchester/publications/MLPA

MRC Holland: www.mlpa.com

Mutation Surveyor software: http://www.softgenetics.com/mutationSurveyor.html

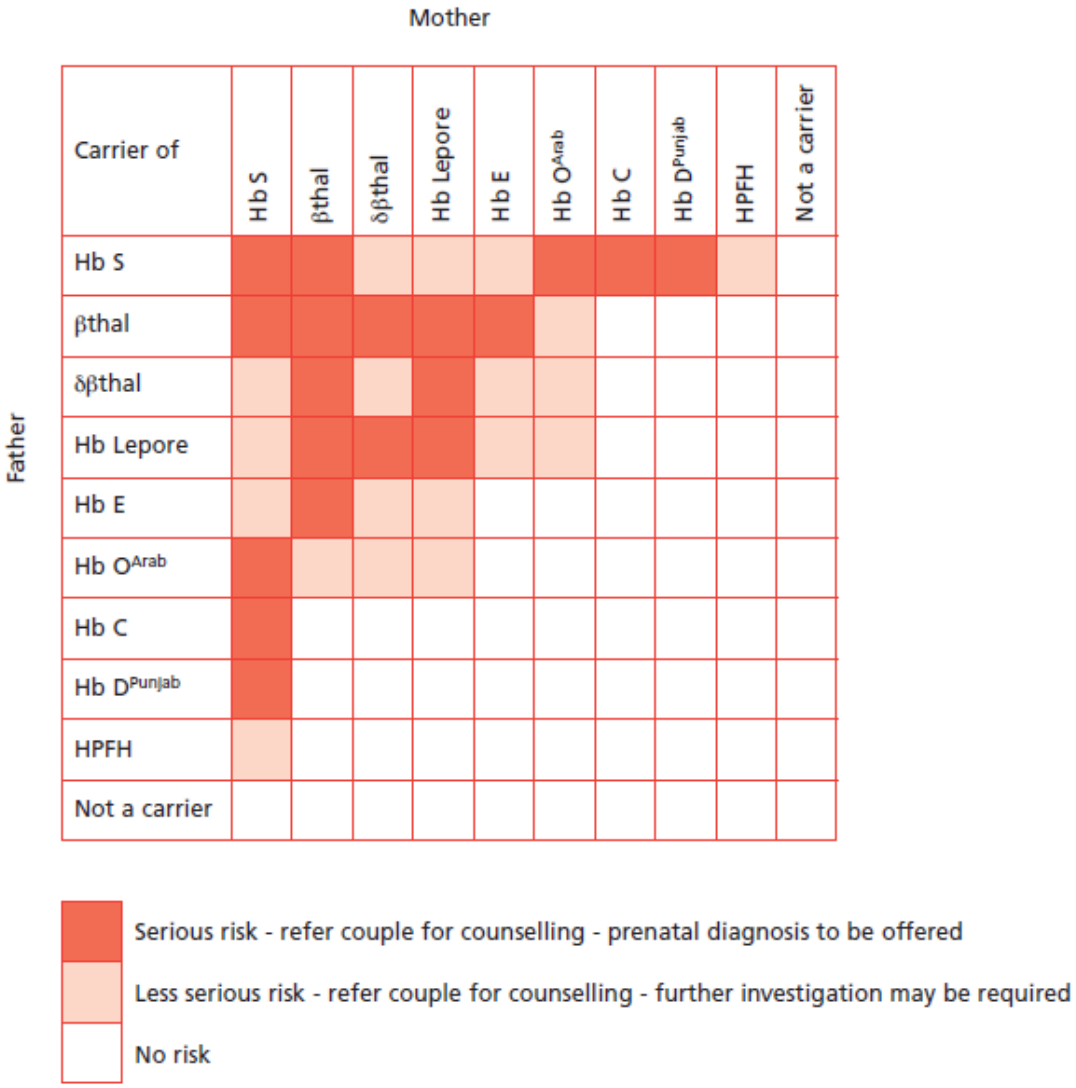
NHS SCT Screening programme: http://sct.screening.nhs.uk/policy

SNP check software (version 2): http://ngrl.manchester.ac.uk/SNPCheckV2/snpcheck

Primer 3 software: http://frodo.wi.mit.edu/

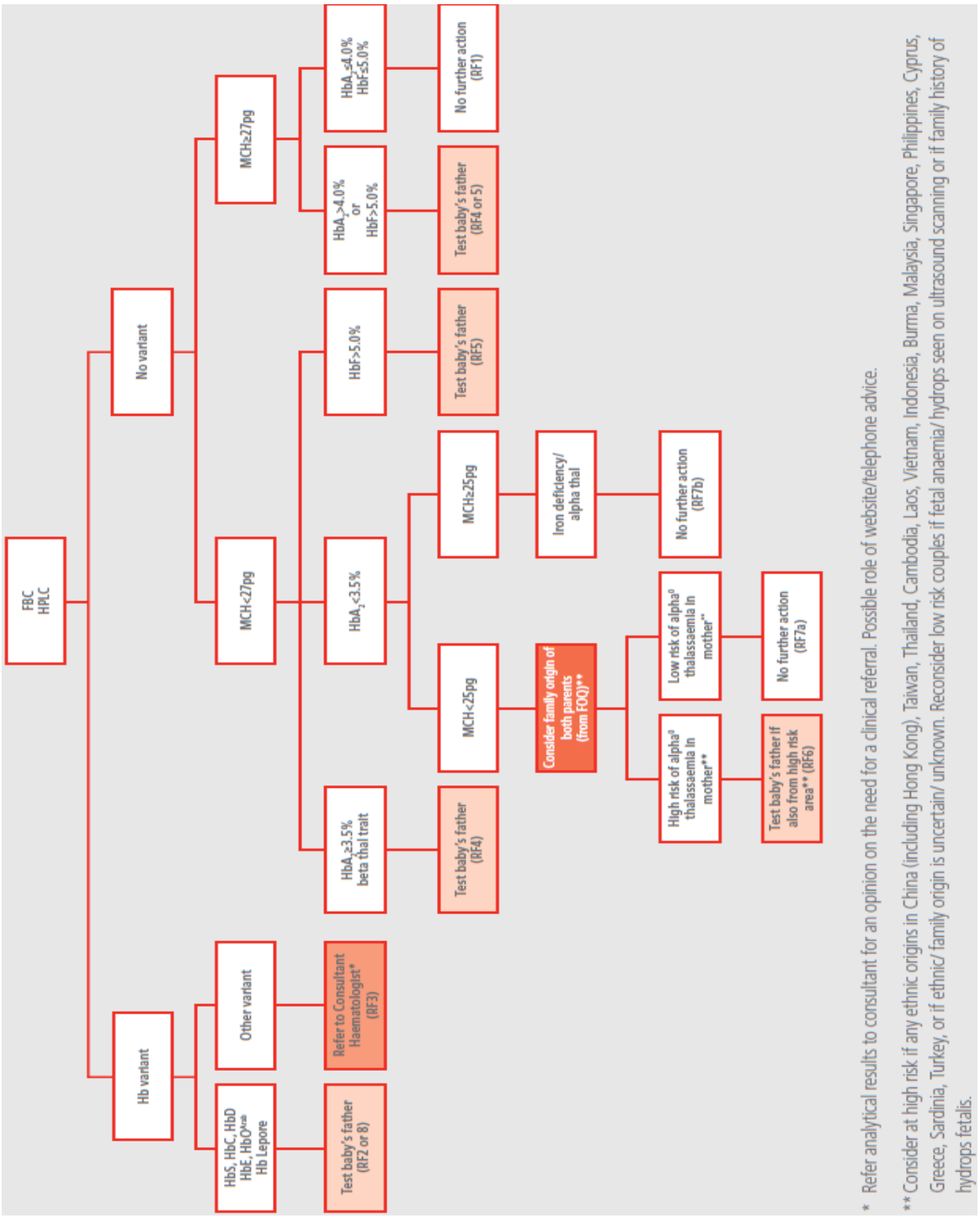
Appendices:

Appendix 1: Table of parental carrier state combinations that give rise to the risk of a fetus with significant sickle cell disease or β-thalassaemia:



Obtained from: NHS SCT screening programme: Handbook for Laboratories September 2009 (http://sct.screening.nhs.uk).

Appendix 2: Testing algorithm for laboratory screening in high prevalence areas:



RF=Report format: (laboratory report formats as recommended by the NHS SCT screening programme)
Obtained from: NHS SCT screening programme: Standards for the linked Antenatal and Newborn Screening Programme. Second Edition. October 201.(http://sct.screening.nhs.uk).

