1- Introduction and Literature Review

1.1. Historical background:

From a historical point of view, Beja lands span the region between the Nile Valley and the Red Sea shore. Both are areas that have brought more interest to them than the Beja area itself, being located in outskirts from both view points. We have to consider the conditions first to the west, then to the east of the latter (Dahl and Orans, 2006). Today, the name Beja is applied to the people occupied the region known as Eastern Sudan or Eastern district which is boarded by Eretria and Ethiopia from the west, Egypt from the north, the Red Sea from the east, and the Nile and the Central district from the west. They all speak a shared language known as "Tibdaawyi". From a semi stable traditional point that is based on ethnicity, Beja is divided into ten tribes from the north to the south Bishariyoon, Amaraar, Artaiga, Ashraf, Kumailab, Hadandawa, Melhaitkenab, Halanga, Banu Amer and Habab. Each of these major tribes includes many sub-tribes. The Hadandawa, for example include the Sharaab and the Fadlab among many others (Ohaj, 1986). The Hadandawa also include the Dignab from which the Prince Osman Digna was originated (Dirar, 1975). Very roughly around 1000 BC a group of people, referred to in the archeological texts as the "C-group", migrated from lower Nubia (the area between present-day Aswan and Wadi Halfa) and settled in upper Nubia (the Nile Valley north of Dongola in Sudan) where they developed the Kingdom of Napta from about 750 BC (Dahl and Orans, 2006). To the Beja, however, they are commonly known as the Bellou by the reason of the fact that on arrival they spoke a strange tongue, the Beja for which (as it still is) was Bellaweit, though it was not until they had been driven south from the Atbai in the 15th century that the later name replaced the former in common usage. At some time in the 9th century began a slow and by no means a continuous process of infiltration by Arab tribes, never at any time of great proportions, yet sufficiently pronounced as to result in all the Beja peoples becoming Moslem (Paul, 1953). The account given by the Arab invaders about their relations with the Beja are the best sources for Beja history available. Modern technological change and systems of colonialism and capitalist organization have meant a new distinct transformation of the contextual constraints of Bedawiet-speaking pastoralists roaming the eastern deserts

of Sudan were not untouched by the political and economical structures of faraway places before the modern era: in particular, they must have been affected by those regional trends that regulated the trade in the Nile Valley on one hand and the orient through the Red Sea or via the Baghdad route (Dahl and Orans, 2006).

1.2. The Hemoglobin molecule:

Human hemoglobins are formed from two pairs of globin chains each with a haem group attached. Seven different globin chains are synthesized in normal subjects; four are transient embryonic hemoglobins referred to as Hb Gower 1, Hb Gower 2, Hb Portland 1, and Hb Portland 2. Hb-F is the predominant hemoglobin of fetal life and comprises the major proportion of hemoglobin found at birth. Hb-A is the major hemoglobin found in adults and children. Hb-A_2 and Hb-F are found in small quantities in adult life (approximately 2-3.3% and 0.2-1.0% respectively). The adult proportion of Hb-A, A_2 and F are usually attained by 6-12 months of age (Bain *et al*, 2012). Hemoglobin is a conjugated protein of molecular weight 64000 (Firkin *et al*, 1989). Each of the different hemoglobins synthesized in the embryo, fetus and adult is adapted to their particular oxygen requirements (Hoffbrand *et al*, 2005). The individual chains synthesized in postnatal life are designated α, β, γ, and δ. Hb-A has two α chains and two β chains ($\alpha_2\beta_2$); Hb-F has two α chains and two γ chains ($\alpha_2\gamma_2$), and Hb-A_2 has two α chains and two δ chains ($\alpha_2\delta_2$). The α chains is thus common to all types of hemoglobin molecules. α chain synthesis is directed by two α genes, α1 and α2, on chromosome 11. γ chain synthesis is directed by two genes, $^{G}\gamma$ and $^{A}\gamma$, also on chromosome 11. The four chains are associated in the form of a tetramer: The $\alpha_1\beta_1$ (and equivalent $\alpha_2\beta_2$) contact is the strongest and involves many amino acids with many interlocking side chains; the $\alpha_1\beta_2$ (and equivalent $\alpha_2\beta_1$) contact is less extensive, and the contacts between similar chains are relatively weak. The binding of haem group into the haem pocket in each chain is vital for the oxygen carrying capacity of the molecule and stabilizes the whole molecule. If the haem attachment is weakened, the globin chains dissociates into dimmers and monomers. There are many naturally occurring, genetically harmless, some have serious clinical effects. Collectively, the clinical syndromes resulting from disorders of hemoglobin synthesis are referred to as "haemoglobinopathies". They can be grouped into three main categories;

Those owing to structural variants of hemoglobin such as Hb-S, Those owing to failure to synthesize one or more of the globin chains of hemoglobin at a normal rate as in thalassaemia, and, those owing to failure to complete the normal neonatal switch from fetal hemoglobin (Hb-F) to adult hemoglobin (Hb-A). Those comprise a group of disorders referred to as Hereditary Persistence of Fetal Hemoglobin (HPFH). An individual can also have a combination of more than one of these abnormalities (Bain *et al*, 2012).

1.3. Genetics of haemoglobin synthesis:

The HBB gene that codes for the globin chains of haemoglobin is located on the short (p) arm of chromosome 11 at position 15.5.

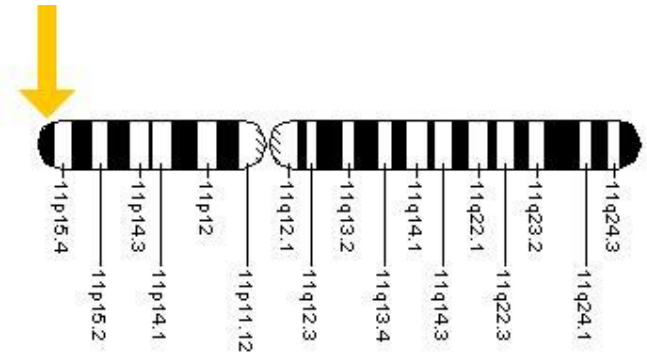


Figure 1-1: The Hemoglobin beta (HBB) gene is indicated by the arrow at position 15.5 (Adapted from Parker and Parker , 2007).

All the human globin genes have three coding sequences (exons) and two intervening non-coding sequences (intervening sequences or introns) and are flanked by 5'and 3' non-coding sequences (referred to as untranslated regions, UTRs). The two α genes differ in structure in intron 2 and the 3' UTR, but the coding sequences are identical. As for all genes, coding is by means of triplets of nucleotides, known as codons, which code for a specific amino acid. 5' to each gene is the promoter, a sequence that binds RNA polymerase and transcription factors and is necessary for the initiation of transcription. Globin gene promoters share several conserved DNA sequences that bind crucial transcription factors. Transcription is the process by which RNA is synthesized from a DNA template by the action of RNA polymerase. The entire globin gene, including the introns and the 5' and 3' UTRs, is transcribed. Transcription is controlled by interaction between the genes and transcription factors that bind to both promoters and upstream regulatory elements, referred to as the β-locus control region (β-LCR) for the β cluster and HS –40 for the α cluster. The

β-LCR includes four erythroid-specific DNase sites, designated HS1, HS2, HS3 and HS4, of which HS3 is probably the most important in opening the chromatin structure to permit access of transcription factors and HS2 is probably the most important in enhancing globin chain synthesis. There are also enhancers within introns of genes and downstream of the β and α genes. *Trans*-acting factors, encoded by genes on chromosomes other than 11 and 16, are vital for the expression of globin genes. Relatively erythroid-specific *trans*-activating factors, including GATA1, NFE2, EKLF, SSP, Nrf-1, Nrf-2 and LCR-F1, contribute to the regulation of gene expression by interacting either with the LCRs or with the globin gene promoters to increase gene expression. EKLF (erythroid Kruppel-like factor) is an enhancer of β chain synthesis and SSP (stage selector protein) is an enhancer of δ and γ chain synthesis (Bain, 2006).

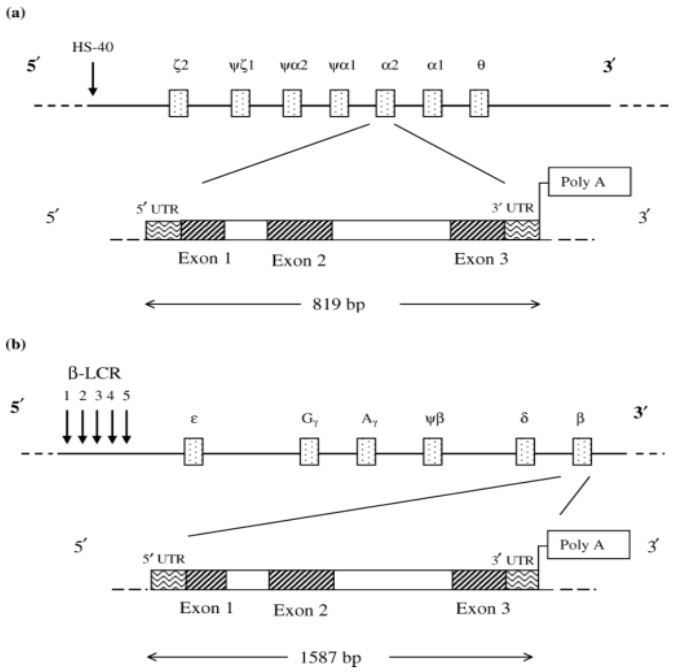


Figure 1-2: The α- and β-like globin gene clusters (With kind permission from professor: Swee Lay Thein).

Figure 1-2 illustrates the Alpha and Beta globin gene clusters with each globin gene having three exons and two introns. (a) The alpha locus on chromosome 16 p13.3. Of the seven genes shown only three genes are expressed at a clinically significant level, namely the ζ2 gene (expressed early in foetal development) and the α1 and α2 genes (the focus of most haemoglobinopathy tests). The other genes shown are pseudogenes (ψζ1, ψα1 and ψα2) or are not expressed at a significant level (θ). The vertical arrow indicates the location of the upstream hypersensitive site (HS-40), important for locus gene expression. The α2 gene structure is shown below indicating the three coding exons (striped boxes), the two introns (open boxes), and the untranslated regions (zigzag boxes) common to all globin genes. (b) The beta locus on chromosome 11 p15.4 with the ε, $^{G}\gamma$ and $^{A}\gamma$, δ and β genes, arranged in the order of their developmental expression. During development two switches are made in beta like globin expression. The first occurs early in foetal development from ε globin to the γ globin genes (both $^{G}\gamma$ and $^{A}\gamma$) at 6–8 weeks gestation, and the second switch occurs shortly prior to birth, from the γ genes to the δ and β genes. Again ψβ denotes a nonexpressed pseudogene. Upstream of the beta globin cluster is the β locus control region (βLCR) comprised of the five hypersensitive sites indicated by vertical arrows. As in panel A, the β globin gene is expanded to show the gene structure covering 1.6 kb.

This is bigger than the alpha genes because of a larger second intron. In addition to transcription factors that are relatively specific to erythroid cells, globin gene expression is also influenced by general transcription factors, including AP-1, Sp1, YY1, USF and TAL-1/SCL. Nascent RNA molecules resulting from transcription are large and unstable and are modified in the nucleus. Initially, the 5' end acquires a 7-methylguanosine cap (CAP), which is probably added during transcription and has a role during translation; during this 'capping' process, methylation of adjacent ribose residues also occurs. Following this, the majority of transcripts acquire a 3' polyadenosine tail with the addition of 75 to several hundred adenylate residues. There is an AAUAA sequence near the 3' end (within the 3' UTR) that serves as a signal for 3' cleaving of the transcript and polyadenylation. Polyadenylation may have a role in transfer of the mRNA from the nucleus to the cytoplasm. The polyadenylate tail is also important for mRNA stability and enhances translation. Finally, the introns are excised to give a functional mRNA molecule which, in most cases, contains a single continuous open reading frame (ORF), encoding the sequence of the relevant protein, flanked by 5' and 3' UTRs. Molecules of mRNA move from the nucleus to the cytoplasm where they bind to ribosomes and serve as templates for the assembly of the polypeptide sequences of the globin chain. Each nucleotide triplet serves as a template for a specific amino acid that is covalently bound to, and transported to the ribosome by, transfer RNA (tRNA). tRNAs are specific for both a nucleotide triplet and an amino acid. Amino acids are thus assembled in the correct sequence, forming a polypeptide. This process is known as translation. An initiation codon, AUG, is essential for the initiation of translation; it is the first codon after the 5' UTR and encodes methionine. Initiation requires the amino acid methionine, tRNA specific for methionine, guanosine triphosphate (GTP) and an initiation factor. When the nascent molecule reaches 20–30 amino acid residues, the methionine is removed through the action of methionine aminopeptidase. When the chain reaches 40–50 residues, cotranslational acetylation of the N-terminal residue can occur through the action of several acetyl transferases. Whether this occurs to any great extent depends on the nature of the N-terminal residue. Thus the glycine of the γ chain is 10–15% acetylated, whereas the valine of normal α, β and δ chains is resistant to acetylation. There are 64 possible nucleotide triplets or codons, 61 of which encode amino acids (20 in all) and three of which do not; the latter serve as stop or termination codons, leading to termination of globin chain synthesis. Transcription thus continues until a termination codon, UAA, UAG or UGA, is encountered. The termination codon is followed by the 3' UTR. The rate-limiting step of globin chain translation is the commencement of elongation, i.e. the next step after

initiation. Transcription from the two α genes is equal up to the eighth week of gestation, but thereafter the α2 gene becomes dominant and, in adult life, the ratio of α2 to α1 mRNA is 2.6–2.8:1. The translational efficiency differs somewhat so that the α2 gene directs the synthesis of about twice as much α chain as the α1 gene. There is more α than β mRNA, probably about 2.5 times as much, but β chain synthesis is more translationally efficient than α chain synthesis and α chains are therefore produced only slightly in excess of β chains. The control of globin chain synthesis is probably mainly at the level of transcription, with translational control being less important. Translation is dependent on the presence of haem. In iron deficiency, the reduced availability of haem leads to inactivation of the initiation factor and thus reduced synthesis of globin chains. The α and β globin chains are synthesized on different polyribosomes. The combination of a free α chain with a β chain that is still attached to the polyribosome, to form an αβ dimer, may contribute to the release of the β chain from the ribosome. Incorporation of haem probably occurs after release from the polyribosome. Globin mRNA is unusually stable so that translation can continue for up to 3 days after cessation of transcription. Both the α and β globin genes have structural determinants in their 3' UTRs that are important for mRNA stability (Bain, 2006).

1.4. The Hemoglobin types:

Hemoglobin A (Hb-A) comprises about 97% of the hemoglobin of adult red cells. Hemoglobin A_2 (Hb-A_2) is the minor hemoglobin in the adult red cells. Hb-A_2 is present in very small amounts at birth and reaches the adult level of 1.5-3.2% during the first year of life. Hemoglobin F (Hb-F) is the major respiratory pigment from early intra-uterine life up to term. At term, Hb-F accounts for 70-90% of the total hemoglobin. It then falls rapidly to 25% at one month and 5% at 6 months. The adult level of about 1% is not reached in some children until puberty (Firkin *et al*, 1989).