Appendix 3: *HBB* genomic DNA sequence with primer locations.

The forward and reverse primer binding sites are indicated in red, for each of the three separate amplification products of *HBB* (the common N13 tail sequences have not been included). Exonic sequences are highlighted in CAPITAL letters and intronic sequences in lower case (*HBB*: NM_000518.4).

HBB Fragment-1

actcctaagccagtgccagaagagccaaggacaggtaCGGCTGTCATCACTTAGACCTCACCCTGT
GGAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGCTGGGC
ATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGACACAACTGTGTTCACTAGCAA
CCTCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGC
AAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGgttggtatcaaggttacaagacagg
tttaaggagaccaatagaaactgggcatgtggagacagagaagactcttgggtttctgataggcac
tgactctctctgcctattggtctattttcccacccttagGCTGCTGGTGGTCTACCCTTGGACCCA
GAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAA
GGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGG
CACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGgt
gagtctatgggacccttgatgttttctttccccttcttttctatggttaagttcatgtcataggaa
ggggagaagtaacagggtacagtttagaatgggaaacagacgaatgattgcatcagtgtggaagtc
tcaggatcgtttt

HBB Fragment-2

tttctttccccttcttttctatggttaagttcatgtcataggaaggggagaagtaacagggtacag
tttagaatgggaaacagacgaatgattgcatcagtgtggaagtctcaggatcgttttagtttcttt
tatttgctgttcataacaattgttttcttttgtttaattcttgctttctttttttttcttctccgc
aatttttactattatacttaatgccttaacattgtgtataacaaaaggaaatatctctgagataca
ttaagtaacttaaaaaaaaacttacacagtctgcctagtacattactatttggaatatatgtgtg
cttatttgcatattcataatctccctactttatttcttttattttaattgatacataatcatta
tacatatttatgggttaaagtgtaatgttttaatatgtgtacacatattgaccaaatcagggtaat
tttgcatttgtaatttaaaaaatgctttcttctttaatatactttttgtttatcttatttcta
atactttccctaatctctttctttcagggcaataatgatacaatgtatcatgcctctttgcaccat
t

HBB Fragment-3

tggaatatatgtgtgcttatttgcatattcataatctccctactttatttcttttatttttaatt
gatacataatcattatacatatttatgggttaaagtgtaatgttttaatatgtgtacacatattga
ccaaatcagggtaattttgcatttgtaatttaaaaaatgctttcttctttttaatatacttttttg
tttatcttatttctaatactttccctaatctctttctttcagggcaataatgatacaatgtatcat
gcctctttgcaccattctaaagaataacagtgataatttctgggttaaggcaatagcaatatttct
gcatataaatatttctgcatataaattgtaactgatgtaagaggtttcatattgctaatagcagct
acaatccagctaccattctgcttttatttatggttgggataaggctggattattctgagtccaag
ctaggccctttgctaatcatgttcatacctcttatcttcctcccacagCTCCTGGGCAACGTGCT
GGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCACCCCACCAGTGCAGGCTGCCTATCAGAA
AGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCACTAAGCTCGCTTTCTTGCTGTCCA
ATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATATTATGAAGGGCC
TTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCAAtgatgtatttaaattatt
tctgaatattttactaaaaagggaatgtgggaggtcagtgcatttaaaacataaaga

Appendix 4: *HBG2* genomic DNA sequence with primer locations.

The forward and reverse primer binding sites are indicated in red for the *HGB2* gene amplification products (the common N13 tail sequences have not been included). Exonic sequences are highlighted in CAPITAL letters and intronic sequences in lower case (*HBG2:* NM_000184.2).

HBG Fragment

cagccagcacatacacttatccagtgataaatacacatcatcgggtgcctacatacatacctgaat
ataaaaaaaatacttttgctgagatgaaacaggcgtgatttatttcaaataggtacggataagtag
atattgaagtaaggattcagtcttatattatattacataacattaatctattcctgcactgaaact
gttgctttataggatttttcactacactaatgagaacttaagagataatggcctaaaaccacagag
agtatattcaaagataagtatagcacttcttatttggaaaccaatgcttactaaatgagactaaga
cgtgtcccatcaaaaatcctggacctatgcctaaaacacatttcacaatccctgaacttttcaaaa
attggtacatgctttaactttaaactacaggcctcactggagctacagacaagaaggtgaaaaacg
gctgacaaaagaagtcctggtatcttctatggtgggagaagaaaactagctaaagggaagaataaa
ttagagaaaaattggaatgactgaatcggaacaaggcaaaggctataaaaaaaattaagcagcagt
atcctcttgggggccccttccccacactatctcaatgcaaatatctgtctgaaacggtccctggct
aaactccacccatgggttggccagccttgccttgaccaatagccttgacaaggcaaacttgaccaa
tagtcttagagtatccagtgaggccaggggccggcggctggctagggatgaagaATAAAAGGAAGC
ACCCTTCAGCAGTTCCACACACTCGCTTCTGGAACGTCTGAGGTTATCAATAAGCTCCTAGTCCAG
ACGCCATGGGTCATTTCACAGAGGAGGACAAGGCTACTATCACAAGCCTGTGGGGCAAGGTGAATG
TGGAAGATGCTGGAGGAGAAACCCTGGGAAGgtaggctctggtgaccaggacaagggagggaagga
aggaccctgtgcctggcaaaagtccaggtcgcttctcaggatttgtggcaccttctgactgtcaaa
ctgttcttgtcaatctcacagGCTCCTGGTTGTCTACCCATGGACCCAGAGGTTCTTTGACAGCTT
TGGCAACCTGTCCTCTGCCTCTGCCATCATGGGCAACCCCAAAGTCAAGGCACATGGCAAGAAGGT
GCTGACTTCCTTGGGAGATGCCATAAAGCACCTGGATGATCTCAAGGGCACCTTTGCCCAGCTGAG
TGAACTGCACTGTGACAAGCTGCATGTGGATCCTGAGAACTTCAAGgtgagtccaggagatg

Appendix 5: *HBA1* and *HBA2* genomic DNA sequence with primer locations

The forward and reverse primer binding sites are indicated in red, for each of the *HBA1* and *HBA2* gene amplification products (the common N13 tail sequences have not been included). Exonic sequences are highlighted in CAPITAL letters and intronic sequences in lower case (*HBA1*: NM_000558.3 and *HBA2*: NM_000517.4.*HBA1)*.

gtggagggtggagacgtcctggcccccgccccgcgtgcacccccaggggaggccgagcccgccgcc
cggccccgcgcaggccccgcccgggactcccctgcggtccaggccgcgccccgggctccgcgccag
ccaatgagcgccgcccggccgggcgtgcccccgcgccccaagcataaaccctggcgcgctcgcggc
ccggcACTCTTCTGGTCCCCACAGACTCAGAGAGAACCCACCATGGTGCTGTCTCCTGCCGACAAG
ACCAACGTCAAGGCCGCCTGGGGTAAGGTCGGCGCGCACGCTGGCGAGTATGGTGCGGAGGCCCTG
GAGAGgtgaggctccctcccctgctccgacccgggctcctcgcccgcccggacccacaggccaccc
tcaaccgtcctggccccggacccaaaccccacccctcactctgcttctccccgcagGATGTTCCTG
TCCTTCCCCACCACCAAGACCTACTTCCCGCACTTCGACCTGAGCCACGGCTCTGCCCAGGTTAAG
GGCCACGGCAAGAAGGTGGCCGACGCGCTGACCAACGCCGTGGCGCACGTGGACGACATGCCCAAC
GCGCTGTCCGCCCTGAGCGACCTGCACGCGCACAAGCTTCGGGTGGACCCGGTCAACTTCAAGgtg
agcggcgggccgggagcgatctgggtcgaggggcgagatggcgccttcctcgcagggcagaggatc
acgcgggttgcgggaggtgtagcgcaggcggcggctgcgggcctgggccctcggccccactgaccc
tcttctctgcacagCTCCTAAGCCACTGCCTGCTGGTGACCCTGGCCGCCCACCTCCCCGCCGAGT
TCACCCCTGCGGTGCACGCCTCCCTGGACAAGTTCCTGGCTTCTGTGAGCACCGTGCTGACCTCCA
AATACCGTTAAGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCCCCCCAGCCCCTCC
TCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCAgcctgtgtgt
gcctgagtttttccctcagcaaacgtgccaggcatgggcgtggacagcagctgggacacacatgg
ctagaacctctc