Table 1-1: Normal human hemoglobins (Modified from Firkin *et al*, 1989):

	Hemoglobin	Structural formula
Adult	Hb-A	$\alpha_2\beta_2$
	Hb-A_2	$\alpha_2\delta_2$
Fetal	Hb-F	$\alpha_2\gamma_2$
	Hb-Bart's	γ_4
Embryonic	Hb-Gower1	$\zeta_2\varepsilon_2$
	Hb-Gower2	$\alpha_2\varepsilon_2$
	Hb-Portland	$\zeta_2\gamma_2$

1.5. Structural variants of hemoglobin:

The inherited diseases of hemoglobin are the commonest single gene disorders; the world Health Organization estimates that about 7% of the world populations are carriers (Hoffbrand *et al*, 2005). Alterations in the structure of hemoglobin are usually brought about by point mutations affecting one or, in some cases, coding for amino acids of globin chains. An example of such a point mutation is Hb-S caused by the substitution of valine for glutamic acid in position 6 of the β globin chain ($\beta^{6Glu \rightarrow Val}$). Less commonly, structural change is caused by shortening or lengthening of the globin chain. For example, five amino acids are deleted in the β chain of Hb-Gun Hill, whereas in Hb-Constant Spring 31 amino acids are added to the α chain. Mutations associated with a frame shift can also lead to synthesis of structurally abnormal hemoglobins, which may be either shorter or longer than normal. There may also be combinations of segments of β and γ or δ chains resulting in hybrid hemoglobins; the β and δ combinations are known as the Lepore and anti-Lepore haemoglobins respectively. Some structural variants are associated with severe clinical phenotypes in the homozygous or even heterozygous state; these mutations affect the physical or chemical properties of the hemoglobin molecule, resulting in changes in hemoglobin solubility, stability, or oxygen-binding properties. Some of these variants separate on electrophoresis or chromatography, whereas some do not (Bain *et al*, 2012).

Table 1-2: Types of hemoglobins (Adapted from Firkin *et al*, 1989):

	Hemoglobin	Structure	Comment
Normal	A	$\alpha_2\beta_2$	97% of adult hemoglobin.
	A_{1c}	$\alpha_2\beta_2$	≤ 5% of Hb-A (glycosylated Hb).
	A_2	$\alpha_2\delta_2$	2% of adult Hb, elevated in β-thalassaemia.
	F	$\alpha_2\gamma_2$	Normal Hb in fetus from 3rd to 9th month; increased in β-thalassaemia.
Abnormal chain production	H	β_4	Found in α-thalassaemia.
	Bart's	γ_4	Found in homozygous α-thalassaemia. Biologically useless.
Abnormal chain structure	S	$\alpha_2\beta_2$	Substitution of valine for glutamic acid in position 6 of the β chain.

1.5.1. Genetic regulation of haemoglobin variants:

HBB gene is located on the short (p) arm of chromosome 11 at position 15.5, base pairs 5,246,695 to 5,248,300. More than 250 mutations in the HBB gene have been found to cause β-thalassaemia. Most of the mutations involve a change in a single DNA building block within or near the HBB gene (Blane *et al,* 2013). Abnormal hemoglobins are inherited as autosomal co-dominants. Thus, subjects who inherit one normal and one abnormal gene are heterozygotes, and those who have two identical abnormal genes are homozygotes. Double heterozygotes are subjects who have inherited two different abnormal genes. The homozygous state usually referred to as the 'disease', and the heterozygous state as the 'trait' (Firkin *et al,* 1989).

Table 1-3: The common abnormal hemoglobins (Firkin *et al,* 1989):

Hemoglobin	Structural formula
Hb-S	$\alpha_2\beta_2^{6Glu\rightarrow Val}$
Hb-C	$\alpha_2\beta_2^{6Glu\rightarrow Lys}$
Hb-E	$\alpha_2\beta_2^{26Glu\rightarrow Lys}$
Hb-D^{Punjab}	$\alpha_2\beta_2^{121Glu\rightarrow Gln}$

1.5.2. Hemoglobins with reduced solubility:

By far the most common hemoglobin variant in this group is the sickle cell hemoglobin or Hb-S. Hb-S has poor solubility in the deoxygenated state and can polymerize within the red cells.

1.5.2.1. Sickle cell disease (SCD):

Sickle cell disease (also called disorder) is a collective name for a group of conditions causing clinical symptoms that are characterized by the formation of sickle red cells. It is common in people originating from Africa, but it is also found in considerable numbers of people of India, Arabic, and Greek descent (Abbas, 2014). Sickle cell anemia is caused by an HBB mutation that produces an abnormal version of the hemoglobin beta chain called hemoglobin S or (HbS). HbS results from a switch in building blocks (amino acids) used to make the beta chain. Specifically, the amino acid glutamic acid is replaced with the amino acid valine at position 6 in the beta chain. The shorthand for this amino acid switch is Glu6Val. Replacing glutamic acid with valine causes the abnormal HbS chains to stick together and form long, rigid molecules. The rigid HbS

molecules bend red blood cells into a sickle (crescent) shape. The sickle-shaped cells die prematurely, which can lead to a shortage of red blood cells (anemia). The sickle-shaped cells can also block small blood vessels, causing pain and organ damage (Parker and Parker, 2007). The homozygous state or sickle cell anemia (β genotype SS) causes moderate to severe hemolytic anemia. Sickle cell trait (β genotype AS), the heterozygous state, is very common, affecting millions of people worldwide. In vivo sickling occurs only at very high attitudes and at low oxygen pressures, spontaneous haematurea, owing to sickling in the renal papillae, is found in about 1% of people with sickle cell trait (Bain *et al*, 2012). In Africa, SCD is the most common inherited hematological disease with high mortality rate at age one to five years. This disease was discovered early in Sudan, the peak occurrence of SCD is among the population from the Western Sudan. The highest prevalence of SCD in Sudanese is among the population from the Western Sudan. It is believed that the sickle cell gene has brought to Sudan through immigrants from West African tribes, especially from Howsa, Folani and Bargo. (Abbas, 2014).

Figure 1-3: The map of the Sudan, the red circles show the distribution of the sickle gene. (With kind permission from Abbas, 2014).

Sickle cell disease (SCD) is one of the most common inherited disorders of haemoglobin in Africa and it is expected that sickle cell trait varies in frequency in different areas in Sudan. An extensive literature search was carried out accessing the US National Library of Medicine, the WHO Eastern Mediterranean Region resources, the Catalogue for Transmission Genetics in Arabs and papers and documents published in Sudan that included data on the prevalence of sickle cell anaemia and trait. Rates of SCA and trait varied in different areas in Sudan with the highest rates reported from Western and Eastern Sudan where one in every 123 children born in Messeryia tribe in Western Sudan is at risk of having SCD. High consanguinity rates and malaria endemicity are strong related factors with sickle cell gene in Sudan (Sabahelzain and Hamamy, 2014). Sickle cell disease (SCD) in children with a history of repeated hospitalization is distressing for children as well as their parents leading to anxiety and has negative effects on the psychological state of children and their families. Children with SCD who had history of repeated hospitalization are at an increased risk of developing behavioral problems. Psychological counseling, social support, and proper pain management could minimize these behavioral consequences (Bakri *et al*, 2015).

1.5.2.1.1. Genetic bases of haemoglobin S:

The sickle cell mutation appears to have arisen spontaneously at least five times in the history of mankind. Such independent mutations can be recognized by their association with different β globin gene haplotypes, demonstrated by the analysis of restriction fragment length polymorphisms (RFLPs). There are three foci of haemoglobin S in Africa, associated with different haplotypes, the haplotypes being defined by RFLP analysis. They are in Senegal (Senegal type), the Central African Republic and southern Africa (Bantu or Central African Republic type) and Benin, Central, West and North Africa (Benin type) (See figure 1-4 page 13). The Benin type has also spread to Spain, Portugal, Sicily (perhaps from Greece, perhaps from Sudanese soldiers in Arab armies) and southern mainland Italy, Greece (particularly Macedonia), Albania, Turkey, north-western Saudi Arabia and Oman. The Bantu type has spread to Kenya, Zambia and the Sudan. In addition to the three major foci, there may have been a further independent mutation amongst the Eton people in southern Cameroon. A fifth mutation is associated with

further foci in eastern Saudi Arabia and in extensive areas of central and southern India, particularly amongst the scheduled tribes (a group living outside the caste system). It may have arisen initially in the Indus valley. The prevalence of the haemoglobin S gene is up to 25% in eastern Saudi Arabia and as high as 30% in some tribal populations in central India. The Indian/Saudi Arabian haplotype has also been found in Afghanistan, Oman, Kuwait, Bahrain and Iran and amongst Bedouin Arabs in Israel. Migration from Africa has led to the sickle cell gene occurring also in Central and South America, in Afro-Americans and in Afro-Caribbeans in Canada, the UK and other European countries. There is a high prevalence in some populations in Mexico, Colombia, Venezuela, Guyana, Surinam, French, Brazil and Peru. All three major African haplotypes are represented in the USA, the Caribbean and the UK. The sickle cell gene is also found in Madagascar, Mauritius (both Bantu and Arab–Indian haplotypes), Abu Dhabi, United Arab Emirates, Lebanon, Iraq, the southern part of the former USSR and amongst North African Arabs (Bain, 2006). SCD is believed to be the most frequent inherited blood disorder on the globe affecting an estimated 100 million people world-wide and, in particular, the black races and persons of Mediterranean origin. In Africa, SCD is the most prevalent genetic disease with high mortality rate at age one to five years. In the United States of America (USA), sickle cell anemia has been found to be the most frequent autosomal recessive gene disorder affecting approximately 1:375 persons of African ancestry (Abbas, 2014). The wide geographical spread of this potentially deleterious gene has been attributed to the protection of heterozygotes from premature death from falciparum malaria. In areas in which malaria is endemic, the β-S and β-A genes may exist as a balanced polymorphism, i.e. death or serious disability from sickle cell anaemia before the age of reproduction is balanced by a decreased death rate from malaria amongst heterozygotes. A second mutation has occasionally occurred in a β-S gene leading to the synthesis of variant haemoglobin with two amino acid substitutions. Such variant haemoglobins retain their ability to sickle, but may have a different electrophoretic mobility from haemoglobin S. In heterozygous, homozygous and compound heterozygous states, these mutations have a similar, although not necessarily identical, significance to haemoglobin S. At least 10 such double mutations are known. The most common,

haemoglobin C-Harlem (initially described under the name of haemoglobin C-Georgetown), $\alpha 2\beta 2^{6Glu \rightarrow Val,\ 73\ Asp \rightarrow Asn}$, is less prone to polymerization than haemoglobin S itself. The rare double substitution haemoglobin, haemoglobin S-Antilles, is even more prone to sickle than haemoglobin S itself, as is haemoglobin Jamaica Plain. There are other haemoglobins unrelated to haemoglobin S that can polymerize *in vitro*, e.g. haemoglobin I ($\alpha 16^{Lys \rightarrow Glu}$) and haemoglobin Setif ($\alpha 22^{Asp \rightarrow Tyr}$). Although they are not associated with any relevant clinical abnormality, haemoglobin Setif can cause a false positive sickle solubility test. Homozygosity for haemoglobin S ($\beta^S\beta^S$) causes a serious condition referred to as 'sickle cell anaemia'. Heterozygosity for haemoglobin S ($\beta\beta^S$), referred to as sickle cell trait, is usually asymptomatic. The β^S gene may also be co-inherited with another β chain variant. When there is deleterious interaction between the sickle cell haemoglobin and the second variant haemoglobins, as is the case, for example, with haemoglobin C and haemoglobin D-Punjab, a clinically significant sickling disorder occurs. Subjects who are heterozygous for β thalassaemia and haemoglobin S likewise suffer from the clinico-pathological effects of sickle cell formation and consequent vascular occlusion. The term 'sickle cell disease' is often used as a generic term to include sickle cell anaemia and other conditions in which a clinically significant disorder results from sickle cell formation and the associated pathological processes (Bain, 2006).

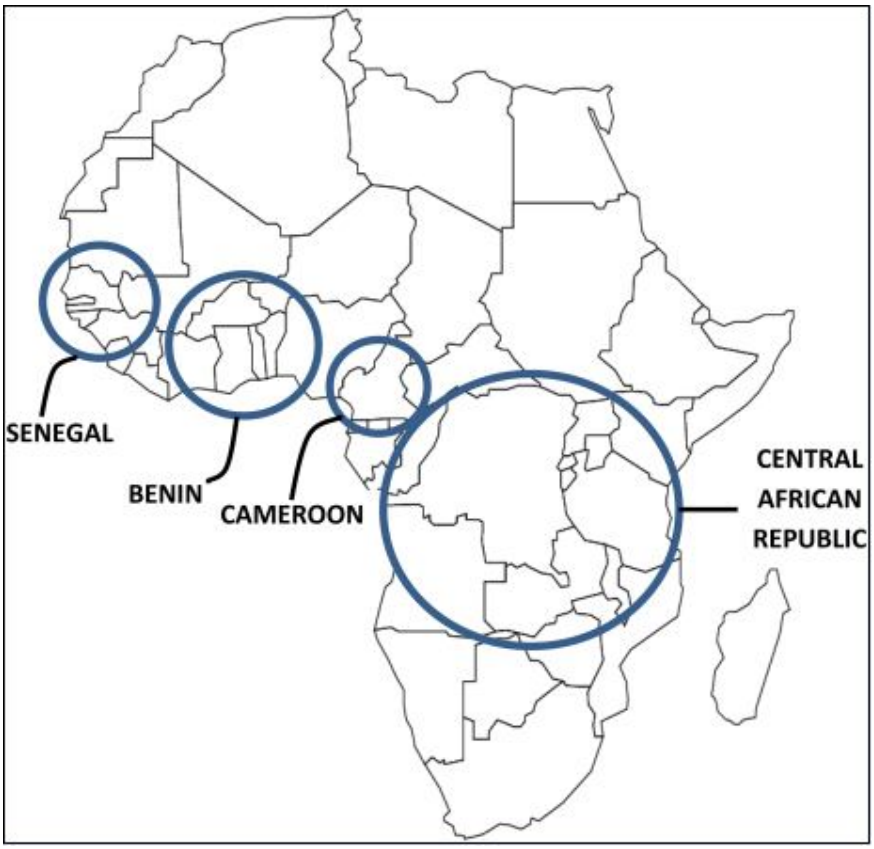


Figure 1-4: Multiple origins of the beta globin Glu6Val (β^s, sickle) mutation within the African continent. Malaria selection pressure led to the emergence of the sickle mutation at four distinct locations in sub-Saharan Africa, which are distinguishable by flanking DNA polymorphisms that define beta globin SCA haplotypes from Senegal (SEN), Benin (BEN), Cameroon (CAM), and Central African Republic (CAR). An additional sickle mutation occurred in the Arabian-Indian region, not shown. (Adapted from Ware, 2013).

1.5.2.1.2. Should SCA be considered a neglected tropical disease (NTD)?

An NTD can be operationally defined as any serious medical affliction with the following characteristics: (1) worldwide distribution, typically affecting millions; (2) highest burden among the most impoverished and disadvantaged populations; (3) serious morbidity, diminished quality of life, and even mortality; (4) co-morbidity for other life-threatening diseases; (5) relatively simple diagnostic testing; and (6) inexpensive treatment options. Although NTDs are usually infectious and communicable, perhaps those should not be strict requirements. Indeed, based on these criteria, SCA is long overdue for membership in this elite but tragic club of global medical maladies.

Based on these six criteria, there are compelling reasons to consider SCA an NTD. First, SCA has a worldwide distribution in Africa, Asia, and the Americas, with over 300,000 births occurring annually. Second, the highest burden of SCA falls on the poorest and most disadvantaged populations in sub-Saharan Africa and tribal Indian populations. Third, SCA causes serious morbidity and contributes substantially to under-5 mortality rates. Fourth, SCA is an important co-morbid factor for other life-threatening diseases, especially malaria and invasive pneumococcal disease. Fifth, rapid and inexpensive testing at birth or in the neonatal period can accurately determine the presence of SCA. Sixth, simple interventions with penicillin prophylaxis, pneumococcal immunizations, and education can be life-saving and lead to improved survival. Most importantly, however, this life-threatening hematological disease has been long overlooked by almost all medical organizations and governments, and is only now receiving modest interest and attention as a substantial global healthcare issue. With better awareness and funding, SCA can change from an invisible killer of children to a neglected one, which will help promote the need for life-saving diagnostic and treatment efforts to reach affected infants and children. We hematologists should learn from NTD specialists and experts to develop strategies and policies that integrate testing and treatment, which will allow improved survival and quality of life for disadvantaged children with SCA wherever they live (Ware , 2013).

1.5.2.1.3. Other forms of sickle cell disease:

Sickle cell/hemoglobin C disease is a compound heterozygous state for Hbs S and C. This state usually results in a milder form of sickle cell disease. *Sickle cell β thalassaemia* arises as a result of inheritance of one Hb-S and one β thalassaemia gene. Africans and Afro-Caribbeans with this condition are often heterogeneous for a mild β^{+}-thalassaemia allele resulting in the production of about 2% of Hb-A. This gives rise to a mild sickling disorder. Inheritance of Hb-S and β°-thalassaemia trait is associated with severe sickle cell disease. Interaction of Hb-S with Hb-D^{Punjab} ($Hb^{Los\ Angeles}$) or with Hb-O^{Arab} gives rise to severe sickle cell disease. (Bain *et al*, 2012). In blacks, the condition is milder, and red cells contain 10-30% Hb-A, many patients have little disability and may be detected by a chance hematological screening examination. Some have associated painful crises, but lead to a normal life otherwise (Firkin *et al*,

1989). Different variants of sickle cell haemoglobinopathies were identified; AS, AS/C, S/ß Thalassaemia, S/C, S/D and SS patterns were reported. Haemoglobin A2 have no significant difference in patients with sickle cell disease, while Hb F and Hb S show significant elevation (Ajjack *et al*, 2014).

Table 1-4: Concentration of haemoglobin types in different sickle cell conditions:

Disorder	Hemoglobin%			
	A	**A_2**	**F**	**S**
Sickle cell trait (AS)	55-70	2-4	N	38-45
Homozygous sickle cell disease (SS)	0	2-5	1-20	75-95
Sickle cell β-thalassaemia:				
S-β^+	30-10	4-8	2-10	60-85
S-β°	0	4-8	5-30	70-90
S-HPFH	0	N	18-30	60-90
Sickle cell Hb-C disease (SC)	0	35-50 (C+A_2)	1-5	50-65
Sickle cell Hb-D disease (SD)	0	N	1-5	95(S+D)
Sickle cell trait α-thalassaemia trait	65-75	N	N	20-30

*Hb-C cannot be separated from HbA_2, and Hb-D cannot be separated from Hb-S, on routine cellulose acetate electrophoresis. The electrophoretic phenotypes associated with other abnormal hemoglobins follow the same general pattern as detailed for Hb-S. The above data are taken from de Gruchy, 2011. N=Normal; HPFH=Hereditary persistence of fetal hemoglobin.

1.5.2.1.4. Other sickling hemoglobins:

In addition to Hb-S, there are nine hemoglobins (Hb-$S^{Antilles}$, Hb-$C^{Ziquinchor}$, Hb-C^{Harlem}, Hb-$S^{Providence}$, Hb-S^{Oman}, Hb-$S^{Trravis}$, Hb-$S^{South\ End}$, Hb-$S^{Jamica\ Plain}$, Hb-$S^{Cameroon}$) that have both the β 6 glutamic acid to valine mutation and an additional single point mutation in the β globin chain. They have clinical significance similar, but not necessarily identical, to Hb-S: for example, Hb-$S^{Antilles}$ is associated with an even greater propensity to sickling than Hb-S (Bain *et al*, 2012).

1.5.3. Hemoglobin C (Hb-C):

Hb-C is the most common structural hemoglobin variant in people of African descent (Lewis S. M. *et al*, 2006). It appears to have originated in West Africa where it affects 20% of the population in some areas (Firkin *et al*, 1989). The substitution of glutamic acid in position 6 by lysine

results in a hemoglobin molecule with a highly positive charge, decreased solubility, and tendency to crystallize. Homozygotes may have mild anemia with numerous target cells and irregularly contracted cells. Interactions with β° and β^{+}-thalassaemia trait results in mild or moderate hemolytic anemia (Bain *et al*, 2012). Hb-C is less soluble than Hb-A, and if present in sufficient amounts, tends to form crystals within the red cell. Hb-C is a slow-moving hemoglobin on cellulose acetate electrophoresis, migrating in the same position as Hb-E and Hb-A_2. Hb-C trait (heterozygous state) is generally asymptomatic; in the blood film the presence of numerous target cells is characteristic, but in occasional cases they may not be a marked feature. Red cell life span is normal. Hb-C disease (homozygous state) is usually a benign illness characterized by compensated hemolysis and a normal hemoglobin level or a mild to moderate anemia. Hb-C β-thalassaemia has been described in American blacks and more recently in Italians, Africans and Turkish people (Firkin *et al*, 1989).

1.5.3.1. Genetic consideration of Hb-C:

Haemoglobin C is a variant haemoglobin with a mutation in the β globin gene at the same site as the mutation in the sickle cell haemoglobin. It was first recognized by Itano and Neel in 1950. Its structure is $\alpha 2\beta 2^{6Glu \rightarrow Lys}$. It may be present in the heterozygous state (haemoglobin C trait), in the homozygous state (haemoglobin C disease) and in a variety of compound heterozygous states, such as sickle cell/haemoglobin C disease and haemoglobin C/β thalassaemia. Haemoglobin C is thought to have originated in West Africa, west of the Niger River. In northern Ghana, the proportion of individuals with haemoglobin C is as high as 40% and, in northern Cote d'Ivoire (Ivory Coast), up to 50%. In Burkina Faso (previously Upper Volta), it is 15–40%. It would appear that haemoglobin C arose in the region spanning Burkina Faso and the north of Cote d'Ivoire and Ghana. It is found in individuals of African descent in the Caribbean (3.5% prevalence), USA (2% prevalence amongst US blacks), Canada and the UK. A high incidence has been noted in a Bedouin tribe in northern Israel. There is also a significant incidence of haemoglobin C in North Africa and southern Europe. However, it should be noted that some early reports of the presence of haemoglobin C based only on electrophoresis at alkaline pH may have been a misidentification of haemoglobin O-Arab as haemoglobin C. Haemoglobin C appears to have

had an independent origin in Oman and Thailand. Haemoglobin C appears to protect against severe falciparum malaria, with homozygosity giving very high protection and heterozygosity giving moderate protection. Oxyhaemoglobin C is prone to crystallization, but crystals dissolve on deoxygenation, so that obstruction of capillaries by cells containing crystals is not likely. In comparison with haemoglobin A, crystallization is inhibited by haemoglobins F and A2 (Bain, 2006).

1.5.3.2. Haemoglobin C trait:

Haemoglobin C trait describes the heterozygous condition in which there is one normal β gene and one β^C gene. It is of no clinical significance, but is of importance in counseling prospective parents. This is largely because of the possibility of sickle cell/haemoglobin C disease if one parent has haemoglobin C trait and the other has sickle cell trait (Firkin *et al*, 1989).

1.5.3.3. Haemoglobin C disease:

Haemoglobin C disease describes the homozygous state in which there are two β^C genes and no normal β gene. As a consequence, about 95% of total haemoglobin is haemoglobin C, with the remainder being haemoglobins A2 and F. Homozygosity for haemoglobin C leads to a clinically mild, chronic, haemolytic anaemia (Firkin *et al*, 1989).

Table 1-5: Prevalence (%) of Haemoglobins S and C in different populations (modified from Bain, 2006):

Country or people	Haemoglobin S	Haemoglobin C
West Africa:		
Senegal	3–15	<1–6
Gambia	6–28	<1–2
Guinea Bissau	<1–25	<1–1.5
Guinea	13–33	
Sierra Leone	22–30	
Liberia	<1–29	1–3
Ivory Coast	2–26	<1–50
Mali	5–17	
Burkina Faso	2–34	15–40
Ghana	3–25	8–40
Togo	6–28	7–17
Benin	5–31	7–27
Niger	5–23	1–8
Nigeria	10–41	<1–9
Central Africa:		
Gabon	8–32	
Cameroon	<1–31	<1
Central African Republic	2–24	
The Republic of the Congo	7–32	
Democratic Republic of the Congo	1–46	
East Africa:		
Kenya	<1–34	
Uganda	1–39	
Tanzania	1–38	
Rwanda		
Tutsi	<1–5	
Hutu	5–15	
Burundi	1.5–26	
Southern Africa:		
Angola	8–40	
Zambia	<1–30	
Zimbabwe	<1–11	
Malawi	3–18	
Mozambique	<1–40	
Madagascar	<1–23	
Botswana	<1	

Table 1-5: *Continued:*

Country or people	Haemoglobin S	Haemoglobin C
Namibia	0–15	
South Africa		
Bantu	<1–4	
Indian	2–10	
Cape Coloured	<1	<1
North Africa:		
Morocco	<1–7	<1–6
Algeria	<1–15	<1–13
Tunisia	<1–2	
Libya	<1–70	
Egypt	<1	
Sudan	<1–17	
Horn of Africa:		
Ethiopia	0–1	
Djibouti	0	
Somalia	0	
Afro-Americans	6–15	1–3.5
Afro-Caribbeans:		
Jamaica	3.5–12	2–4
Bahamas	1-4	3
Barbados	4	3–5
Cuba	0–23	0–2.5
Haiti	7–17	1–3
Dominican Republic	6–12	3
Puerto Rica	<1–8	<1–2
Lesser Antilles	1–14	1–4.5
Guadeloupe	4.4	
Middle East:		
Turkey	<1–34	
Syria	<5–25	
Lebanon	<1	
Jordan	4–6	
Iraq	0–25	
Iran	0	
Saudi Arabia	<1–36	
Kuwait	2	
Bahrain	2.5	
Oman	5	Rare

Table 1-5: *Continued:*

Country or people	Haemoglobin S	Haemoglobin C
Yemen	1–2	
Abu Dhabi	2	
Abu Dhabi	2	
Asia:		
India	0–35	
Pakistan	0.5–1	
Sri Lanka		Rare
Thailand		Rare

1.5.4. Hemoglobin E (Hb-E):

Hb-E is found predominantly in south-east Asia, India, Burma, and Sri Lanka and amongst immigrant Indo-Chinese population in western countries. Hb-E arises from the substitution of lysine to glutamic acid in the 26^{th} position of the β chain. On cellulose acetate electrophoresis, Hb-E is slow moving and migrates in the same position as Hb-C and Hb-A_2. Agar gel electrophoresis permits differentiation, as Hb-E does not separate from Hb-A on this medium. Hb-E trait is asymptomatic and the hemoglobin level is normal. The red cells are microcytic with a mildly reduced MCV and MCH. Hb-E disease may also be asymptomatic in some patients. It is characterized by compensated hemolysis with a normal hemoglobin level. There may be mild jaundice. In the blood film there is marked hypochromia usually with many target cells. The MCV and MCH are reduced. Hb-E β thalassaemia is relatively common in Thailand. It is a more severe condition than Hb-E disease. The clinical and hematological features resemble those of β-thalassaemia major in most cases. Hb-E α thalassaemia is also common in Thailand. Several disorders of variable severity involving interactions between the α-thalassaemia1 and α-thalassaemia2 genes and Hb-E have been described (Firkin *et al*, 1989).

Haemoglobin E is a β chain variant, $\alpha 2\beta 2^{26Glu \rightarrow Lys}$. It was first described by Chernoff and colleagues in 1954 and independently in the same year by Itano and colleagues. The highest prevalence occurs in some parts of Thailand, Kampuchea (previously Cambodia) and Laos. Thailand and Myanmar (previously Burma) have an overall prevalence of around 14–15%. Gene frequency in Thailand varies from 8% to 50–70%, being highest in northeastern Thailand. Haemoglobin E is also found in Sri

Lanka, north-eastern India (Bengal and Assam), Bangladesh, Pakistan, Nepal, Vietnam, Malaysia, the Philippines, Indonesia and Turkey. Although haemoglobin E is prevalent in Sri Lanka, it is not prevalent in southern India; it is thought to have reached Sri Lanka during migration from north-eastern India during the 5th century BC. Occasional cases have been observed in individuals of apparent Northern European Caucasian descent and a single affected family has been observed in the former Czechoslovakia. The β E chain is synthesized at a reduced rate in comparison with β A. This is because the mutation creates a false splicing site towards the 3' end of exon 1 so that there is a proportion of abnormally spliced messenger RNA (mRNA). Post-transcriptional processing of the latter is abnormal. The result of the reduced rate of synthesis of the β E chain, and therefore of haemoglobin E, is that heterozygotes, compound heterozygotes and homozygotes show some thalassaemic features. Haemoglobin E may therefore be regarded as a thalassaemic haemoglobinopathy. The α: non-α chain synthesis ratio is 1.2–2.1 in heterozygotes. Haemoglobin E also has weakened α1β1 contacts, leading to instability in conditions of increased oxidant stress. The most significant clinical consequences occur if haemoglobin E is coinherited with β thalassaemia trait, leading to thalassaemia major or thalassaemia intermedia. Homozygosity for haemoglobin E produces a clinically mild condition and is thus of much less significance. Haemoglobin E may protect against severe falciparum malaria (Bain, 2006).

1.5.5. The haemoglobin D (Hb-D) haemoglobinopathies:

Hb-D occurs mainly in north-west India, Pakistan, and Iran, although it was first described in the United States. About 3% of Sikhs living in the Punjab are affected. It is also found sporadically in Blacks and Europeans. The original Hb-D was called Hb-D Los Angeles, and it was later shown to be identical to Hb-D Punjab found in India and Pakistan. Hb-D Punjab arises from the substitution of glutamine for glutamic acid in the 121st position of the β chain. The electrophoretic mobility of Hb-D on cellulose acetate is identical to that of Hb-S. On Agar gel electrophoresis, Hb-D migrates with Hb-A, whereas Hb-S separates from Hb-A. Hb-D disorders were reviewed by Yella and Lehman (1974) (Firkin *et al*, 1989).