HBA2

gtggagggtggagacgtcctggcccccgccccgcgtgcacccccaggggaggccgagcccgccgcc
cggccccgcgcaggccccgcccgggactcccctgcggtccaggccgcgccccgggctccgcgccag
ccaatgagcgccgcccggccgggcgtgcccccgcgccccaagCATAAACCCTGGCGCGCTCGCGGG
CCGGCACTCTTCTGGTCCCCACAGACTCAGAGAGAACCCACCATGGTGCTGTCTCCTGCCGACAAG
ACCAACGTCAAGGCCGCCTGGGGTAAGGTCGGCGCGCACGCTGGCGAGTATGGTGCGGAGGCCCTG
GAGAGgtgaggctccctcccctgctccgacccgggctcctcgcccgcccggacccacaggccaccc
tcaaccgtcctggccccggacccaaaccccacccctcactctgcttctccccgcagGATGTTCCTG
TCCTTCCCCACCACCAAGACCTACTTCCCGCACTTCGACCTGAGCCACGGCTCTGCCCAGGTTAAG
GGCCACGGCAAGAAGGTGGCCGACGCGCTGACCAACGCCGTGGCGCACGTGGACGACATGCCCAAC
GCGCTGTCCGCCCTGAGCGACCTGCACGCGCACAAGCTTCGGGTGGACCCGGTCAACTTCAAGgtg
agcggcgggccgggagcgatctgggtcgaggggcgagatggcgccttcctctcagggcagaggatc
acgcgggttgcgggaggtgtagcgcaggcggcggctgcgggcctgggccgcactgaccctcttctc
tgcacagCTCCTAAGCCACTGCCTGCTGGTGACCCTGGCCGCCCACCTCCCCGCCGAGTTCACCCC
TGCGGTGCACGCCTCCCTGGACAAGTTCCTGGCTTCTGTGAGCACCGTGCTGACCTCCAAATACCG
TTAAGCTGGAGCCTCGGTAGCCGTTCCTCCTGCCCGCTGGGCCTCCCAACGGGCCCTCCTCCCCTC
CTTGCACCGGCCCTTCCTGGTCTTTGAATAAAGTCTGAGTGGGCAGCAgcctgtgtgtgcctgggt
tctctctatccccggaatgtgccaacaatgg

Appendix 6: Details of SALSA MLPA P140-B2 *HBA* probemix, with their corresponding chromosomal positions:

Length (nt)	SALSA MLPA probe	Chromosomal position
64-70-76-82	Q-fragments: DNA quantity; only visible with less than 100 ng sample DNA	
88-92-96	D-fragments: Low signal of 88 OR 96 nt fragment indicates incomplete denaturation	
100	X-fragment: Specific for the X chromosome	
105	Y-fragment: Specific for the Y chromosome	
130	Reference probe 0797-L00463	5q31
136	**HBA region**, probe S0290-L09493	**Signal only on samples containing the Constant Spring mutation !**
142 †	**HBA region**, probe 4630-L04011	**HBA1 +HBA2 exon 1**
148	Reference probe 3707-L03161	9q22
154	**HBA region**, probe 8499-L08423	0.2 kb downstream HBA1
160 ±	**HBA region**, probe 8498-L08422	HBA2 intron 2
166 †	**HBA region**, probe 4632-L06292	**HBA1+HBA2 exon 3**
172	Reference probe 2020-L01539	15q11
178	**HS-40**, probe 4799-L04797	HS-40
184	**HBA region**, probe 4637-L04018	Between HBA2P and HBA1P
190	**HBA region**, probe 4626-L04740	Between HBA2 and HBA1
196	**HBA region**, probe 8491-L08414	End of HBA2 exon 3
201	**HBA region**, probe 4627-L04007	Between HBA1P and HBA2
208	Reference probe 2813-L02242	17q21
214	**HBA region**, probe 8492-L08415	Between HBA1P and HBA2
220 ¥	**HBA region**, probe 8493-L08416	Between HBA2 and HBA1
229	**HBA region**, probe 4628-L04008	Between HBA1P and HBA2
236	**POLR3K** probe 4913-L01316	16p13, 60 Kb telomeric of HS40
240	**HBA region**, probe 4633-L06249	HBA2 intron 2
247	Reference probe 2869-L02336	1p21
256	**HBA region**, probe 8494-L08417	Between HBA2 and HBA1
265	Reference probe 3075-L02475	5p15
274	Reference probe 1542-L00985	5q22
283	**HBA region**, probe 4638-L04019	0.5 kb downstream HBA1
292	**HBA region**, probe 4624-L04004	Between HBZ and HBZP
301	Reference probe 3255-L02692	11p13
310	**HBA region**, probe 4639-L04020	2.4 kb downstream HBA1
318	**HBA region**, probe 4625-L04005	Between HBZ and HBZP
325	Reference probe 3085-L04948	16p13, 2.6 Mb centromeric of HBA1
339 *	**HBA region**, probe 8497-L08420	Between HBA2 and HBA1
346	**HBA region**, probe 4622-L04001	3.5 kb upstream HBZ
355	Reference probe 0547-L00116	11q22
364	**HBA region**, probe 4926-L04017	9.3 kb upstream HBZ
373	**HBA region**, probe 8488-L08410	Between HBA1P and HBA2
382	**HS-40**, probe 4800-L04175	HS-40
391	Reference probe 1701-L01469	5q22
402	**HBA region**, probe 6707-L06294	3.7 kb downstream HBA1
409	Reference probe 3272-L02709	3q29

* The 339 nt probe is not entirely specific and will be replaced in the future. Samples that are homozygous for the alpha-3.7 deletion do not give any signal for 8 out of 9 probes that are deleted. The 339 nt probe however does give a (reduced) signal on such samples.
† The 142 & 166 nt probes detect both *HBA1* and *HBA2*.
¥ SNP rs35557 could influence the probe signal of the 220 nt probe. In case of apparent deletions, it is recommended to sequence the region targeted by this probe. Population frequency is not known.
± SNP rs2362746 influences the 160 nt probe signal. This probe is located at a 1 nt difference between the *HBA1* and *HBA2* intron sequence. We recommend ignoring apparent deletions of only this probe.

(Description version 22; 24/12/2011. Obtained from MRC-Holland: www.mlpa.com).

Appendix 7: Details of SALSA MLPA P102-B1 *HBB* probemix, with their corresponding chromosomal positions:

Length (nt)	SALSA MLPA Probe	Chromosomal position
64-70-76-82	Q-fragments: DNA quantity; only visible with less than 100 ng sample DNA	
88-92-96	D-fragments: Low signal of 88 OR 96 nt fragment indicates incomplete denaturation	
100	X-fragment: Specific for the X chromosome	
105	Y-fragment: Specific for the Y chromosome	
130	Reference probe 0797-L00463	5q31
136	Reference probe 0662-L00158	6p21
142 *	Reference probe 04243-L03598	
148	**HBB promoter** probe 05827-L06319	Promotor
154 *	**HBB** probe 11883-L12683	Intron 1
160	Reference probe 02611-L02082	6p21
166 *°	**HBB** probe 11884-12684	Exon 3
173	**HBB region** probe 05836-L06321	0.8 kb downstream of exon 3
184	Reference probe 03786-L03295	17q11
189	**HBB** probe 05828-L05332	Exon 1
196	**HBB** probe 05833-L05335	Intron 2
206 *	**HBB** probe 11885-L13080	Exon 3
214 * §	***HBB** probe 05830-L05307*	*Hemoglobin S Mutation-specific probe*
223	Reference probe 03862-L03611	3p22
229	**HS3** probe 05810-L05314	
238	**HS1** probe 05812-L06323	
247	Reference probe 01164-L00720	11q13
256	**HS5 region** probe 05804-L05308	
265	**HS4** probe 05807-L05311	
274 *	**HBB** probe 11980-L12803	9 kb downstream of HBB
284	Reference probe 3592-L03997	3q21
292	**HS2** probe 06395-L05315	
301	**HS4** probe 05806-L06325	
310 *	**HBD** probe 11886-L12686	Exon 1
320	Reference probe 03581-L02942	3p21
328	**HS3 region** probe 05808-L05312	
337	**HBB region** probe 05824-L05328	
346	**HS3** probe 05809-L05313	
355	Reference probe 03318-L02736	22q12
364 *	**HBB** probe 11982-L12805	1 kb upstream of HBB
373	**HBG2 region** probe 05815-L05319	
382	**HBBP1** probe 05820-L05324	Exon 3
391	Reference probe 02580-L02042	22q12
399 °	**HBD** probe 06397-L05327	Exon 3
409	**HBG region** probe 05817-L05321	
420	**HBD region** probe 05821-L06327	
427	Reference probe 03929-L03750	15q21
436 †	**HBG2** probe 05816-L05320	Exon 3
445	**HBBP1** probe 06400-L05323	Exon 1
454	Reference probe 04075-L03310	17q11
463	**HBE1** probe 05813-L05317	Exon 1
472	Reference probe 02757-L02206	11p13
483	SMPD1 probe 06676-L06254	1.12 Mb centromeric from HBB

* New in version B1 (from lot 0508 onwards)

§ The 214 nt probe will only give a signal when the sickle cell-specific point mutation (rs334) is present.

° The 166 nt and 399 nt probes have been reported to be variable.

† Please disregard apparent copy number changes detected only by the 436 *HBG2* exon 3 probe. In the sequence detected by this probe (17 nt after the stop codon), there is only a single nucleotide difference between the *HBG1* and *HBG2* gene. Due to their close proximity (5 kb), it is likely that the *HBG2* sequence at this position in some healthy individuals is changed in a *HBG1* sequence (and vice versa) due to gene conversion, without any consequences.

(Description version 14; 24/08/2011. Obtained from MRC-Holland: www.mlpa.com)

Appendix 8: Raw table of all patient information and mutational data

PATIENT INFORMATION				DNA MUTATION DATA					
Patient	**M /F**	**Ethnicity**	**Clinical reason for referral**	**α-multiple Gap-PCR**	***HBB* Sequencing**	***HBA1/HBA2* Sequencing**	**Xmn**	**External DNA analysis**	**MDC Further Analysis**
1	M	Any other	Beta thal ?severity	αα/ααα	Cd 8/9 (+G)	No mutation			
2	F	Any Other: Chilean	Abnomal variant -IVF	αα/αα	Hb Hofu	No mutation			
3	F	ARABIC: A/O Oman	? Alpha zero	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	n/a			
4	M	ARABIC: A/O Oman	? Alpha zero	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	n/a			
5	F	ARABIC: Iranian	?Alpha/HbH	αα/ααα	Cd5 (del CT) (Het)	No mutation	Het -158		
6	F	ARABIC: Iranian	?HbH. HbF=9.2. ↑ for age	αα/ααα	Cd5 (del CT) (Het)	No mutation	Het -158		
7	M	ARABIC: Iranian	?alpha	$\alpha^{-3.7}/\alpha^{-4.2}$	No mutation	n/a			
8	M	ARABIC: Iranian	?silent α-thal trait. HbA2 1.9	αα/αα	No mutation	No mutation			
9	M	ARABIC: Iranian	? Alpha thal	$\alpha\alpha/\alpha^{-3.7}$	No mutation	No mutation			
10	M	ARABIC: Iraq	Beta thal major	αα/αα	Cd8 (del AA:Homo)	No mutation			
11	F	ARABIC: Iraq	Unknown variant 19.9%	αα/αα	No mutation	Hb Q-Iran			
12	M	ARABIC: Iraq	?alpha thal	αα/αα	No mutation	No mutation			
13	M	ARABIC: Kurdish	Alpha thal (non-del)	$\text{--MED}/\alpha^{-}_{3.7}$	No mutation	n/a			
14	M	ARABIC: Kurdish	Hb variant 16.5%	αα/αα	No mutation	Hb Setif			
15	F	ARABIC: Kurdish	IVF-α chain variant ? G-Pest	αα/αα	No mutation	Hb Kudistan			
16	F	ARABIC: Other (Middle Eastern)	↑HbA2. Norm indicies	αα/αα	No mutation	Cd146 (3'UTR G>A)	Poly-369,309		
17	F	ARABIC: Other (Saudi Arabia)	Unknown variant 13.9%	αα/αα	No mutation	Hb Dunn			
18	M	ASIAN	Beta thal ?β° or β+	αα/αα	IVS1-5 (G-C)	n/a	Het -158		

19	M	ASIAN	?δβ thal.HbA2=4.4% HbF=18.5%	αα/$\alpha^{-3.7}$	Minus 88 (C-T)	No mutation	Het-158		
20	F	ASIAN	HbA_2 4% ?mild beta	αα/αα	CAP+1 (A-C)	No mutation			
21	F	ASIAN	?HbE abnormal variant	αα/αα	Hb D-Iran	No mutation			
22	M	ASIAN	HbDD/HbDβ thal	$\alpha^{-3.7}/\alpha^{-3.7}$	Hb D-Punjab (Homo)	n/a			
23	F	ASIAN	HbEE/HbE/B-thal	αα/αα	HbEE	No mutation			
24	M	ASIAN	?βthal major(trace HbA)	αα/αα	IVS1-5 (G>C) Homo	n/a			
25	F	ASIAN	Beta thal ?β° or β+	αα/αα	IVS1-5 (G-C)	n/a			
26	F	ASIAN	?Alpha or δβ thal	αα/αα	No mutation	No mutation			
27	M	ASIAN	Alpha thal	αα/αα	No mutation	No mutation			
28	F	ASIAN	HbEE/E Beta thal	n/a	HbEE	n/a			
29	F	ASIAN: Any Other	?alpha thal	αα/αα	No mutation	IVS1donor.det 5nts(TGAGG)			
30	M	ASIAN: Any Other	A-typical thal M.I raised HbF	αα/αα	No mutation	No mutation		Normal/εγδβ°	
31	F	ASIAN: Bangladeshi	HbE and ?alpha thal	αα/αα	HbAE	No mutation			
32	M	ASIAN: Bangladeshi	HbEE/E/B--thal	αα/αα	HbEE	n/a			
33	F	ASIAN: Bangladeshi	?B-thal at risk couple	αα/αα	IVS1-5 (G-C)	No mutation			
34	M	ASIAN: Bangladeshi	?Hb Lepore. ANC couple	αα/αα	No mutation	No mutation		Hb Lepore	Hb Lepore
35	F	ASIAN: Indian	HbAS ?alpha	αα/αα	HbAS	n/a	Het -158		
36	F	ASIAN: Indian	?HbAS ?alpha	αα/αα	HbAS	No mutation	Het -158		
37	M	ASIAN: Indian	F/Sonly ?HbS/b-thal	αα/αα	HbAS and Cd8/9 (+G)	n/a	Het -158		
38	F	ASIAN: Indian	F/Sonly ?HbS/b-thal	αα/αα	HbAS and Cd8/9 (+G)	n/a	Het -158		
39	M	ASIAN: Indian	?Beta thal major	αα/αα	IVS1-5 (G-C) (Homo)	Cd19 (del G)	Homo-158		
40	F	ASIAN: Indian	?Beta	αα/αα	CAP+1 (A-C)	No mutation			
41	F	ASIAN: Indian	BTT-PND failed	αα/αα	Cd 39 (C-T)	No mutation			
42	M	ASIAN: Indian	BTT-PND failed	αα/αα	Cd 39 (C-T)	No mutation			
43	F	ASIAN: Indian	Known β-thal.α-genetics	αα/αα	Cd 39 (C-T)	No mutation			
44	F	ASIAN: Indian	?Hb Manitoba	αα/αα	No mutation	Hb Manitoba II			

45	M	ASIAN: Indian	?Beta thal couple	αα/αα	Cd 41/42 (-TTCT)	No mutation		
46	M	ASIAN: Indian	?$B^{+}/^{o}$thal (Partner HbD)	αα/αα	Cd 5 (del CT)	No mutation		
47	F	ASIAN: Indian	?Beta thal couple	αα/αα	Cd 8/9 (+G)	No mutation		
48	M	ASIAN: Indian	?Beta thal trait	αα/αα	Cd 8/9 (+G)	No mutation		
49	M	ASIAN: Indian	Beta thal. + or 0	αα/αα	Cd 8/9 (+G)	No mutation		
50	M	ASIAN: Indian	? Alpha+	αα/$\alpha^{-3.7}$	n/a	n/a		
51	M	ASIAN: Indian	Beta thal HbA2=5%	αα/$\alpha^{-3.7}$	IVS1-5 (G-C)	No mutation		
52	F	ASIAN: Indian	Hb Manitoba	αα/αα	No mutation	Hb Manitoba II		
53	F	ASIAN: Indian	?Hb variant 24.6%.Hb Q-India	αα/$\alpha^{-3.7}$	No mutation	Hb Q-India		
54	F	ASIAN: Indian	IVF-variant 17.9% Q-India	αα/αα	No mutation	Hb Q-India		
55	F	ASIAN: Indian	?Variant. 17%	αα/αα	No mutation	Hb Q-India		
56	F	Asian: Indian	?δβ/HPFH. HbF19.2% M.I	αα/αα	No mutation	n/a	Indian 32kb	
57	F	ASIAN: Indian	Beta ?Alpha	αα/αα	No mutation	n/a	Indian 619bp del	Not detected
58	F	ASIAN: Indian	Hb variant HbA2 peak	αα/αα	No mutation	No mutation		
59	F	ASIAN: Indian	?HPFH HbF 28%	αα/αα	No mutation	No mutation	HPFH-3	HPFH-3
60	F	ASIAN:Indian (Sri-Lanka)	Indeterminate HbA2	αα/αα	CAP+1 (A-C)	No mutation		
61	F	ASIAN:Indian (Sri-Lanka)	?Hb variant. ?Shape of A peak	αα/αα	Hb Alperton	No mutation		
62	M	ASIAN:Indian (Sri-Lanka)	? Alpha	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	n/a		
63	F	ASIAN: Other	?thal intermedia HbH bodies	αα/αα	HbAE and Hb Monroe	No mutation Homo-158		
64	F	ASIAN: Other	HbSS or S/βthal.Multi Tx	αα/αα	HbAS/IVS1-110 (G-A)	No mutation		
65	F	ASIAN: Other (Afgan)	Hb D/G variant	αα/αα	Hb D-Punjab	No mutation		
66	M	ASIAN: Other (Afgan)	Hb D/G variant	αα/αα	Hb D-Punjab	No mutation		
67	M	ASIAN: Other (Afgan)	Hb variant (shoulder pk)	αα/αα	No mutation	Santa Barnabas		