Haemoglobin D-Punjab is a β chain variant, $\alpha 2\beta 2^{121Glu \rightarrow Gln}$, initially described under the name of haemoglobin D-Los Angeles. The latter is the correct designation and is used in North America, whereas in the UK the designation D-Punjab tends to be employed. This variant haemoglobin has also been designated haemoglobin D-Cyprus, D-Conley, D-Chicago, D-North Carolina, D-Portugal and haemoglobin Oak Ridge. Its only major importance is because of its interaction with haemoglobin S. Its highest incidence is amongst Sikhs in the Punjab who show a prevalence of 2–3%. Gujeratis have a prevalence of about 1% and a similar prevalence is found in British Pakistanis. There is a low but significant prevalence amongst Afro-Americans (0.4%) and Afro-Caribbeans. It has also been reported in individuals with American Indian ancestry. There is a low prevalence amongst Caucasian populations in England, France and Portugal who have had a close contact with India. In England, the highest prevalence is in Norfolk where it has been attributed to the sojourn of the IXth Foot Regiment in India. Haemoglobin D-Punjab is also found in Pakistan, Afghanistan, Iran, China, Turkey, Yugoslavia, Greece, Holland, Australia and North Africa. It is important to distinguish haemoglobin D-Punjab from other α and β chain variants with similar electrophoretic properties, but with less or no clinical significance. Haemoglobin D has a normal stability and a normal or slightly increased oxygen affinity. Haemoglobin D-Punjab trait Heterozygosity for haemoglobin D is of genetic but no other clinical significance (Bain, 2006).

1.5.6. Unstable hemoglobins:

The unstable hemoglobin disorders are a group of inherited hemolytic anemias that result from structural changes in the hemoglobin molecule, which cause its intracellular precipitation with the formation of Heinz bodies (Hoffbrand *et al,* 2005). Over 100 unstable hemoglobins have been described. They include Hb-Koln, Hb-Zurich, Hb-Hammersmith, and Hb-Sydney. Most arise from β chain substitution, and affected patients are heterozygous for the unstable hemoglobin and Hb-A. The condition is not limited to any particular racial group. The disorder is fully reviewed by White (1976) (Firkin *et al*, 1989). Amino acid substitution close to the heme group, or at the point of contact between globin chains, can affect protein stability and result in intracellular precipitation of globin chains. There is considerable heterogeneity in the

hematological and clinical effects of unstable hemoglobins. Many are almost silent and are detected only by specific tests, whereas others are severe, causing hemolytic anemia in the heterogeneous state (Bain *et al*, 2012).

The term 'unstable haemoglobin' is best restricted to those variant haemoglobins that cause clinically recognizable haemolysis. Other haemoglobins, for example haemoglobin E, are unstable *in vitro*, but this probably does not contribute significantly to the associated clinical features. Both α and β chain variant haemoglobins may be unstable. Two γ chain variants have been described which led to haemoglobin instability in a neonate, and two γ chain variants that were methaemoglobins were also unstable. δ chain variants can also be unstable but, because of their low concentration, this is of no clinical significance. Several patients have been described with two unstable haemoglobins, together with haemoglobin A or a β^o thalassaemia determinant. This results not from three β genes in the genome, but from posttranslational modification of an unstable haemoglobin. Both haemoglobin Sydney and haemoglobin Atlanta can be modified in this manner, producing haemoglobin Sydney-Coventry and haemoglobin Atlanta-Coventry, respectively. Haemoglobin instability can result from an unstable globin chain or an unstable haemoglobin molecule. The causative primary abnormality can affect the secondary structure (a helix and intervening turns), the tertiary three dimensional structure of the monomer or the quaternary structure (relationship between monomers).

Haemoglobin instability may be consequent on:

• An abnormality of the haem pocket, so that haem is not firmly bound and water can enter the normally hydrophobic haem pocket; haem-depleted dimers and tetramers are present;

• Interference with the a helical structure, often because an amino acid is replaced by the imino acid proline, or interference with the interhelical bends (abnormality of the secondary structure);

• Replacement of an internal non-polar amino acid with a polar amino acid, which must be oriented outwards and thus disrupts the molecule (interference with the tertiary structure);

• Interference of the binding of α and β subunits to each other, specifically impairment of the α1β1 dimeric bonds (interference with the

quaternary structure); this can result in dissociation into monomers, favouring methaemoglobin formation;

• Elongation of the β chain. The first unstable haemoglobins identified were two β chain variants, haemoglobin Zurich and haemoglobin K, reported in the early 1960s. Subsequently, a case of 'congenital non spherocytic haemolytic anaemia', reported in the 1950s, was found to be attributable to haemoglobin Bristol. Since these early reports, more than 100 unstable haemoglobins have been described of which haemoglobin ln appears to be the most common. Many more unstable β variants K have been reported than unstable α variants, perhaps because α variants, being a lower proportion of total haemoglobin, are more likely to go unrecognized.

Unstable haemoglobins may result from:

• Point mutations leading to replacement of one amino acid by another or replacement of an amino acid by the imino acid proline;

• Point mutation followed by post-translational modification of the haemoglobin encoded; leucine in an abnormal haem pocket is modified to hydroxyleucine.

• Deletions of one to eight codons leading to the deletion of a small number of amino acids, e.g. haemoglobin Gun Hill has a β chain which lacks five amino acids including the haem-binding site, and haemoglobin J-Biakra (unstable only *in vitro*) has an α chain which lacks eight amino acids;

• Tandem duplication of codons leading to duplication of a small number of amino acids, e.g. haemoglobin Fairfax has five extra amino acids, the same amino acids that are deleted in haemoglobin Gun Hill;

• Frame shift or stop codon mutations leading to the synthesis of an elongated β chain. Of these various mechanisms, by far the most common is a single amino acid substitution. Unstable haemoglobins show a greater or lesser tendency to Heinz body formation. Heinz bodies bind to the inner surface of the red cell membrane, probably by hydrophobic interactions rather than covalent bonds.

The proportion of the variant haemoglobin is determined by:

• The rate of synthesis of the variant chain;

• The rate of breakdown of the β variant chain before association with the α chain can occur; the rate of association of α and β variant chains in

comparison with the rate of association of α chains and normal β A chains;

• The rate of breakdown of αβ variant dimers before assembly of dimers to tetramers can occur;

• The rate of denaturation of unstable haemoglobin (with consequent removal as Heinz bodies). In the case of unstable α chain variants, the proportion of variant haemoglobin is usually less than 15%, but may be 1–2% or even less. The explanation for the variable proportion of the variant haemoglobin is the same as for the β chain variants. In some instances, the proportion of an unstable haemoglobin is also affected by the fact that the mutation occurs in *cis* to an α thalassaemia determinant (haemoglobin Suan-Dok, haemoglobin Petah Tikva) or a β thalassaemia determinant (haemoglobin Leiden). Depending on the degree of instability of a variant globin chain or variant haemoglobin, various degrees of abnormality are possible: (i) a very unstable α or β chain is destroyed so rapidly that no variant globin chain or haemoglobin is detectable and the phenotype is that of thalassaemia; (ii) an unstable globin binds haem, but cannot bind to other globin chains, so that it precipitates as Heinz bodies in erythroid precursors; this leads to dyserythropoietic and ineffective erythropoiesis, as in dominant β thalassaemia intermedia resulting from heterozygosity for a hyper-unstable β chain; (iii) a lesser degree of instability permits a variant haemoglobin to be synthesized, leading to a haemolytic anaemia and the clinical features recognized in association with an unstable haemoglobin. Highly unstable α chain variants are usually clinically silent but, if they interact with an α thalassaemia determinant, the phenotype may be either that of haemoglobin H disease or a haemolytic anaemia or may simulate β thalassaemia intermedia; in the latter instance, there is no haemoglobin H detectable, there is dyserythropoiesis and globin chain synthesis studies show a paradoxically increased α:β ratio. A phenotype with features of both thalassaemia and a Heinz body haemolytic anaemia can be produced either by a reduced rate of synthesis of an unstable variant or by marked instability leading to destruction of the variant globin chain or of αβ dimers before assembly of haemoglobin tetramers can occur (Bin, 2006).

1.5.7. Hemoglobins with altered oxygen affinity:

Hemoglobin variants with altered oxygen affinity are a rare group of variants that result in increased or reduced oxygen affinity (Bain *et al*, 2012).

High oxygen-affinity hemoglobin variants: Homozygotes for hereditary persistence of hemoglobin F are mainly polycythaemic because of the high oxygen affinity of Hb-F.

Low oxygen-affinity hemoglobin variants: At least six hemoglobin variants with reduced oxygen affinity have been reported. The first to be described, Hb-Kansas, was found in a mother and son with unexplained cyanosis (Hoffbrand *et al*, 2005).

1.5.8. Hemoglobin M (Hb-M):

The Hb-M group is another rare group of variants. Such hemoglobins have a propensity to form methemoglobin, generated by the oxidation of ferrous iron in heme to ferric iron, which is incompatible of binding oxygen. Methemoglobinaemia is also found in congenital NAPH methemoglobin reductase deficiency, as well as after exposure to oxidant drugs and chemicals (nitrates, nitrites, quinines, chlorates, phenacetin, dapsone, and many others (Bain *et al*, 2012).

The designation haemoglobin M is given to a variety of α, β and γ chain haemoglobin variants that show an increased tendency to oxidation to methaemoglobin, with consequent cyanosis or pseudocyanosis. However, with haemoglobins M-Boston and M-Iwate, which have a low oxygen affinity, there is also an element of true cyanosis. In Japan, terms meaning ‘black mouth’, ‘black blood’ or ‘black child’ were used for carriers of haemoglobin M-Iwate. The first haemoglobin M was recognized in 1948, a year before the first description of haemoglobin S. The molecular abnormality is usually the replacement of a histidine residue in the haem pocket by tyrosine, so that the iron of the haem molecule is stabilized in the ferric (Fe^{+3}) form. Either the proximal or the distal histidine may be involved. The exception is haemoglobin M-Milwaukee in which the longer side-chain of glutamic acid, substituted for valine at β 67, reaches the haem molecule and stabilizes an abnormal ferric state. Once the variant globin chain has been oxidized, it becomes non-functional from the viewpoint of oxygen transport. In the case of the

three α chain variants, M Boston, M-Iwate and M-Milwaukee, only the β chains can bind oxygen; in the case of the two β chain variants, M-Saskatoon and M-Hyde Park, only the α chains can bind oxygen (Bain, 2006).

1.6. The Thalassaemia syndromes:

The thalassaemia syndromes are a heterogeneous group of inherited conditions characterized by defects in the synthesis of one or more of the globin chains that form the hemoglobin tetramer. The clinical syndromes associated with thalassaemia arise from the combined consequences of inadequate hemoglobin production and of unbalanced accumulation of one type of globin chain. Clinical manifestations range from completely asymptomatic microcytosis to profound anemia that is incompatible with life and can cause death in utero (Bain *et al*, 2012). The thalassaemias are the commonest single-gene disorders. The condition was first recognized in 1925 by a Detroit physician, Thomas B. Cooley who described a series of infants who become profoundly anemic and developed splenomegally over the first year of life. In 1936, Whipple and Bradford, in describing the pathological changes of the condition for the first time, recognized that many of their patients came from the Mediterranean region and hence invented the word thalassaemia for the Greek word θαλασσα, meaning 'the sea' (Hoffbrand *et al*, 2005). Thalassaemias are generally inherited as alleles of one or more of the globin genes located on either chromosome 11 (for β, γ and δ chains) or on chromosome 16 (for α chains). They are encountered in every population in the world but are most common in the Mediterranean littoral and near equatorial regions of Africa and Asia. Gene frequencies for the α and β thalassaemias on a global basis range from 1% to more than 80% in areas where malaria is endemic (Bain *et al*, 2012). They are divided into the α-, β-, δβ- or γδβ-thalassaemias, according to which globin chain is synthesized at all, and those are called α°- or β° thalassaemias; in others, designated α^{+}- or β^{+}-thalassaemias, the globin chain is produced at a reduced rate. The δβ-thalassaemias are designated in the same way. Clinically, the thalassaemias are classified according to their severity into major, intermediate and minor forms. The thalassaemia major is a severe, transfusion-dependent disorder, thalassaemia intermedia is characterized by anemia and splenomegally, though not of such severity as to require regular transfusion. Thalassaemia minor is the symptomless carrier state

(Hoffbrand *et al*, 2005). Prevention of β-thalassaemia implies knowledge of molecular spectrum occurring in the population at risk. This knowledge is necessary when a prevention protocol is applied to multiethnic population. So the detection of all mutations of β-thalassaemia in each population is the major goal in prevention, and is especially helpful for prenatal diagnosis. (Roudknar *et al*, 2003).

Table 1- 6: The β-thalassaemias and related syndromes: electrophoretic phenotypes:

Disorder	Hemoglobin %		
	A	A_2	F
β –thalassaemia minor	90-95	3.5-7	1-5
δβ-thalassaemia minor	80-95	1.0-3.5	50-20
β-thalassaemia major:			
β-thal^{+}	10-90	1.5-4.0	10-90
β-thal$^{\circ}$	0	1.5-4.0	98
HPFH (Black heterozygotes)	60-88	1.0-2.0	15-35

HPFH= Hereditary persistence of fetal hemoglobin (Source, Firkin *et al*, 1989)

1.6.1. β-thalassaemia syndromes:

The β-thalassaemia occurs widely in a broad belt, ranging from the Mediterranean and parts of North and West Africa through the Middle East and India subcontinent to south-east Asia. It should be remembered that β-thalassaemia is not confined entirely to those high incidence regions and it occurs sporadically in every racial group (Hoffbrand *et al*, 2005). Beta thalassaemia is a type of inherited blood disorder that can cause anemia (a low number of red blood cells). It affects a person's ability to produce hemoglobin, the protein in red blood cells that delivers oxygen to all parts of the body. Worldwide, beta thalassaemia is considered a fairly common blood disorder, affecting thousands of infants each year. Beta thalassaemia occurs most frequently in Mediterranean countries, North Africa, the Middle East, India, and south East Asia. In North America, the disorder is less common; an estimated 750-1000 people have beta thalassaemia (Parker and Parker, 2007). Many different mutations cause β-thalassaemia and related disorders. These mutations can affect every step in the pathway of globin gene expression. More than 300 mutations have been described. β-thalassaemia major is a severe transfusion-dependent, inherited anemia. There is a profound defect of β chain production. Excess α-chains accumulate and precipitate in the red

cell precursors in the bone marrow resulting in ineffective erythropoiesis. If untreated, 80% of children with β-thalassaemia die within the first 5 years. Homozygotes for β-thalassaemia alleles usually have either a normal hemoglobin with microcytosis or a mild microcytic hypochromic anemia; Hb-A_2 is elevated and Hb-F is sometimes also elevated (Bain *et al*, 2012). Hb-A level is always elevated and it is heterogeneously distributed among the red cells. In β°-thalassaemia, there is no Hb-A, whereas in β^{+}-thalassaemia the level of Hb-F ranges from 30% to 90%. Carriers for β-thalassaemia are usually symptom free except in periods of stress such as pregnancy, where they may become more anemic. Splenomegally is rare and hemoglobin values are in the 9-11g/dl range. The red cells show hypochromia and microcytosis with characteristically low values. The reticulocyte count is normal. The bone marrow shows moderate erythroid hyperplasia. The characteristic finding is an elevated Hb A_2 level in the 4-6% range. There is a slight elevation of Hb-F in the 1-3% range in about 50% of cases (Hoffbrand *et al*, 2005). Beta-thalassaemia mutations are characterized according to the extent of defect in formation of beta-globin chains with β^{o} referring to the mutations that prevent the synthesis of β-globin chains while β^{+} to mutations that allow some synthesis of β-globin chains to occur. In either situation, there is a relative excess of α-globin chains, which do not form homo tetramers and instead, they bind to red blood cell membrane causing reduced red cell lifespan. Therefore, mutation analysis of beta-globin gene is useful for the prediction of clinical phenotype of beta-thalassaemia, pre-symptomatic diagnosis of family members at-risk, prenatal diagnosis and also for prevention and control programs. However, due to the marked heterogeneity in mutations, screening of mutations in beta-thalassaemia is time consuming and expensive (Dassanayake *et al*, 2013). Beta thalassaemia major and thalassaemia intermedia are inherited in an autosomal recessive pattern, which means two copies of the gene in each cell, are altered. Most often, the parents of an individual with an autosomal recessive disorder are carriers of one copy of the altered gene but do not show signs and symptoms of the disorder. Sometimes, however, carriers of the altered HBB gene have a mild anemia referred to as thalassaemia minor. In a small percentage of families, the HBB mutation is inherited in an autosomal dominant manner. In these cases,

one copy of the altered gene in each cell is sufficient to cause the disorder (Parker and Parker, 2007).