68	F	ASIAN: Pakistan	?b-thal intermedia	αα/αα	Cd 8/9 (+G)	No mutation		
69	F	ASIAN: Pakistan	β-thal inter.(HbF 59%)	αα/αα	Minus 88 (C-T) Homo	No mutation		
70	F	ASIAN: Pakistani	HbSS/HbS/Beta thal or HPFH	αα/αα	HbSS	No mutation	Homo-158	
71	M	ASIAN: Pakistani	No HbA. 100% HbF	αα/$\alpha^{-3.7}$	IVS2-1 (G>A) (Homo)	No mutation	Homo-158	
72	F	ASIAN: Pakistani	β-thal.rpt sample	αα/αα	CAP+1 (A-C)	No mutation		
73	F	ASIAN: Pakistani	B-thal trait	$\alpha^{-3.7}/\alpha^{-3.7}$	CAP+1 (A-C)	n/a		
74	F	ASIAN: Pakistani	b-thal trait	αα/αα	Cd 41/42 (-TTCT)	n/a		
75	M	ASIAN: Pakistani	B-Thal	αα/αα	Cd 41/42 (-TTCT)	n/a		
76	M	ASIAN: Pakistani	?Beta thal/ααα	αα/ααα	Cd 8/9 (+G)	n/a		
77	F	ASIAN: Pakistani	Beta thal/intermedia	αα/αα	Cd 8/9 (+G)	No mutation		
78	M	ASIAN: Pakistani	Beta thal	αα/αα	Cd 8/9 (+G)	No mutation		
79	F	ASIAN: Pakistani	B-thal intermedia	αα/αα	Cd 8/9 (+G)	No mutation		
80	F	ASIAN: Pakistani	Hb variant 54.7%	αα/$\alpha^{-3.7}$	Hb Camden	No mutation		
81	F	ASIAN: Pakistani	HbDD/HbD/Bthal	αα/αα	Hb D-Punjab/IVS1-5G>C	No mutation		
82	F	ASIAN: Pakistani	Hb variant (39.4%)	αα/$\alpha^{-3.7}$	Hb D-Iran	n/a		
83	F	ASIAN: Pakistani	Hb variant ?D Iran	αα/αα	Hb D-Iran	No mutation		
84	M	ASIAN: Pakistani	?Hb D trait	αα/$\alpha^{-3.7}$	Hb D-Punjab	IVS2-55 poly		
85	F	ASIAN: Pakistani	?HbE trait. 24%	αα/αα	Hb D-Punjab	No mutation		
86	F	ASIAN: Pakistani	Hb D/G ?D disease	$\alpha^{-3.7}/\alpha^{-3.7}$	Hb D-Punjab (Homo)	n/a		
87	F	ASIAN: Pakistani	HbDD/HbDβ thal	αα/αα	Hb D-Punjab/Cd15 G>A	Poly A (del AA)		
88	M	ASIAN: Pakistani	Hb J variant	αα/$\alpha^{-3.7}$	Hb Malmo	No mutation		
89	F	ASIAN: Pakistani	Hb J variant	αα/αα	Hb Malmo	No mutation		
90	M	ASIAN: Pakistani	variant 46.7% parents Hb Malmo	αα/αα	Hb Malmo	No mutation		
91	M	ASIAN: Pakistani	variant 45.9% parents Hb Malmo	αα/$\alpha^{-3.7}$	Hb Malmo	IVS2-55 poly		
92	F	ASIAN: Pakistani	?B-thal	αα/$\alpha^{-3.7}$	Hb Monroe -42 poly	n/a		

93	M	ASIAN: Pakistani	ANC partner.?B-thal	αα/-$\alpha^{4.2}$	Hb Monroe -42 poly	n/a		
94	F	ASIAN: Pakistani	? Non-deletional alpha thal	αα/αα	Normal	n/a	No mutation: Normal HbF/A2. Unlikely B-thal	
95	F	ASIAN: Pakistani	?thal intermedia/major	αα/αα	Minus 88 (C-T) Homo	IVS1-117 (G-A)		
96	F	ASIAN: Pakistani	?thal intermedia/major (Tx)	αα/αα	Minus 88 (C-T) Homo	No mutation		
97	F	ASIAN: Pakistani	?Hb J variant (19.5%)	αα/αα	No mutation	Hb J-Meerut		
98	M	ASIAN: Pakistani	B-thal trait ? Normal	αα/αα	No mutation	n/a		
99	F	ASIAN: Pakistani	HbF 5.4% (husband βtrait)	αα/αα	No mutation	n/a	Non-del HPFH unidentified	
100	F	ASIAN: Pakistani	?Alpha thal/Fe def	αα/$\alpha^{-3.7}$	No mutation	No mutation		
101	F	ASIAN: Pakistani	?alpha thal	αα/$\alpha^{-3.7}$	No mutation	No mutation		
102	F	ASIAN: Pakistani	Hb variant 14.4%	αα/αα	No mutation	Hb Fontainebleau		
103	F	ASIAN: Pakistani	Indeterminate HbA2	αα/αα	No mutation	No mutation		
104	F	BLACK: African	Raised HbF	$\alpha^{-3.7}/\alpha^{-3.7}$	Minus 88 (C-T)	n/a	Het -158	
105	M	BLACK: African	ANC HbAS	αα/αα	HbAS	n/a	Het 6nt del	
106	M	BLACK: African	?HbS/HPFH	αα/αα	HbSS	No mutation	Poly:-369,309.-158	
107	M	BLACK: African	HbC.?SAT/δβthal	αα/$\alpha^{-3.7}$	HbCC	No mutation	Poly-309	HPFH 1 or 2	HPFH.2
108	M	BLACK: African	HbA2 variant 1.9%.HbF1.7%	$\alpha^{-3.7}/\alpha^{-3.7}$	Minus 29 (A-G)	n/a	Het-309/-378(C>T)	Prob δ–chain variant	
109	M	BLACK: African	ANC HbAS	$\alpha^{-3.7}/\alpha^{-3.7}$	HbAS	n/a	Het-369+ Homo-309		

110	F	BLACK: African	ANC HbAS=33.5%	αα/$\alpha^{-3.7}$	HbAS	n/a	hom 6nt del-307 A-G		
111	F	BLACK: African	?HbS/HPFH/δβ	αα/$\alpha^{-3.7}$	HbSS (?Hemi)	No mutation	Homo poly-309+-369.	HbS/$\alpha^{3.7}$HPFH/δβ deletion un-i.d	
112	M	BLACK: African	HBAS/Beta thal	αα/αα	Hb AS/Minus 88 (C-T)	No mutation	Homo-309		
113	M	BLACK: African	?HbS/HPFH/δβ HbS47%.HbF21.7%	αα/$\alpha^{-3.7}$	HbAS	No mutation	Minus 202(C>G) HPFH		
114	F	BLACK: African	ANC HbAS	αα/αα	HbAS	n/a	Minus 369 and -309		
115	F	BLACK: African	?βthal trait/intermedia	αα/αα	Cd 39 (C-T)	No mutation			
116	F	BLACK: African	IVF-Unknown Hb 49.8%	αα/αα	Hb Camden	No mutation			
117	F	BLACK: African	Abnormal Hb:fast band 40%	αα/αα	Hb J-Guantanamo	No mutation			
118	F	BLACK: African	?Hb Stanleyville	αα/αα	n/a	Hb Stanleyville-II			
119	M	BLACK: African	ANC HbAS-M.I	$\alpha^{-3.7}/\alpha^{-3.7}$	HbAS	n/a			
120	F	BLACK: African	HbAS and ?alpha	αα/$\alpha^{-3.7}$	HbAS	n/a			
121	F	BLACK: African	ANC.HbAS	αα/$\alpha^{-3.7}$	HbAS	n/a			
122	F	BLACK: African	ANC.HbAS	αα/$\alpha^{-3.7}$	HbAS	n/a			
123	F	BLACK: African	? Hb Variant ?G-Phil	αα/αα	Normal	Hb Stanleyville-II			
124	M	BLACK: African	HbAS	n/a	HbAS	Homo +832			
125	F	BLACK: African	HbAS	n/a	HbAS	n/a			
126	M	BLACK: African	HbAS	n/a	HbAS	n/a			
127	M	BLACK: African	ANC.HbAS	αα/$\alpha^{-3.7}$	HbAS	n/a			
128	F	BLACK: African	HbAS/Alpha	αα/$\alpha^{-3.7}$	HbAS	No mutation			