1.6.1.1. β- thalassaemia in association with Hemoglobin variants: In many populations, because there is a high incidence of β-thalassaemia and various hemoglobin variants; it is quite common for individual to inherit a β-thalassaemia gene from one parent and a gene for a structural hemoglobin variant from the other. Although numerous interactions of this type have been described, only three are common: sickle cell β-thalassaemia, Hb-C β-thalassaemia and Hb-E β-thalassaemia (Hoffbrand *et al*, 2005). β-thalassaemias are prevalent heritable single gene disorders affecting the quantity of the hemoglobin molecule. Rarely, a co-inheritance of these impairments with α-thalassaemia and/or a haemoglobinopathy occurs and makes an important double heterozygote or homozygous state. Finding these patients is essential for genetic counseling, since, unlike the typical β-thalassaemia carriers, these individuals coupled with another α-thalassaemia heterozygote will be at risk of having progeny with homozygous α-thalassaemia. Moreover, clinically significant diseases can occur from α and β gene defect coinheritances. Patients with both HbS and β-thalassaemia can produce a variable severity of sickling syndrome. A patient carrying Hb E and β-thalassaemia mutations simultaneously (Hb E/β-thalassaemia) will clinically be similar to a β-thalassaemia major or intermedia (Alizadeh, 2014).

1.6.1.2. Delta beta thalassaemia:

A type of thalassaemia resulting from complete absence of both β- and δ-chains, in most cases due to extensive deletion of DNA in the β-globin gene complex, occurs in areas where the β-thalassaemia gene is prevalent. The homozygous form of the condition is rare, and is characterized by a β-thalassaemia major-like illness of moderate severity and 100% Hb-F in the red cells. The heterozygous form is more common, and closely resembles β-thalassaemia minor, however the Hb-A_2 level is normal or reduced, and the only abnormality of hemoglobin constitution is an elevation of Hb-F which ranges from 5-20% (Firkin *et al*, 1989).

1.6.1.3. The hemoglobin-Lepore syndrome:

Occasional patients with apparent classical β-thalassaemia minor or major have abnormal hemoglobin, Hb-Lepore. Structural analysis has shown that Hb-Lepore is made up of normal α-chains combined with

chains that consist of parts of both δ and β chains. Hb-Lepore syndromes are found in many population groups and have a similar mobility on cellulose acetate electrophoresis. It does not separate from Hb-A on agar gel (Firkin *et al*, 1989).

1.6.2. α-Thalassaemia Syndromes:

The α-thalassaemia occur widely throughout Sub-Saharan Africa, the Mediterranean region, the Middle East, the Indian Subcontinent and South-East Asia in a line stretching from southern China through Thailand, the Malay peninsula and Indonesia, to the Pacific island populations (Hoffbrand *et al*, 2005). There are four syndromes of α-thalassaemia: α^{+}-thalassaemia trait, where one of the two globin genes on a single chromosome fails to function; α°-thalassaemia trait, where two genes on a single chromosome fail to function; Hb-H disease, with three genes affected; and Hb-Bart's hydrops fetalis, where all four are defective. Theses syndromes are usually a result of deletions of one or more genes, although approximately 20% of the mutations described are nondeletional. α^{+}-thalassaemia is particularly common in Africa, and α°-thalassaemia is common in Southeast Asia (Bain *et al*, 2012). The α-thalassaemias are disorders in which there is defective synthesis of α chains with resulting depression of production of the hemoglobins that contain α chains, i.e. Hb-A, Hb-A_2 and Hb-F. The deficiency of α chains leads to an excess of γ chains in the fetus, and β chains in the adult. The γ chains form the tetramer Hb-Bart's (γ_4), and the unstable β chains precipitate and form Hb-H (β_4). The clinical and hematological studies of α-thalassaemia in the 1950s and 1960s identified four forms of the disorder (Table 1-7) (Firkin *et al*, 1989).

1.6.2.1. Classification of α-thalassaemia syndrome:

Table 1-7: Types of α-thalassaemia:

Clinical	Genetic	Genotype	Number of genes
α-thalassaemia 2	Heterozygous α^{+}-thalassaemia	αα/-α	3
α-thalassaemia 1	Homozygous α^{+}-thalassaemia Heterozygous α°-thalassaemia	-α/-α αα/--	2 2
Hb-H disease	Double heterozygote $\alpha^{+}/\alpha^{\circ}$ thalassaemia	-α/--	1
Hb-Bart's hydrops fetalis	Homozygous α° thalassaemia	--/--	0

Source (Firlin *et al*, 1989).

1.6.2.2. Hemoglobin H disease:

Interaction of the α^{+}(-α/) and the α°(--/) determinants gives rise to this form of α-thalassaemia (-α/--) which is common in south-east Asia and is also seen in the middle East and some Mediterranean countries. It is rare in Blacks. The blood film shows marked red cell morphological changes including severe hypochromia and microcytosis, target cell formation, and basophilic stippling. The hemoglobin pattern consists of 2-40% Hb-H, the remainder being Hb-A, Hb-A_2 and Hb-F (Firkin *et al*, 1989).

1.6.2.3. Hemoglobin Bart's hydrops fetalis:

The most severe manifestation of the α-thalassaemia gene is hemoglobin Bart's hydrops fetalis which is common in south-east Asia, but rare in other parts of the world where the α-thalassaemia gene is found. Affected infants are homozygous for the α° determinant (--/--), both parents having heterozygous α°-thalassaemia. There is almost total suppression of α chain synthesis with a gross excess of γ chains. The γ chain tetramer, Hb-Bart's, has a high oxygen affinity and severe tissue hypoxia results. Affected infants are either born dead or die within a few hours of birth. The hemoglobin pattern consists of 80-90% Hb-Bart's, with a small amount of Hb-H and a third 'fast' component, Hb-Portland. There is usually no Hb-A, Hb-A_2 or Hb-F (Firkin *et al*, 1989).

1.6.2.4. Increased Hb-F in adult life:

Hemoglobin production in man is characterized by two major switches in the hemoglobin composition of the red cells. During the first 3 months of gestation, human red cells contain embryonic hemoglobins, whereas during the last 6 months of gestation, red cells contain predominantly fetal hemoglobin. The major transition from fetal to adult hemoglobin synthesis occurs in the perinatal period, and by the end of the first year of life red cells have a hemoglobin composition that usually remains constant throughout adult life. The major Hemoglobin is then Hb-A, but there are small amounts of Hb-A_2 and Hb-F. Only 0.2-1.0% of total hemoglobin in human red cells is Hb-F, and it is restricted to a few cells called "F" cells (Lewis *et al*, 2006). Hereditary Persistence of Fetal Hemoglobin (HPFH) is a heterogeneous group of conditions in which there is persistent fetal hemoglobin production in adult life in the absence of major hematological abnormalities. Some forms of HPFH result from long deletions of the β globin gene cluster, similar to those that cause δβ-thalassaemia. Homozygotes, with 100% Hb-F, have a mild thalassaemia like blood picture and are mildly polycythaemic, reflecting the high oxygen affinity of Hb-F. Heterozygotes have no hematological abnormalities but have 20-30% Hb-F (Hoffbrand *et al*, 2005). The general organization of the production of the two γ genes differs in only one amino acid: $^{G}\gamma$ has glysine in position 136, whereas $^{A}\gamma$ has alanine. In fetal red cells, the ratio of $^{G}\gamma$ to $^{A}\gamma$ is approximately 3:1; in adult red cells it is approximately 2:3. The most common, the African type of HPFH, is associated with a high concentration of Hb-F, pan cellular distribution of Hb-F in Kleihauer staining, and normal red cell indices (Bain *et al*, 2012). In most population groups, a small number of apparently healthy adult subjects with normal or near normal hematological findings have a raised Hb-F level (Firkin *et al*, 1989).

1.7. Natural selection and the haemoglobin disorders:

There are several reasons why the haemoglobin disorders have reached such extra ordinarily high gene frequencies in many tropical countries. First and foremost, it appears that their frequency reflects natural selection through protection of heterozygotes against severe malaria, a mechanism first suggested by Haldarre (1949) and latter confirmed independently in the case of sickle cell anaemia by Allison (1954 a, b, c). Evidence has steadily grown to confirm that malaria is indeed the primary

force behind the high frequency of inherited haemoglobin disorders that is seen in many tropical and sub-tropical populations. The evidence supporting this assertion comes from four main sources: the similarity between the distributions of malaria and specific haemoglobin disorders at local, regional and global scales; from population genetic predictions of their historic age; from clinical studies concluded in malaria-endemic areas; and from mechanistic studies conducted both in-vitro and ex-vivo. Although such evidence is strongest for Hb-S and for α-thalassaemias, which has been the focus of the most research, there can now be little doubt that malaria is responsible for the current distributions of all the major Haemoglobinopathies (Williams and Weatherall, 2012). The haemoglobinopathies have reached high frequencies in certain geographical regions because of the selective advantage conferred by the carrier state against the malarial parasite *Plasmodium falciparum*. The sickle-cell mutation (β^s), is found in many world populations, the highest frequencies occurring in Africans and their descendants in the Americas. The β^s gene is not found, however, in Southeast Asia (reviewed by Nagel and Fleming 1992). The β^s gene has been found in linkage disequilibrium with five major β-globin gene cluster haplotypes: four of these occur homogeneously and virtually exclusively in four segregated regions of Africa and the fifth β^s haplotype is almost exclusively present in individuals from eastern Saudi Arabia and India. Thus, knowledge of the β^s haplotype that occurs in a population may be used to gauge that population's origin(s) (Hewitt *et al.* 1996). The gene for sickle hemoglobin (HbS) is a prime example of natural selection. It is generally believed that its current prevalence in many tropical populations reflects selection for the carrier form (sickle cell trait [Hb AS]) through a survival advantage against death from malaria. Nevertheless, >50 years after this hypothesis was first proposed, the epidemiological description of the relationships between Hb AS, malaria, and other common causes of child mortality remains incomplete (Williams *et al*, 2005).

Table 1-8: Prevalence of α and β-thalassaemia in different African populations (modified from Bain, 2006):

Country or population	α^{o}-thalassaemia	α^{+}-thalassaemia	β-thalassaemia
West Africa	Nil	Gambia 8–15%, Togo 46%, Nigeria 8–58%, Senegal 22%, Benin and Burkina Faso (formerly Upper Volta) 29%, Ivory Coast 39%	Overall 1–14%: Senegal <1–5%, Liberia <1–9%, Ivory Coast 1–12%, Mali < 1%, Burkina Faso 2–12%, Ghana 1–11% (overall 1–2%), Togo <1–2%, Nigeria 1–4% (overall 0.8%), Cameroon <1–2%
East Africa	Nil	Kenya 19–34%, Tanzania 2%	Rare in Kenya, Uganda and Tanzania; 2% in Mauritius and Reunion Island
Central Africa	Nil	Central African Republic 39% (23% of pygmy population), Republic of the Congo 36–40% (29% of pygmy population)	Central African Republic <1%, Republic of the Congo <1% (pygmy population of 6.5%)
Southern Africa	Nil in South African blacks	Zambia 20–27%, Malawi 39%, Namibia 11.5%, South African Cape Coloured population 7%, South African black 12% (San) to 36% (Venda), Mozambique 5–6%, Madagascar <1–3%, Comoros 2% or more	Comoros 3%
Africans in UK		Probably 25–30%	0.9%
Afro-Americans		25%	1–2%

1.8. Investigation of a suspected haemoglobinopathy:

1.8.1. Basic Hematological tests:

Initial investigation should include determination of hemoglobin concentration and red cell indices. A detailed examination of a well-stained blood film should be carried out. In some instances, a reticulocyte count and a search for red cell inclusions give valuable information. Other important basic tests are hemoglobin electrophoresis or chromatography, a sickle solubility test and measurement of Hb-A_2 and Hb-F percentage. The blood count, including Hb and red cell indices provides valuable information useful in the diagnosis of both α and β thalassaemia interactions with structural variants. A film examination may reveal characteristic red cell changes such as target cells in Hb-C trait, sickle cells in sickle cell disease, and irregularly contracted cells in the presence of an unstable hemoglobin (Lewis *et al*, 2006).

Besides the use of DNA technology as a confirmative tool, some thalassaemic syndromes can only be diagnosed by that technology.

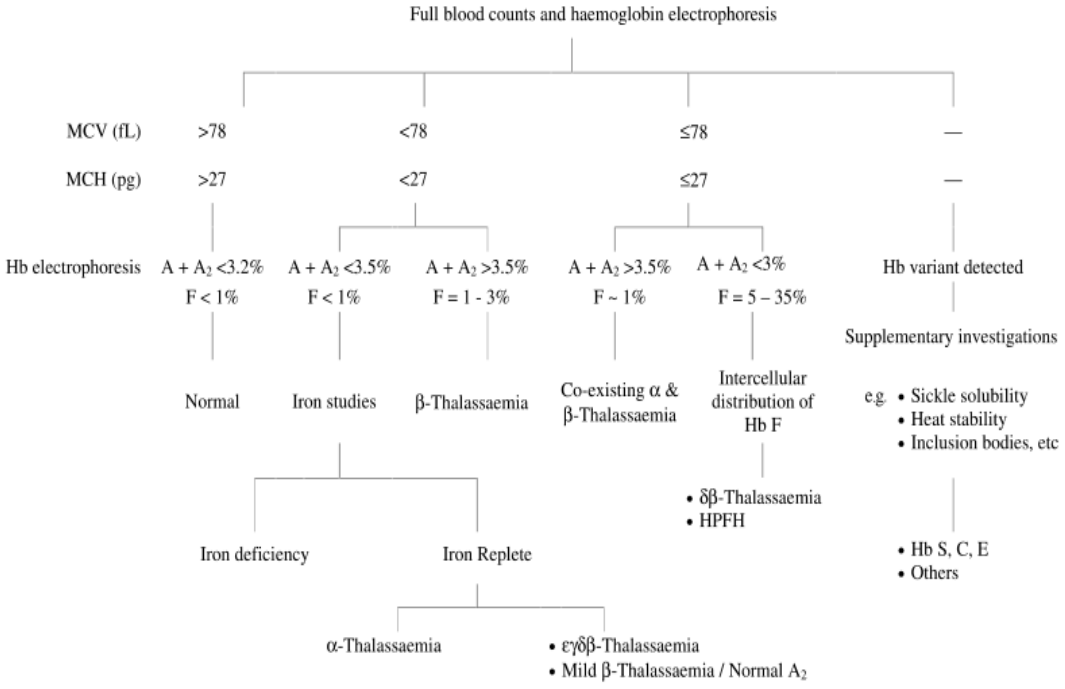


Figure 1-5: A guide to the diagnosis of the different forms of haemoglobinopathies in carriers (adapted from Clark and Thein, 2004).

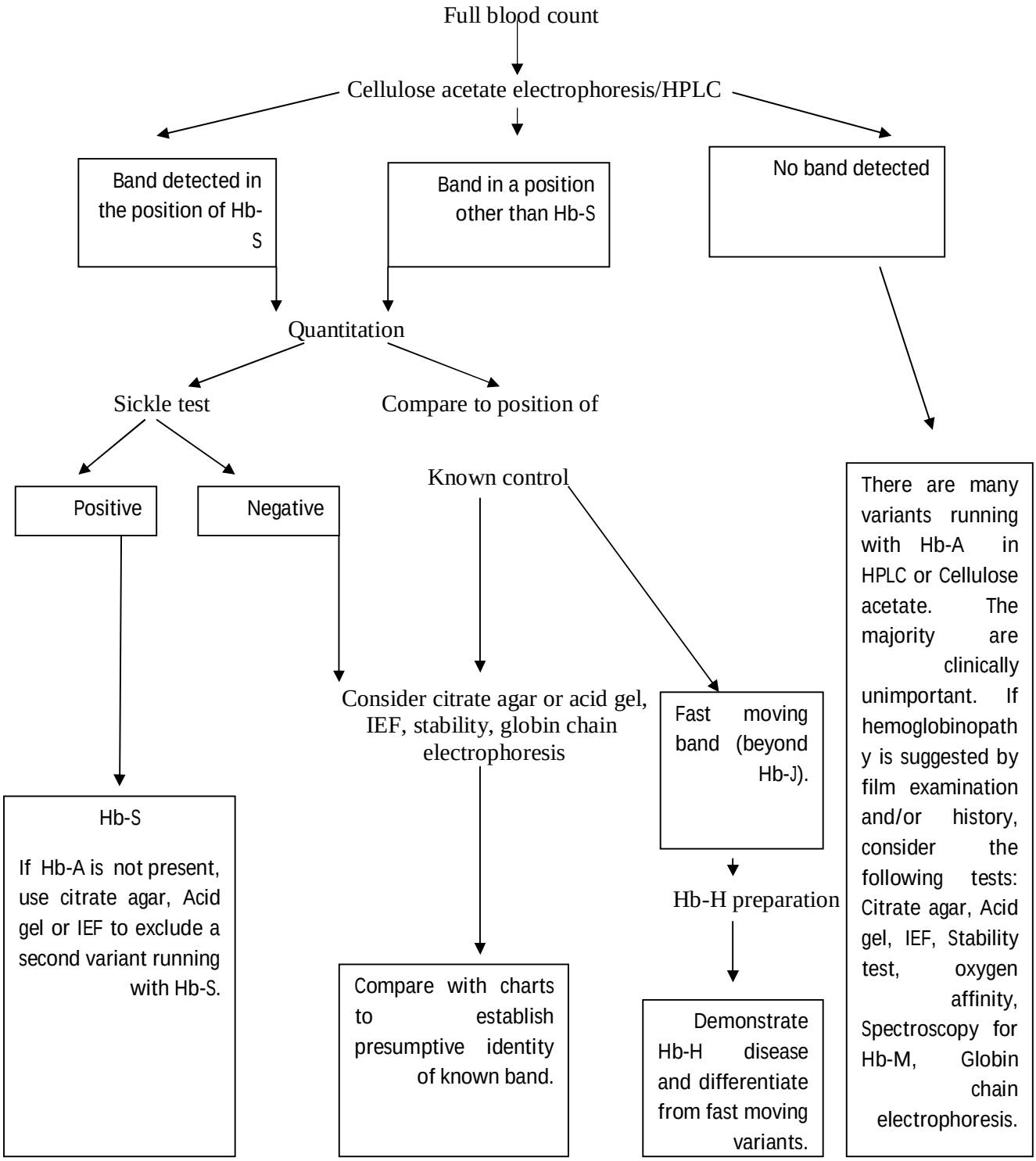


Figure 1-6: Suggested scheme of investigation for structural variants (Barbara et al 2012).

1.8.2. Hemoglobin electrophoresis:

Hemoglobin electrophoresis is the most useful method for the demonstration of abnormal hemoglobins. The hemoglobins are separated on a variety of supporting media on the basis of electric charge differences. Cellulose acetate electrophoresis at pH 8.6 is the method of choice in most clinical laboratories. Agar gel electrophoresis using a citrate buffer at pH 6.0 is useful in supplementing (but not replacing) the information gained from other methods, as the mobility of some abnormal hemoglobins on agar gel differs from that on other supporting media (Firkin *et al*, 1989). Electrophoresis has long been the method of choice in hematological laboratories for qualitative and quantitative hemoglobin analyses. Currently, four different electrophoresis techniques are routinely used in the lab setting: 1) alkaline and acid gel electrophoresis, 2) isoelectric focusing (IEF), 3) high-pressure liquid chromatography (HPLC) and 4) capillary electrophoresis (CE) (Brant, 2010). In the proposal of this study it was proposed to carry out the haemoglobin electrophoresis analysis by High Performance Liquid Chromatography (HPLC) technique, but we have deviated and made it by Capillary Electrophoresis (CE) technique. That was due to the unavailability of HPLC and that CE is preferred to HPLC in that: In addition to that in one analysis, separation of Hb S from Hb D, and Hb C from Hb E (and from Hb A2) is obtained;

1- It is precise, quick in the quantification of Hb F and Hb A2, even in the presence of Hb S.
2- Superior to HPLC in the measurement of Hb A2 in the presence of Hb E.
3- Delta chain variants, alpha chain variants, and other minor Hb fractions are readily visualized.
4- Hb H and Hb Bart's are more readily detected and measured by CE than by the HPLC method (Brant, 2010).

Automated capillary zone electrophoresis is a new and complementary screening technique to HPLC for the routine detection and measurement of haemoglobins and variants. CE patterns for abnormal haemoglobins are simple and easy to read (as long as Hb A is present in the sample), as the method does not separate haemoglobin derivatives such as glycated fractions. CE is able to detect and quantify Hb H and Hb Bart's. Should be noted that two out of three Bio-Rad HPLC methods are also able to

detect and quantify Hb H and Hb Bart's; if needed a software function can allow quantification with the third method. CE is accurate in the absence of variants (see above) and high-throughput and has as an advantage over HPLC in that it separates Hb E from HbA_2, thus providing a clean measurement of HbA_2 in patients with Hb E (Old *et al*, 2012).

Table 1-9: Hb patterns in Haemoglobin disorders (Modified from Drew *et al*, 2004:

% Haemoglobin	A	F	A_2	S	Other
Normal	**97**	**<1**	**2–3**		
β thalassaemia trait	80–95	1–5	3–7		
β thalassaemia intermedia	30–50	50–70	0–5		
β thalassaemia major	0–20	80–100	0–13		
HPFH (Black heterozygote)	60–85	15–35	1–3		
HPFH (Black homozygote)		100			
α thalassaemia trait	85–95				Bart's 0–10% at birth
Hb H disease	60–95				H 5–30%, Bart's 20–30% at birth
Hb Bart's hydrops					Bart's 80–90%
Hb E trait	60–65	1–2	2–3		E 30–35
Hb E disease	0	5–10	5		E 95
Hb E/β thalassaemia	0	30–40	–		E 60–70
Hb E/α thalassaemia	13				E 80
Hb D trait	50–65	1-5	1–3		D 45–50
Hb D disease	1-5	1-3			D 90–95
Hb D/β thalassaemia	0-7	1-7			D 80–90
Hb C trait	60–70				C 30–40
Hb C disease		Slight ↑			C 95
Sickle trait	55–70	1	3	30–45	
Sickle cell anaemia	0	7	3	9	
Sickle/β^+ thalassaemia	5–30	5–15	–	60–85	
Sickle/β^o thalassaemia	0	5–30	4–8	70–90	
Sickle/D	0	1–55	0		D 50%
Sickle/C	0	1		50–65	C 50%
Hb Lepore trait	80–90	1–3	2–2.5		
Hb Lepore disease	0	70–90	0		Lepore 8–30%
Hb Lepore/b thalassaemia		70–90	2.5		Lepore 5–15%

1.8.3. DNA analysis techniques (Genotyping):

The haemoglobinopathies are a diverse group of inherited recessive disorders that include the thalassaemias and sickle cell disease. They were the first genetic diseases to be characterized at the molecular level and, consequently, have been used as a prototype for the development of new techniques of mutation detection. There are now many different polymerase chain reaction (PCR) based techniques that can be used to diagnose the globin gene mutations, including dot-blot analysis, reverse dot-blot analysis, the amplification refractory mutation system (ARMS), denaturing gradient gel electrophoresis (DGGE), mutagenically separated PCR, gap-PCR, and restriction endonuclease analysis (Old, 2003). Most of the HBB gene mutations arise from single nucleotide polymorphisms (SNPs). SNP detection encompasses two broad areas: scanning DNA sequences for previously unknown polymorphisms and screening (genotyping) individuals for known polymorphisms. Scanning for new SNPs can be further divided into the global, or random, approach and the regional, or targeted, approach. Although the technologies capable of scanning DNA for new polymorphisms can be used in screening individuals for known polymorphisms, there are many more options for SNP genotyping. Indeed, many of the most exciting genome technologies are those developed for genotyping tens of thousands of SNPs in thousands of individuals. Furthermore, SNP detection technologies are applicable not only to humans, but to any living organism on earth (Kwok and Chen, 2003). Each method has its advantages and disadvantages, and the particular one chosen by a laboratory to diagnose point mutations depends not only on the technical expertise available in the diagnostic laboratory but also on the type and variety of the mutations likely to be encountered in the individuals being screened (Old, 2003). The ideal genotyping method must be easily and reliably developed from sequence information, robust, inexpensive, flexible, automated, and produces data that are easily analyzed. Although no such ideal genotyping method exists, a number of promising SNP genotyping methods are currently available and further improvements in biochemistry, engineering, and analytical software will bring the existing methods closer to the ideal. The evolution of SNP genotyping technologies is characterized by improvements in the three aspects of the genotyping process: allelic

discrimination, reaction format, and signal detection (Kwok and Chen, 2003).

The main diagnostic approaches for the PCR diagnosis of the haemoglobinopathies are provided in Table 1-10:

Table 1-10: DNA Diagnosis of Haemoglobinopathies:

Disorder and mutation type	Diagnostic method
α^{o}-Thalassaemia	Gap-PCR, Southern blotting
α^{+}-Thalassaemia	
Deletion	Gap-PCR, Southern blotting
Non-deletion	ASO, RE, DGGE
β-Thalassaemia	
Deletion	Gap-PCR
Non-deletion	ASO, ARMS, DGGE
δβ-Thalassaemia	Gap-PCR
HPFH	
Deletion	Gap-PCR
Non-deletion	ASO, ARMS, RE, DGGE
Hb Lepore	Gap-PCR
Hb S	ASO, ARMS, RE
Hb C	ASO, ARMS
Hb E	ASO, ARMS, RE
Hb D Punjab	ASO, ARMS, RE
Hb O Arab	ASO, ARMS, RE
Hb variants	RT-PCR and DNA sequencing

Gap-PCR, gap polymerase chain reaction; ASO, allele-specific oligonucleotide; RE, restriction endonuclease; DGGE, denaturing gradient gel electrophoresis; RT-PCR, reverse transcriptase polymerase chain reaction. (Modified from Old, 2003).

1.8.3.1. Amplification Refractory Mutation System (ARMS):

Amplification refractory mutation system (ARMS) is a PCR-based method utilizing allele specific priming, where specific wild and mutant primers are used in two separate reaction tubes to determine whether a particular mutation is present (in hetero or homozygous state) or it is absent (Nerweiy *et al*, 2010). ARMS is also known as Allele-specific amplification (ASA), or allele-specific PCR (AS-PCR). It is a PCR method used to detect known point mutations or known polymorphisms at the same locus. The method is based on the fact that only a primer with exactly the same sequence as the target DNA will be amplified and primers that have a mismatch at the 3' terminal nucleotide will not be

elongated during PCR reaction. Allele-specific primers are designed to match, at their 3′ ends, the two variants (A/B) at the polymorphic site. The genotyping is run in two complementary PCR reactions, one containing one allele-specific primer (e.g. wild-type specific) and the common reverse primer and the other containing the other allele-specific primer (e.g. mutant specific) and the common reverse primer. DNA polymerase used in this assay must lack the 3'→5' proofreading activity so as not to eliminate the 3' mismatches (Karkun and Bhagat, 2011).

1.8.3.2. Restriction Fragment Length Polymorphism Analysis of PCR- Amplified Fragments (PCR-RFLP):

PCR-RFLP is a method that couples PCR amplification with restriction analysis of the produced DNA fragment. This technique, also known as cleaved amplified polymorphic sequence (CAPS), is very useful in genotyping. Single nucleotide polymorphisms (SNP) often create or abolish specific restriction sites, thus a specific nucleotide variation in this position may be detected using restriction endonucleases that recognize such sites. First, a specific fragment of a gene containing the variation is amplified by PCR. Next, the fragment is digested with an appropriate restriction enzyme and digestion products are resolved by agarose or acrylamide gel electrophoresis. Depending on the presence or absence of the investigated sequence variation, the PCR products digested by restriction enzyme will give readily distinguishable patterns after electrophoresis (Gapic-Ciminska *et al*, 2013). Restriction Fragment Length Polymorphisms, (RFLP's), are DNA differences that are inherited and can be used as genetic markers for diseases such as sickle cell anemia. In this study RFLPs are used to find out the point mutation in beta globin gene. RFLPs arise because mutations can create or destroy the sites recognized by specific restriction enzymes, leading to variations between individuals in the length of restriction fragments produced from identical regions of the genome (Karkun and Bhagat, 2011).

1.8.3.3. DNA sequencing:

Sequencing is a method for determining the nucleotide sequence of a DNA molecule, thus it is a very precise method for detection of exact mutations responsible for specific genetic disorders. Basic sequencing methods, both developed in 1977, include:

- Maxam-Gilbert sequencing, a method based on chemical modification of DNA followed by cleavage at specific bases,

● Sanger dideoxy sequencing, a method based on selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication. The Sanger dideoxy sequencing method became more dominant because it was more convenient in comparison to the Maxam-Gilbert method and the automation of the sequencing process has been developed (in automated DNA sequencer instruments). In general, the Sanger sequencing method involves the following steps: (a) isolation of DNA, (b) PCR amplification of the DNA fragment of interest, (c) cycle amplification using mix of standard nucleotides and dideoxynucleotides, (d) gel electrophoresis or capillary electrophoresis of amplified products. When a dideoxynucleotide incorporates at the 3' end of synthesized DNA chain, the elongation terminates at A, T, C or G depending on the specific dideoxynucleotide that was incorporated. Automated sequencing technologies involve the usage of fluorescent dye-labeled dideoxynucleotides in a cycle sequencing reaction. Each of four dideoxynucleotides (ddNTPs) is labeled with a different fluorescent dye, thus the reaction generates fragments of increasing size, ending with one of four dideoxynucleotides and the exact base is determined by the fluorescence colour detected by the sequencing instrument. In 1996, a new sequencing technology, called pyrosequencing, has been developed by Mostafa Ronaghi and Pal Nyren at the Royal Institute of Technology in Stockholm. Pyrosequencing is a real-time sequencing method and, as distinguished from the Sanger dideoxy method, pyrosequencing is based on the detection of pyrophosphate release during nucleotide incorporation and not chain termination resulting from dideoxynucleotides incorporation. Four enzymes are involved in the pyrosequencing reaction, i.e. DNA polymerase, ATP sulphurylase, luciferase and apyrase. The method detects pyrophosphate (PPi) release accompanying the DNA synthesis process (Karkun and Bhagat, 2011).