129	F	BLACK: African	?HbS/G-Philadelphia	αα/$\alpha^{-3.7}$	HbAS	G-Philadelphia		
130	F	BLACK: African	HbS/HbS/G-Phil	αα/$\alpha^{-3.7}$	HbAS	G-Philadelphia		
131	F	BLACK: African	HbAS at risk couple	αα/αα	HbAS	n/a		
132	M	BLACK: African	HbAS at risk couple	αα/αα	HbAS	n/a		
133	F	BLACK: African	HbAS-at risk couple	αα/αα	HbAS	n/a		
134	M	BLACK: African	HbAS-at risk couple	αα/αα	HbAS	n/a		
135	F	BLACK: African	HbAS with severe pain crisis?α-variant	αα/$\alpha^{-3.7}$	HbAS	No mutation		
136	M	BLACK: African	HbAS. Multiple Tx.HbS=7%	αα/αα	HbAS	IVS1-5 (G-A)		
137	M	BLACK: African	Variant@HbA_2/Alpha	αα/$\alpha^{-3.7}$	HbAS/Minus 83poly	Fort Worth:Homo		
138	M	BLACK: African	HbAS+?α-thal/Fe def	αα/αα	HbAS/ Hb City of Hope	No mutation		
139	M	BLACK: African	?Sickle/beta thal	αα/αα	HbSS	IVS2-55 poly		
140	F	BLACK: African	Beta thal/intermedia	αα/$\alpha^{-3.7}$	Int Codon (A-G)	IVS2-55 poly		
141	F	BLACK: African	abnormal variant	αα/αα	No mutation	Hb Buffalo		
142	M	BLACK: African	Hb D/G. 42.2%	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	G-Philadelphia		
143	M	BLACK: African	Hb variant. ?Hb D/G	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	G-Philadelphia		
144	F	BLACK: African	Hb variant (HbA_2 19.3%)	αα/αα	No mutation	Hb Toulon		
145	M	BLACK: African	Variant@HbA_2/Alpha	αα/αα	Minus 83poly (G-A)	Fort Worth:Homo		
146	M	BLACK: African	?α-zero partner of above	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	n/a		
147	F	BLACK: African	Abnormal variant 21.6%	αα/αα	n/a	Hb Matsue-Oki		
148	F	BLACK: African	Hb variant:23.6% ?α-thal	αα/αα	No mutation	Hb Stanleyville II		
149	F	BLACK: African	?α not at risk-not sent Oxford	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	n/a		
150	M	BLACK: African	?Unstable Hb variant	αα/$\alpha^{-3.7}$	No mutation	No mutation		
151	F	BLACK: African	HPFH/δβ	αα/αα	No mutation	No mutation	δβ trait rare novel del	

152	F	BLACK:Afro-Caribbean	?alpha	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	n/a			
153	M	BLACK:Afro-Caribbean	?HbS/Beta thal	αα/αα	HbAS/Minus 88 (C-T)	No mutation			
154	F	BLACK: British	?δβ-mother is δβ-trait	αα/-$\alpha^{4.2}$	No mutation	No mutation	Het -158, -309 (A-G)	12Kb Black-δβ	
155	F	BLACK: British	?δβ-mother is δβ-trait	αα/αα	No mutation	No mutation	Homo-158	12Kb Black-δβ	
156	M	BLACK: British	HbAS/?alpha (HbS20%)	$\alpha^{-3.7}/\alpha^{-3.7}$	HbAS	n/a			
157	F	BLACK: British	Bthal trait	αα/αα	IVS2-849 (A-G)	No mutation			
158	M	BLACK: Caribbean	?C/βthal	αα/αα	Minus-88 (C-T)/HbAC	IVS2-55 poly	Het-309		
159	F	BLACK: Caribbean	Hb SS, persistent ↑HbF	αα/$\alpha^{-3.7}$	HbAS	No mutation	Het-368	?Form of HPFH del	HPFH.2
160	F	BLACK: Caribbean	BTT (partner HbAS)	αα/αα	Cd 24 (T-A)	No mutation			
161	F	BLACK: Caribbean	C Beta ? Exisiting thal	αα/αα	Hb CC	n/a		HPFH 1 or 2	HPFH.2
162	F	BLACK: Caribbean	HbS 25% ?Alpha M.I	αα/$\alpha^{-3.7}$	HbAS	No mutation			
163	F	BLACK: Caribbean	Mother of X	αα/$\alpha^{-3.7}$	HbSS	No mutation			
164	F	BLACK: Caribbean	BTT-considering PND	αα/$\alpha^{-3.7}$	Minus 29	No mutation			
165	F	BLACK: Caribbean	↑Hb A2. Normal indicies	αα/$\alpha^{-3.7}$	Minus 88 C>T(Homo)	No mutation			
166	F	BLACK: Caribbean	?HPFH HbF 11.6%	αα/αα	No mutation	n/a		Non-del HPFH γ-gene (unidentified)	No mutation
167	F	BLACK: Caribbean	Var:14.8% co-elute HbA_2	αα/αα	No mutation	Spanish Town			
168	M	BLACK: Caribbean	?β (M.I HbF 23.1%)	αα/αα	No mutation	No mutation		HPFH-1 or 2	HPFH 2
169	M	BLACK: Carribean	?δβ/HPFH	αα/$\alpha^{-3.7}$	No mutation	No mutation	Het-158		HPFH.1
170	F	BLACK: Carribean	?HbJ variant 42.7%	αα/αα	Hb Alamo	No mutation			
171	M	MEDITERRANEAN: Greek	ANC-? Beta	αα/αα	Minus-87 (C>G)	No mutation	Homo-158		
172	F	MED: Greek /Turkish	?Alpha thal	αα/--MED	No mutation	No mutation			
173	M	MED: Greek Cypriot	Raised HbA2 and F	αα/$\alpha^{-3.7}$	IVS1-110 (G-A)	No mutation			

174	F	MED: Other (Bulgaria)	?Hb variant	αα/αα	No mutation	Hb O-Padova		
175	M	MED: Other (Portugal)	?δβ/HPFH (Partner HbAS)	αα/αα	No mutation	No mutation		HPFH.2
176	F	MED: Turkish	HbAS+Poly A alpha thal	αα/αα	HbAS	No mutation		
177	F	MIXED: Afro/E.U	HbD/S and Hb X detected	αα/$\alpha^{-3.7}$	HbAS	G-Philadelphia		
178	F	MIXED: Asian/E.C	HbEE/E Beta thal	αα/αα	HbEE	No mutation		
179	F	Mixed: Black Caribbean	β variant ? Osu Chistiansborg	αα/$\alpha^{-3.7}$	Hb Osu Christiansborg	n/a		
180	F	MIX: British/Pakistani	BTT/HPFH	αα/αα	Minus 88 (C-T:Homo)	n/a		
181	F	MIX: British+Pakistan	Hb variant 16.2%	αα/αα	No mutation	Hb West One		
182	F	MIX: Chinese/Afro	?Alpha and beta	αα/--SEA	No mutation	No mutation		
183	F	MIX: Chinese/British	Thal trait. ?Alpha zero	αα/--SEA	No mutation	No mutation		
184	F	MIX: English/Thai	Hb A/E and ?α-zero thal	αα/--SEA	HbAE	No mutation		
185	M	MIX: Half Chinese	Alpha zero	αα/--SEA	No mutation	n/a		
186	F	MIX: Indian/French	?Hb Manitoba	αα/αα	No mutation	Hb Manitoba II		
187	M	MIX: Indian/French	?Hb Manitoba	αα/αα	No mutation	Hb Manitoba II		
188	M	MIX:Middle E/Chinese	?B, partner HbAS/?alpha	αα/αα	No mutation	No mutation		
189	F	MIX: Part chinese	B-thal. LWH HbA24.1 (MRI 3.6%)	αα/αα	Unknown: Cd133 (G-C)	Poly+861 (A-G)		
190	F	MIX: UK/Ital/Spanish	Unknown variant 45.7%	αα/αα	Hb Old Dominion	No mutation		
191	F	MIX:White/Chinese	Hb variant 16.8% ?HbD/G	αα/αα	No mutation	Hb G-Norfolk		
192	F	NOT STATED	?Homozygous E (71%) (HbF 7%)	αα/αα	HbEE	Hb Fontainebleau	Het -158	
193	M	NOT STATED	α°/HbH-3 mth follow up	αα/--SEA	No mutation	No mutation	Het -158	
194	F	NOT STATED	?b=thal/ Normal red cell	αα/αα	No mutation	Poly+861 (A-G)	Het-158/ -317 (G>A)	
195	M	NOT STATED	?δβ thal (HbF 18%)	αα/αα	IVS1-5 (G-C)	Hb Evanston	Hetero-158/ -408	