1.9. Rationale:

Sudan is a diverse country with a complex population originating from different ethnic groups. This reflects a matter of various intermarriage and social interaction status. Yet a local data base table is not completed. Hence, the determination of the hemoglobin profile in Beja ethnic group of eastern Sudan will potentially fill in the gap in our local data base table. Furthermore, thalassaemias and haemoglobinopathy are the most common inherited disorders among humans, and they represent a major public health problem in many areas of the world (Funcharoen *et al*, 2004). It has been estimated that approximately 7% of the world population are carriers of such disorders and that about 300 000 – 400 000 babies with severe forms of these disorders are born each year (Weatherall and Clegg, 2001 and Hoffbrand, 2005). Inherited hemoglobin disorders were originally characteristic of the tropics and subtropics and are now common worldwide due to migration. The diversity and heterogeneous distribution of hemoglobin disorders make it necessary to develop strategies at country level (Modell and Darlison, 2008). Formerly the distribution of thalassaemia had been mainly limited to the areas from the Mediterranean across the Middle East through Southern Asia to Southeast Asia in the so called ‘thalassaemia belt. However, recent migrations of people have spread thalassaemia genes throughout the world. Furthermore, there has been a major transition in the demography of common illnesses over the past years. This change includes a remarkable decline in childhood mortality due to infections and malnutrition in some regions. As a result, many infants with serious genetic disorders such as thalassaemia can now survive the early months of life and live long enough to require treatment for their hematologic disorders. Given the demographic population shifts, thalassaemia is at present considered to be a global health problem (Gaafar *et al*, 2006). And, since no population is completely free from haemoglobinopathies, and the condition is not limited to any particular racial group (Firkin *et al*, 1989), determination of the possible haemoglobinopathies among this ethnic group will add additional information to our local medical knowledge.

1.10. Objectives:

1.10.1. General objective:

To determine the genotype of haemoglobinopathies in Beja tribes and other minor groups living in Port Sudan city, to add medical information to the local data base table.

1.10.2. Specific objectives:

1- To evaluate the peripheral blood picture among the study group.
2- To determine the Normal hemoglobin pattern in the study group.
3- To detect, quantify and genotype the abnormal hemoglobin variants among the study group.
4- To compare the abnormal haemoglobin variants in the study group to those reported for other races.

2- Materials and Methods

2.1. Study design:

This is a non-interventional, descriptive Cross-sectional study.

2.2. Study population and study area:

All adult Sudanese individuals belonging to Beja ethnic group of Red Sea State of Sudan are subjected to be the study population of this research. Beja lands span the region between the Nile valley and the Red Sea shore.

2.3. Criteria of inclusion and exclusion:

The following criteria will be adopted for inclusion in the study:

Included in this study are hospitalized, anaemic Beja patients who show:

1- General clinical symptoms of anemia.

2- Laboratory features of anemia including:

a- Patients with Hemoglobin concentrations lower than 13.5 g/dl when they are adult males.

b- Patients with Hemoglobin concentrations lower than 11.5g/dl when they are adult females.

c- Other age groups with Hemoglobin concentrations bellow the lower limits of normal in respect to age and sex.

d- Anemic patients belonging to other minor ethnic groups who show peripheral blood picture suggestive for haemoglobinopathy.

Exclusion criteria include:

1- Hospitalized Beja patients who do not show clinical or laboratory features of anemia.

2.4. Study Sample:

A sample of 600 anemic patients referred to Haematology laboratories in three major hospitals (Port Sudan teaching hospital, Sea Ports Corporation hospital, and Police forces hospital) in the city during the period from March 2011 to July 2013 were included.

2.5. Study duration:

This study was carried out from the 1st of October 2010 up to the 30th of April 2015.

2.6. Data collection:

A formerly prepared structured questionnaire is used for data collection.

2.7. Research ethics:

Ethical Approval was done by the Sudan University of Science and Technology. Because of the high rate of illiteracy among the population, only verbal consent was obtained from the participants.

2.8. Procedures and techniques:

2.8.1. Collection of blood specimens:

5 ml of venous blood was collected from each volunteer of the study sample using the Vacutainer collection technique.

2.8.1.1. Requirements:

- 2.5 ml EDTA vacutainers.
- Vacutainer needle holder.
- Tourniquet.
- Adhesive dressing.
- 70% ethyl alcohol.

2.8.1.2. Procedure:

- The skin of the area of collection was cleaned with 70% alcohol and allowed to dry.
- The tourniquet is then applied correctly.
- The needle was inserted into an antecubital vein and the tourniquet loosened immediately after the blood appears in the indicator.
- The vacutainer was depressed onto the other point of the needle for the blood to flow automatically into the tube.
- The needle was then removed smoothly and an adhesive dressing was applied to the puncture site.
- The blood container was mixed well and then labeled with the volunteer's name and number.

2.8.2. Preparation and staining of thin blood smears:

Thin blood films were prepared by the wedge technique immediately after blood collection and will be stained with Lieshman's stain.

2.8.3. Hemoglobin estimation and red cell indices:

Hemoglobin estimation and red cell indices were carried out using the automatic blood cell counter Sysmex KX21N (Japan).

2.8.4. Capillary electrophoresis:

209 samples were subjected to Capillary Electrophoresis (CE) with the Sebia CAPILLARYS 2 piercing fully automated system. CA technology utilizes liquid flow electrophoresis-buffer replaces agarose Gel as the

medium. Haemoglobin variants are separated with the principles of electrophoretic flow at an alkaline pH (9.4) with a negatively charged silica capillary support, and high voltage. Multiple samples undergo an 8-minutes high resolution separation, concurrently. A high resolution haemoglobin separation is obtained, similar to IEF separation. The ideal wavelength of 415 nm is utilized for haemoglobin detection with CE, The result, in electropherogram, is made up of 300 consecutive readings (dots) and is divided into 15 zones. To facilitate interpretation, results are automatically positioned with regard to the Hb A fraction in the sample. Haemoglobins (normal and variants) are displayed as peaks and the zone to which a variant belongs is identified automatically by the system. An on-board haemoglobin library is present in the form of dropdown list and lists all of the normal and variant haemoglobins that may be present within a particular zone.

The CAPILLARYS 2 Sebia automated platform, utilizes the same separation principle. With Sebia's CA system, packed red blood cells are utilized for analysis. Plasma is removed from samples and the bar-coded primary sample tube is loaded into the instrument; all other steps of sample processing and separation are performed automatically by the system.

2.8.5. DNA extraction:

DNA extraction was carried out by the commercial kit provided by Jena Bioscience (Germany) that is based on the Sodium dudisyl sulphate (SDS) method with some modifications according to the regulations of the laboratory of the Sudanese training center in Biotechnology (STCB). The procedure was as follows:

- Firstly 90% Isopropanol and 80% ethanol are prepared that are not provided into the kit.

2.8.5.1. Cell Lysis:

- 300 µl of whole blood is added to a 1.5 ml microtube containing 900 µl RBCs lysis solution and inverted 10 times.
- The mixture is incubated for 3 minutes at room temperature with occasional inversion.
- The tube is then centrifuged for 10 minutes at 15000g.
- The supernatant is then removed with a pipette leaving behind the visible white cell pellet and about 10-20µl of the residual liquid.

- The tube is vortexed vigorously for 10 seconds to resuspend the white cells in the residual liquid.
- 300 µl of cell lysis buffer is added to the resuspended cells and pipetted up and down to lyse the cells until no clumps are visible.
- The tube is then heated at 60 ^{o}C waterbath for 30 minutes.

2.8.5.2. Protein precipitation:

- 100µl protein precipitation solution is added to the cell lysate and vortexed vigorously for 20 seconds to mix well.
- The tube is then centrifuged at 15000g for 10 minutes and the precipitated proteins should form a tight, dark brown pellet.
- In cases where the protein pellet is not tight, vortexing is repeated, followed by incubation on ice for 5 minutes and is centrifuged again.

2.8.5.3. DNA precipitation:

- The supernatant is transferred into a clean 1.5ml microtube containing 300µl 99% Isopropanol and incubated for 30 minutes at -20 ^{o}C.
- The tube is centrifuged at 15000g for 10 minutes (DNA should be visible as a small white pellet).
- The supernatant is discarded and the tube drained briefly on a clean absorbent paper. - 300µl 80% ethanol is added and the tube inverted several times to wash the DNA pellet.
- The tube is centrifuged at 15000g for 10 minutes.
- The supernatant is carefully discarded and the tube dried at room temperature for 15 minutes.

2.8.5.4. DNA Hydration:

- 50-100µl DNA hydration solution is added and the tube vortexed for 5 seconds at medium speed to mix.
- The samples are then incubated at 65 ^{o}C oven for 30 minutes to accelerate hydration.
- The DNA extract is then stored at -20 ^{o}C.

2.8.6. Amplified Refractory Mutation System ARMS-PCR:

First described by Newton and colleagues in 1989, amplification refractory mutations system (ARMS) has become a standard technique that allows the discrimination of alleles that differ by as little as 1 bp. The system is simple, reliable, and nonisotopic. It clearly distinguishes heterozygotes at a locus from homozygotes for either allele. The system requires neither restriction enzyme digestion, nor allele-specific

oligonucleotides as conventionally applied, nor the sequence analysis of polymerase chain reaction (PCR) products. The basis of the system is that oligonucleotides with a mismatched 3′-residue will not function as primers in the PCR under appropriate conditions. A standard ARMS PCR consists of two complementary reactions (two tubes) and uses 3 primers. One primer is constant and complementary to the template in both reactions, and the other primers differ at their 3′ terminal residues and are specific to either the wild-type DNA sequence or the mutated sequence at a given base—only one of these primers is used per tube. If the sample is homozygous mutant or homozygous wild type amplification will only occur in one of the tubes, if the sample is heterozygous amplification will be seen in both tubes (John and David, 2011).

2.8.6.1. Equipments:

Autoclave, Microcentrifuge, Microwave oven, pH meter, weighing balance, water path, Agarose gel Electrophoresis unit, Law voltage electrophoresis power supply, Transilluminator, Incubator, PCR Thermocycler, Vortex mixer, Racks, Eppendorf tubes, droppers, Micropipettes (P10, 20, 50, 100, 1000µl), Sterile DNase free pipette tips, Laboratory Para film.

2.8.6.2. DNA Extraction:

- DNA isolation was done to 55 blood samples, (26) of patients with Sickle Cell Anemia and (29) of patients with β-Thalassemia using DNA extraction kits from Jena Biosciences, Germany.
- Samples showed good DNA quality (on agarose gel electrophoresis).
- The DNA yield was determined spectrophotometrically by using Nanodrop Spectrophotometer.

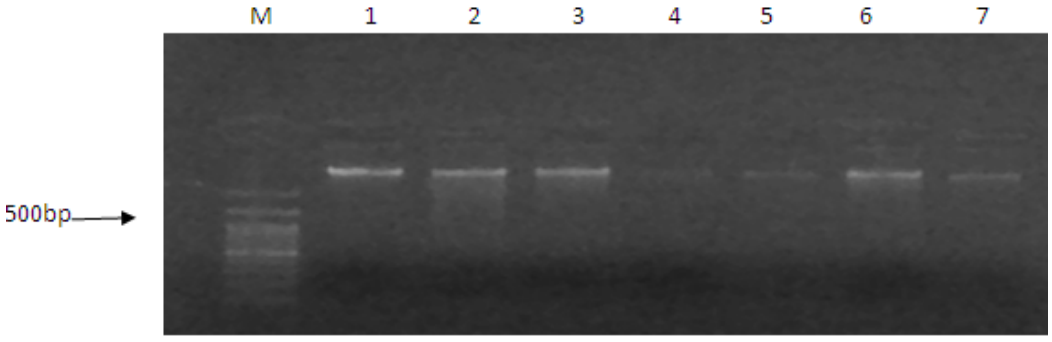


Figure 2-1: Analysis of extracted genomic DNA in 1.5% Agarose gel with 1X TAE buffer:

2.8.6.3. Genotyping of Sickle Cell Anemia patients:

2.8.6.3.1. PCR Analysis:

2.8.6.3.2. Primer design:

Four pairs of primers were designed using the published data for human beta globin chain sequence of the known haplotypes (Cameroon, Benin, Senegal, Arab-Indian and Bantu haplotypes) for sequences to amplify different DNA fragments of the globin S haplotypes, Table (3-1).

- HindII & XmnI for detection of Senegal & Arab-Indian haplotypes
- HindIII for detection of Benin haplotype
- HinfI for detection of Cameroon, Bantu haplotypes and differentiation between Senegal & Arab-Indian haplotypes.

Table 2-1: Enzymes and primers sequences used in this study for sickle cell anaemia:

Restriction enzymes	Primers (´5 - 3')	Product size, bp	Presence of site, bp
1. Xmn1	F. 5' AACTGTTGCTTTATAGGATTTT R. AGGAGCTTATTGATAACCTCAGAC	655	450 205
2.Hind III	F: AGT GCT GCA AGA AGA ACA ACT ACC R: CTC TGC ATC ATG GGC AGT GAG CTC	328	237 91
3. Hinf1´5 to β	F: AGT AGA GGC TTG ATT TGG AGG R: GTT AAG GTG GTT GAT GGT AAC	638	336 302

2.8.6.3.3. Reagents:

- dNTPs
- Primers
- Gotaq polymerase (5u/µl, Promega/MADISON WI, Promega USA).
- 100 bp DNA ladder (Cat. No. G2101, Promega Product Information, Promega)
- Agarose
- Boric Acid
- Tris/base
- EDTA

2.8.6.3.4. Standard PCR reaction:

- The experiment consists of the experimental DNA, a negative control. In a 0.1 µl PCR tube the following solutions were placed in a total volume of 50 µl:
- 10X Taq buffer (final concentration 1X)
- 2.5 mM 4dNTP stock (final concentration 200 µmol)
- 10 pmol/µl primer F
- 10 pmol/µl primer R
- 100 ng of genomic DNA template.
- MgCl2 (final concentration 1.5µm)
- H2O (up to the total volume 0f 50µl)
- 1u Taq. Polymerase (Gotaq polymerase 5u/µl, Promega/MADISON WI USA).

2.8.6.3.5. PCR Programming:

- **Initialization**: This step consists of heating the reaction to a temperature of 94 °C which is held for 5 minutes.
- **1. Denaturation step:** This step is the first regular cycling event and consists of heating the reaction to 94°C for 5minute. It causes DNA melting of the DNA template by disrupting the hydrogen bonds between complementary bases
- **2. Annealing step:** The reaction temperature is lowered to 57 °C for 1minute allowing annealing of the primers to the single-stranded DNA template.
- **3. Extension/elongation step:** The temperature at this step depends on the DNA polymerase used; Taq polymerase has its optimum activity temperature at 72 °C. Final elongation: This single step is occasionally performed at a temperature of 72 °C for 5minutes after the last PCR cycle to ensure that any remaining single-stranded DNA is fully extended.