196	F	Not Stated	?M haemoglobin	αα/αα	Minus 56 (G>C)	No mutation Het-158		
197	F	NOT STATED	Hb variant unknown	αα/αα	n/a	Hb Chesapeake		
198	F	NOT STATED	?HPFH/Beta	αα/$\alpha^{-3.7}$	CAP+1 (A-C)/IVS1-5 (G-C)	No mutation		
199	F	NOT STATED	Hb variant unknown	αα/αα	n/a	Hb J-Meerut		
200	F	NOT STATED	?Abnormal Hb (37.2%)	αα/αα	Hb D-Iran	No mutation		
201	F	NOT STATED	?Hb D-Iran	αα/αα	Hb D-Iran	No mutation		
202	M	NOT STATED	?Alpha thal	αα/--FIL	n/a	n/a		
203	F	NOT STATED	Abnormal Variant	αα/αα	Hb K-Woolwich	No mutation		
204	F	Not Stated	?Hb variant 24.8%	αα/αα	Hb Leiden	n/a		
205	M	Not Stated	?Hb Olympia	αα/αα	Hb Olympia	n/a		
206	F	NOT STATED	Abnormal Variant 50%	αα/αα	Hb Pyrgos	No mutation		
207	M	NOT STATED	Hb variant ?flat A pk guthrie	αα/αα	Hb Tyne	No mutation		
208	F	NOT STATED	Hb variant :27%	$\alpha^{-3.7}/\alpha^{-3.7}$	HbAE	n/a		
209	F	NOT STATED	Hb S only on HLPC(69.4%).HbF 20.0%	$\alpha^{-3.7}/\alpha^{-3.7}$	HbAS/IVS1-5 (G-C)	n/a		
210	F	NOT STATED	HbE/β Thal	αα/αα	HbEE	n/a		
211	F	NOT STATED	HbEE/Eβ-thal	αα/αα	HbEE	n/a		
212	F	NOT STATED	HbF 14.2%	$\alpha^{-3.7}/\alpha^{-3.7}$	IVS1-5 (G-C)	n/a		
213	M	NOT STATED	?beta thalassaemia	αα/αα	IVS2-1 (G-A)	n/a		
214	F	NOT STATED	Hb variant. No HbA2.?Manitoba	αα/αα	n/a	Hb Manitoba (I)		
215	F	NOT STATED	? Alpha	αα/$\alpha^{-3.7}$	No mutation	n/a		
216	M	NOT STATED	?alpha. Persis. low MCV	αα/$\alpha^{-3.7}$	No mutation	No mutation		
217	M	NOT STATED	Nomal (not sent) ANC partner	αα/αα	No mutation	No mutation		
218	M	NOT STATED	? Alpha	αα/αα	No mutation	No mutation		
219	M	NOT STATED	?Pyruvate def.?Unstable variant	αα/αα	No mutation	Poly+861 (A-G)		
220	F	NOT STATED	?HbA2 variant. M.I	αα/$\alpha^{-3.7}$	No mutation	Hb Seal Rock		

221	M	NOT STATED	Raised HbA2+HbF	αα/αα	No mutation	No mutation	Afghan 909bp β° del	
222	M	NOT STATED	alpha thal	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	n/a		
223	M	NOT STATED	HbA2 1.8%	$\alpha\alpha/\alpha^{-3.7}$	No mutation	No mutation		
224	F	SEA	?HbEE (96.3%E)	$\alpha\alpha/\alpha^{-3.7}$	HbEE	No mutation		
225	F	SEA	Abnormal variant 19.7%	αα/αα	No mutation	Hb Q-India		
226	F	SEA	?Alpha	αα/αα	No mutation	No mutation		
227	M	SEA	?alpha zero	αα/αα	No mutation	Hb Quong Sze		
228	M	SEA: Chinese	?alpha zero	αα/--SEA	No mutation	No mutation Het -158		
229	F	SEA: Chinese	Beta thal carrier	αα/αα	Cd 41/42 (-TTCT)	n/a		
230	M	SEA: Chinese	?Beta/underlying alpha	αα/αα	Cd 41/42 (-TTCT)	No mutation		
231	F	SEA: Chinese	Beta thal	αα/αα	Cd 41/42 (-TTCT)	No mutation		
232	F	SEA: Chinese	HbH disease	$\text{--SEA}/\alpha^{-4.2}$	n/a	n/a		
233	F	SEA: Chinese	Alpha zero	αα/--SEA	n/a	n/a		
234	F	SEA: Chinese	?Alpha zero	αα/--SEA	n/a	n/a		
235	F	SEA: Chinese	Hb variant (43.5%)	αα/αα	Hb J-Bangkok	n/a		
236	M	SEA: Chinese	βthal (rule out α-thal)	αα/αα	IVS2-654 (C-T)	No mutation		
237	F	SEA: Chinese	βthal (rule out α-thal)	αα/αα	IVS2-654 (C-T)	No mutation		
238	F	SEA: Chinese	βthal/αthal	αα/αα	IVS2-654 (C-T)	No mutation		
239	F	SEA: Chinese	?Hb variant 30.8%	$\alpha\alpha/-\alpha^{4.2}$	No mutation	Hb Q-Thailand		
240	F	SEA: Chinese	?Hb variant Q-Thailand:30%	$\alpha\alpha/-\alpha^{4.2}$	No mutation	Hb Q-Thailand		
241	M	SEA: Chinese	?Hb variant: HbG	αα/αα	No mutation	Hb Ube-4		
242	M	SEA: Chinese	?Alpha	αα/--SEA	n/a	No mutation		
243	M	SEA: Chinese	?alpha thal	αα/--SEA	n/a	No mutation		
244	F	SEA: Chinese	?alpha zero carrier	αα/--SEA	n/a	No mutation		
245	M	SEA: Chinese	?Alpha thal	αα/--SEA	n/a	No mutation		
246	F	SEA: Chinese	?Alpha at risk couple	$\alpha\alpha/-\alpha^{4.2}$	n/a	No mutation		
247	M	SEA: Chinese	?HPFH ?ANC couple	αα/αα	No mutation	n/a	Chinese-δβ°	Chinese-δβ°

248	F	SEA: Chinese	HbH disease	--SEA/$\alpha^{-3.7}$	No mutation	n/a		
249	F	SEA: Chinese	?alpha zero/Hb variant	αα/--SEA	No mutation	No mutation		
250	M	SEA: Chinese	Alpha thal (Son HbH)	αα/--SEA	No mutation	No mutation		
251	M	SEA: Chinese	?Alpha (Father)	αα/--SEA	No mutation	No mutation		
252	F	SEA: Chinese	?alpha (Daughter)	αα/--SEA	No mutation	No mutation		
253	F	SEA: Chinese	?Alpha	αα/--SEA	No mutation	No mutation		
254	M	SEA: Chinese	?Alpha zero thal	αα/--SEA	No mutation	No mutation		
255	F	SEA: Chinese	?Alpha thal (HbA2 2.6%)	αα/--SEA	No mutation	No mutation		
256	F	SEA: Chinese	?Alpha thal	αα/--SEA	No mutation	No mutation		
257	F	SEA: Chinese	Conf of alpha zero	αα/--SEA	No mutation	No mutation		
258	F	SEA: Chinese	?Alpha	$\alpha^{-3.7}/\alpha^{-4.2}$	No mutation	n/a		
259	M	SEA: Chinese	?alpha partner of X	αα/αα	No mutation	Constant Spring		
260	F	SEA: Chinese	Alpha thal (Son HbH)	αα/$\alpha^{-3.7}$	No mutation	No mutation		
261	F	SEA: Chinese	?Alpha thal	αα/$\alpha^{-3.7}$	No mutation	IVS1-55 poly		
262	M	SEA: Chinese	?Hb Variant 29.2%	αα/$\alpha^{-3.7}$	No mutation	Hb G27G-Honolulu (Homo)		
263	F	SEA: Chinese	Unknown variant 27.3%	αα/αα	No mutation	Hb Linwood		
264	M	SEA: Chinese	Hb variant 23.5 % (mum Hb Linwood)	αα/αα	No mutation	Hb Linwood		
265	F	SEA: Chinese	?HbH	αα/--SEA	No mutation	Hb Quong Sze		
266	F	SEA: Chinese	?B-thal borderline HbA_2	αα/αα	No mutation	No mutation		
267	F	SEA: Chinese	alpha thal	αα/--SEA	No mutation	No mutation		
268	M	SEA: Chinese	αthal	αα/--SEA	No mutation	No mutation		
269	F	SEA: Filipino	?Hb variant@42.8% (IVF)	αα/αα	Hb J-Bangkok	No mutation		
270	F	SEA: Filipino	HbH	--SEA/$\alpha^{-3.7}$	No mutation	n/a		
271	M	SEA: Filipino	? Alpha thal	αα/--FIL	No mutation	n/a		
272	M	SEA: Filipino	?Alpha (Daughter HbH)	αα/--FIL	n/a	No mutation		
273	F	SEA: Filipino	?Alpha (Daughter HbH)	αα/$\alpha^{-3.7}$	n/a	No mutation		

274	M	SEA: Filipino	?alpha zero couple	αα/--SEA	No mutation	No mutation		
275	F	SEA: Filipino	?alpha thal	αα/--SEA	No mutation	No mutation		
276	F	SEA: Filipino	?alpha thal	$\alpha^{-3.7}/\alpha^{-3.7}$	No mutation	n/a		
277	F	SEA: Filipino	?alpha zero couple	αα/αα	No mutation	No mutation	No mutation	
278	F	SEA: Indonesian	?Alpha zero	αα/--SEA	No mutation	No mutation		
279	F	SEA: Indonesian	miscarriage ?α-zero	$\alpha\alpha/\alpha^{-3.7}$	No mutation	No mutation		
280	F	SEA: Malaysian	?alpha zero/HbE	αα/αα	Hb AE	PolyA: AATAA>AATA (AA del)		
281	M	SEA: Malaysian	Beta thal ?underlying α-thal	αα/αα	IVS1-5 (G>C)	No mutation		
282	F	SEA: Malaysian	?Alpha zero	αα/--SEA	n/a	No mutation		
283	M	SEA: Malaysian	?Alpha	αα/--SEA	No mutation	No mutation		
284	F	SEA: Malaysian	?Alpha thal	αα/αα	No mutation	No mutation		
285	F	SEA: Other (Mongolia)	?HPFH HbF 12.9	αα/αα	No mutation	n/a	Non-del HPFH γ-gene (inidentified)	No mutation
286	F	SEA: Thai	?EE/Eβ thal and ?αzero	$\alpha^{-3.7}/\alpha^{-3.7}$	HbEE	n/a Het -158/-406A-G		
287	F	SEA: Thai	?HbE (22.6%)	αα/--SEA	Hb AE	No mutation		
288	F	SEA: Thai	HbEE	αα/αα	Hb EE	No mutation		
289	M	SEA: Thai	↑HbA2 (ANC partner)	αα/αα	No mutation	Constant Spring		
290	F	SEA: Thai	Alpha thal	αα/--SEA	No mutation	No mutation		
291	M	SEA: Thai	?Alpha thal	αα/--SEA	No mutation	No mutation		
292	F	SEA: Vietnamese	?Alpha zero	αα/--SEA	No mutation	No mutation		
293	M	WHITE: Caucasian	?alpha aquired/hered.	αα/αα	No mutation	No mutation		
294	F	WHITE: English	?Hb Le Lamentin (23.3%)	αα/αα	n/a	Hb Brugg		
295	M	WHITE: English	B-thal/HPFH (HbF 9.4)	αα/αα	Cd 39 (C-T)	No mutation		
296	M	WHITE: English	?Beta thal trait	αα/αα	Cd 39 (C-T)	No mutation		

297	F	WHITE: English	?Hb var: 34.8%.Hb Bushwick	αα/αα	Hb Bushwick	No mutation		
298	F	WHITE: English	?Hb var:37.2%.Hb Bushwick	αα/αα	Hb Bushwick	No mutation		
299	F	WHITE: English	?Hb D-punjab	αα/αα	Hb D-Punjab	n/a		
300	F	WHITE: English	?Hb D-Punjab 37.6% (antenatal)	αα/αα	Hb D-Punjab	No mutation		
301	M	WHITE: English	variant ?Hb Koln 7.5%. ↑MCV+MCH	αα/αα	Hb Koln	No mutation		
302	F	WHITE: English	Hbvar43.6%.Old Dom	αα/αα	Hb Old Dominion	No mutation		
303	F	WHITE: English	?Hb Osu Christiansborg	αα/$\alpha^{-3.7}$	Hb Osu Christiansborg	IVS2-55 poly		
304	M	WHITE: English	?Hb J Oxford. PartnerHbAS	αα/αα	Hb South Florida	No mutation		
305	F	WHITE: English	Abnormal variant 8.2%	αα/αα	Hb South Florida	No mutation		
306	F	WHITE: English	?Hb Tyne (as in father)	αα/αα	Hb Tyne	No mutation		
307	M	WHITE: English	sister has HbH disease	$\alpha^{-3.7}$/--^{BRIT}	n/a	n/a		
308	M	WHITE: English	?beta thal/MDS	αα/αα	IVS2-654 (C-T)			
309	F	WHITE: English	?Hb Le Lamentin (23.6%)	αα/αα	No mutation	Hb Brugg		
310	M	WHITE: English	?variant LWH-split A pk (norm:MRI)	αα/αα	No mutation	Hb Riccarton		
311	F	WHITE: English	mother has HbH disease	αα/--BRIT	No mutation	n/a		
312	F	WHITE: English	?alpha thal	αα/--BRIT	No mutation	No mutation		
313	M	WHITE: English	?BRIT alpha	αα/--BRIT	No mutation	No mutation		
314	F	WHITE: English	?alpha BRIT thal	αα/--BRIT	No mutation	No mutation	αα/--BRIT	αα/--BRIT
315	M	WHITE: English	?BRIT Alpha	αα/--BRIT	No mutation	Poly IVS2-55	αα/--BRIT	αα/--BRIT
316	F	WHITE: English	Alpha thalassaemia	αα/αα	No mutation	Poly+861 (A-G)		
317	M	WHITE: English	?Alpha-pers. ↓MCV+MCH	αα/αα	No mutation	No mutation		
318	F	WHITE: English	↓MCV unresponsive Fe	αα/αα	No mutation	No mutation		
319	F	WHITE: English	↑HbA2. Norm red cell	αα/αα	No mutation	No mutation		

320	M	WHITE: English	?alpha thal	αα/ααα	No mutation	No mutation		
321	M	WHITE: English	?alpha thal	αα/αα	+1570 (T>C. 5' Poly A)	No mutation		
322	F	WHITE: English	b-thal	αα/αα	Poly A (-AT del)	No mutation		
323	F	WHITE: English (WIGAN)	?Hydrops, severe alpha	αα/αα	No mutation	Poly+861 (A-G)		
324	M	WHITE: European	?delta beta	αα/αα	No mutation	No mutation	No mutation	
325	F	WHITE: European	?α-thal/Fe deficiency	αα/αα	No mutation	Poly+861 (A-G)		
326	F	WHITE: Irish	?Hb variant 22.3%	αα/αα	n/a	Hb J-Toronto		
327	F	WHITE: Irish	?Hb variant (18.8%)	αα/αα	No mutation	Codon 127 (A-G) variant		
328	F	WHITE:N.European	?Hb variant 25.9%	αα/αα	No mutation	No mutation		

***HBA* Sequencing key:**

n/a denotes neither *HBA1* or *HBA2* sequencing was performed

Black text: denotes changes found within *HBA2*

Red text: denotes changes found within *HBA1*

n/a: denotes *HBA2* not amplififed to due to homo/cmp htz for $-\alpha^{+}$ deletion