2.8.6.3.6. Checking of the PCR Products

- To confirm the presence of amplifiable DNA in the samples, The specificity of PCR is typically analyzed by evaluating the production of the target fragment by gel electrophoresis of 5µl PCR products on 2% agarose gel stained with ethidium bromide.

2.8.6.3.6.1. Procedure:

1. Preparation of the 2% Agarose gel:

- 2 gram of agarose

- 100 ml 1X TBE buffer
- Heat the mixture until the agarose completely dissolved, cool to about 60°C and add 4µl of ethidium bromide. Pour the gel on the gel tank and leave it to solidify.

2- Load 5µl PCR products on the gel

3- Run the gel at l00 volts for 20 min,

4- View on the UV system and photograph.

2.8.6.3.6.2. RFLP analysis: Restriction Enzyme Digest Analysis:

This will be done after completion of the PCR amplification. For restriction enzyme analysis, 10 µL of the PCR product will be used. A standard restriction enzyme analysis consists of the following components to a final volume of 15 µL:

- 10 µL of PCR product.
- 0.2 µL (10 U) of restriction enzyme.
- 2.5 µL of 10X buffer
- 12 µL of H_2O. This mixture is incubated at 37ºC for 5-15 minutes followed by enzyme inactivation at 80ºC for 20 minutes, after the incubation is complete; the restriction analysis is carried out in an agarose gel electrophoresis with 1X TAE buffer.

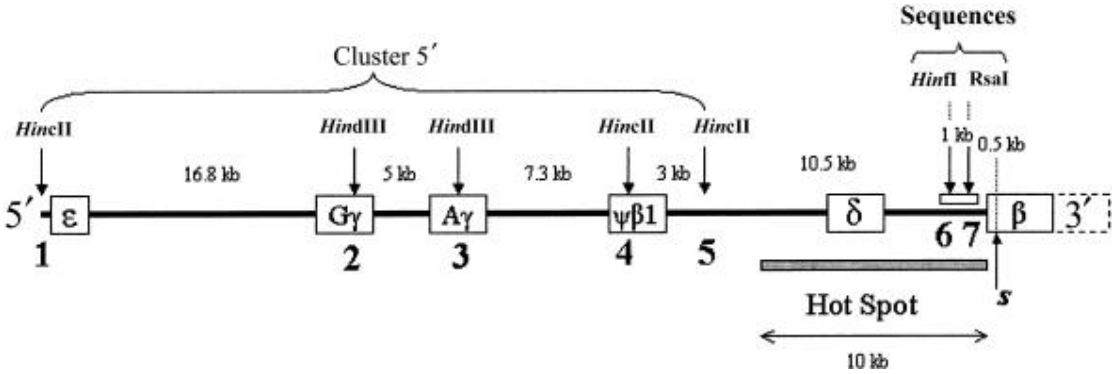


Figure 2-2: Map of the human β-globin gene cluster on chromosome 11. The location of the polymorphic RFLP sites that define the 5′ RFLP cluster haplotypes are shown by arrows and numbers. The sequenced region between positions −1069 and −391 is indicated as a white bar.

Table 2-2: Restriction Fragment Length Polymorphism of patients with sickle cell anemia:

Haplotypes /RE.	β Hind II/Xmn1	Hind III	Hinf1
Senegal	+	+	+
Benin	_	_	+
Cameron	_	+	_
Indian/Arab	+	+	_
Bantu	_	+	+

2.8.6.4. Molecular Testing for β-Thalassaemia:

2.8.6.4.1. ARMS-PCR Analysis:

Primers were designed according to the published ARMS protocol for β - thalassaemia mutations commonly found in African-American and a pair of control primers was also used as internal control for the PC.

Table 2-3: β-thalassaemia Primers sequences used in this study:

Primers (´5 - 3')	Used with	Product size, bp
- African-American mutation primer:		
-88 (C-T) TCA CTT AGA CCT CAC CCT GTG GAG CCT CAT	F	369
-29 (A-G) CAG GGA GGG CAG GAG CCA GGG CTG GGT ATG	F	310
- Control Primers:		
Int-Control A CAA TGT ATC ATG CCT CTT TGC ACC		861
Int-Control B GAG TCA AGG CTG AGA GAT GCA GGA		
Common F CAA TAG GCA GAG AGA GTC AGT GCC TAT CA		

2.8.6.4.2. PCR Standard Mixture:

- PCR reaction mixture was prepared containing the different components, two tubes for each patient (normal and mutant).
- The tubes then placed in PCR Thermo cycler and run using the specific program for ARMS.

2.8.6.4.3. Amplification refractory mutation system (ARMS):

A standard PCR reaction consists of the following components to a final volume of 50 µL:

- 5 µL of 10X PCR buffer.
- 1 µL (10 pm/µL) of forward primer.
- 1 µL (10 pm/µL) of reverse primer.
- 0.25 µL (25 mM) of dNTP mix.
- 1.5 mM (25 mM) $MgCl_2$.
- 1 µL (100 ηg) of DNA.
- 0.25 µL (5 U/µL) of Taq polymerase.
- 40 µL of H2O.

A standard PCR cycling condition consists of the following steps:

- 94°C for 5 minutes.
- 94°C for 30 seconds.
- 55°C for 1 minute.
- 72°C for 45 seconds.
- 72°C for 7 minutes.

2.8.6.4.4. Checking of the PCR Products:

- To confirm the presence of amplifiable DNA in the samples, the specificity of PCR is typically analyzed by evaluating the production of the target fragment by gel electrophoresis of 5µl PCR products on 2% agarose gel stained with ethidium bromide. Visualized using UV transilluminator.

2.8.6.5. Visualization of Test DNA products:

The DNA bands are visualized by transiluumination after Agarose gel electrophoresis according to the following procedure adapted by the Sudan University of Science and Technology research laboratory:

- Samples (DNA extraction) are allowed to thaw to room temperature.
- The appropriate volume of working 1X TAE buffer is prepared from the stock 50X TAE buffer.
- 0.5g of agarose powder is dissolved in 50 ml of the working buffer without mixing.
- The mixture is incubated in the microwave for 5 minutes at medium low temperature.
- The mixture is taken gently and 2 µl ethidium bromide stains is added.
- The electrophoresis machine is the prepared and the Comb is added upright close to the cathode.

- The agar gel is added gently to the tank and allowed to dry for 15 minutes at RT.
- 10 µl of the sample is mixed with 3 µl of loading dye and added carefully to the wells.
- The tank is filled with the working buffer and put on the electric current (100v) for 60 minutes.
- Visualize the DNA bands by the transilluminator.

2.9. Sequencing of β^+ and β^s genes:

DNA purification and standard sequencing was performed for both genes by Macrogen laboratories (Seoul, Korea). Sequence chromatogram was viewed by Finch TV program (http://www.geospiza.com/finchtv.shtml) then the nucleotide sequences of genes were searched for sequence similarity using nucleotide BLAST (http://blast.ncbi.nlm.nih.gov/blast.cgi.) (Eltayeb HN *et al*, 2014). Highly similar sequences were retrieved from NCBI and subjected to multiple sequence alignment using (http://www.phylogen.fr/simple.phylogeny.cgi) software (Dereeper *et al*, 2008 and Dereeper *et al*, 2010).

2.10. Data analysis:

Data analysis was performed using the Statistical Package for Social Science (SPSS) soft ware facilities.

3- Results

This Cross-sectional descriptive study was conducted in Port Sudan city during the period from March 2011 to July 2013 and aimed at the detection and genotyping of haemoglobinopathies (Sickle cell anaemia and thalassaemia) in anemic patients referred to three Haematology laboratories in the three major Hospitals in the city; namely Port Sudan teaching hospital, Sea Ports Corporation hospital and Police Forces hospital. The study included 600 anemic patients aged between 1 year and 85 years old with the mean age being 31 years old. 46% of the patients were females and 54% were males. The mean haemoglobin concentration in the study group was 10.6 g/dl (see table 1). 209 (34.83%) of samples were subjected to capillary electrophoresis testing due to features suggesting haemoglobinopathy such as polychromasia and cell fragments in the thin blood smears. For the rest 391 (65.17%) samples, the mean serum iron concentration was 55.6 mg/dl. The serum iron concentration was normal in 266(67.7%), low in 59 (15.3%) and within lower limits of normal in 66 (16.8%) of the patients. 2.2% of the study population showed features of iron deficiency anaemia and 3.2% showed low serum iron without anaemia. No significant difference (p. value = 0.207) was seen between males and females in the Hb concentration but the difference was significant (p. value = 0.033) between the two groups in the serum iron concentration. The Hb concentration in the study group was neither affected by age nor by sex (p. value= 0.707 and 0.207 respectively. On the other hand, serum iron concentration was not affected by age (p. value= 0.105) but significantly affected by sex (p. value= 0.033). 56.3% of the peripheral blood smears showed normocytic normochromic cells, 38.3% showed microcytic hypochromic cells, 3.3% showed marked anisopoikilocytosis and 2.2% showed microcytosis with a relative numbers and pencil cells. Normal

morphology of white blood cells was seen in 81.7% of cases, leukocytosis in 15.7% and leucopenia in 2.7% while adequate platelets were seen in 82.2% of cases, thrombocytosis in 13.7% and thrombocytopenia in 4.2%.

Table 3-1: Comparison between males and females of the study group in complete blood count parameters and serum iron concentration:

		Mean Values		
NO	**Items**	**Males**	**Females**	**P. value**
1	Hb g/dl	11	10.4	0.207
2	RBCs count mg/dl	4.41×10^{6}	4.2×10^{6}	0.000
3	Serum Iron mg/dl	51	58.2	0.032
4	Hct L/L	35	33.5	0.017
5	MCV Fl	77.6	79.2	0.117
6	MCH Pg	25.7	26.4	0.677
7	MCHC/µl	31.4	31.2	0.211
8	WBCs/µl	8.3×10^{3}	9.2×10^{3}	0.649
9	Platelets count/µl	322×10^{3}	295×10^{3}	0.022

On the other hand, the 209 samples subjected to capillary electrophoresis included 174 (83.3%) patients belonging to Beja tribes and 35 (16.7) patients of other races among whom haemoglobinopathy is known to be frequent as a comparison group. The subjects in the other minor groups were selected according to the same criteria of selection and differ only in that they do not belong to Beja tribes. Regarding complete blood count (CBC) parameters, the only statistically significant difference between the two groups was obtained in leucocytes count (see figure 3-1). The electrophoretic pattern of the test group showed that sickle cell disease (SS) was detected in Bani Amer (2.63%) and Hadandawa (5.26%) while sickle cell trait (AS) was detected in Bani Amer (10.56%), Hadandawa, Attman and Abderahmanab (2.63%) each. Hb D trait (AD) was detected

only in Hadandawa (5.26%). Similarly, Hb E trait (AE) was also detected only in Hadandawa (2.63%). On the other hand, Beta thalassaemia trait was detected in Bani Amer (23.68%), Hadandawa (28.94%), Abderahmanab (7.89%), Almada and Artega (2.63%) each. In the other minor groups, sickle cell disease (SS) was detected in Hawsa (19.06%), Taaesh (14.28%), Fallata (4.76%) and Nuba (4.76%) while sickle cell trait (AS) was detected in Hawsa (19.06%), and Fallata, Foor and Nuba (4.76%) each. Beta thalassaemia trait was also detected in Hawsa, Jaafra, Danagla and Nuba (4.76%) each. Beta thalassaemia major was detected in only one tribe in this group, Rubatab (4.76%). In the test group, the occurrence of haemoglobinopathies was neither affected by sub tribe origin (P. value = 0.59), age (P. value = 0.61), gender (p. value = 0.62), nor by residence (p. value = 0.13). Similarly, in the other minor groups it didn't affected neither by tribe, age, gender nor by residence (P. values = 0.33, 0.18, 0.46, and 0.87 respectively).

Table 3-2: Comparison between Beja tribes and other minor groups based on the results of complete blood count.

Parameter	Results of Beja tribes (Mean ± 1 SD)	Results of minor groups (Mean ± 1 SD)	P. value
Leucocytes count /µl	8400 ± 6405	11812 ± 8899	0.01
Erythrocytes count /µl	3910896 ± 2532548	4016562 ± 2836488	0.84
Haemoglobin concentration g/dl	8.84 ± 2.30	8.91 ± 2.75	0.87
Packed cell volume l/l	28.25 ± 7.66	26.53 ± 7.61	0.24
Mean cell volume fl	73.06 ± 14.03	78.23 ± 14.75	0.05
Mean cell haemoglobin pg	24.04 ± 5.66	25.63 ± 5.30	0.14
Mean cell haemoglobin concentration g/dl	31,97 ± 3.55	32.65 ± 2.94	0.31
Thrombocytes count /µl	298615 ± 232131	346343 ± 154401	0.25

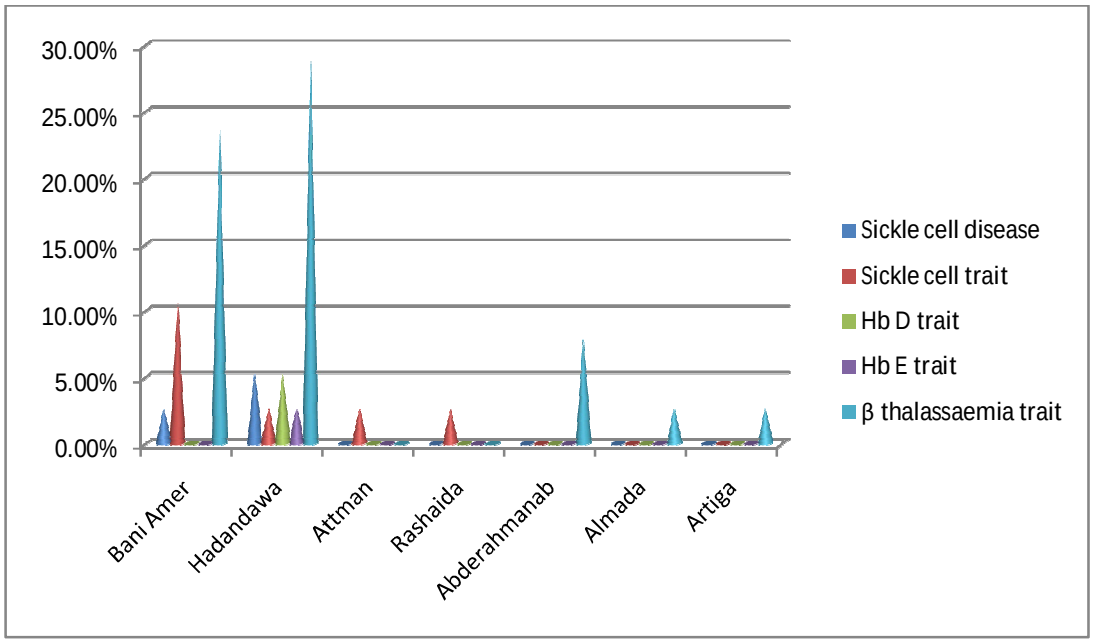


Figure 3-1: Electrophoretic results of Beja tribes.

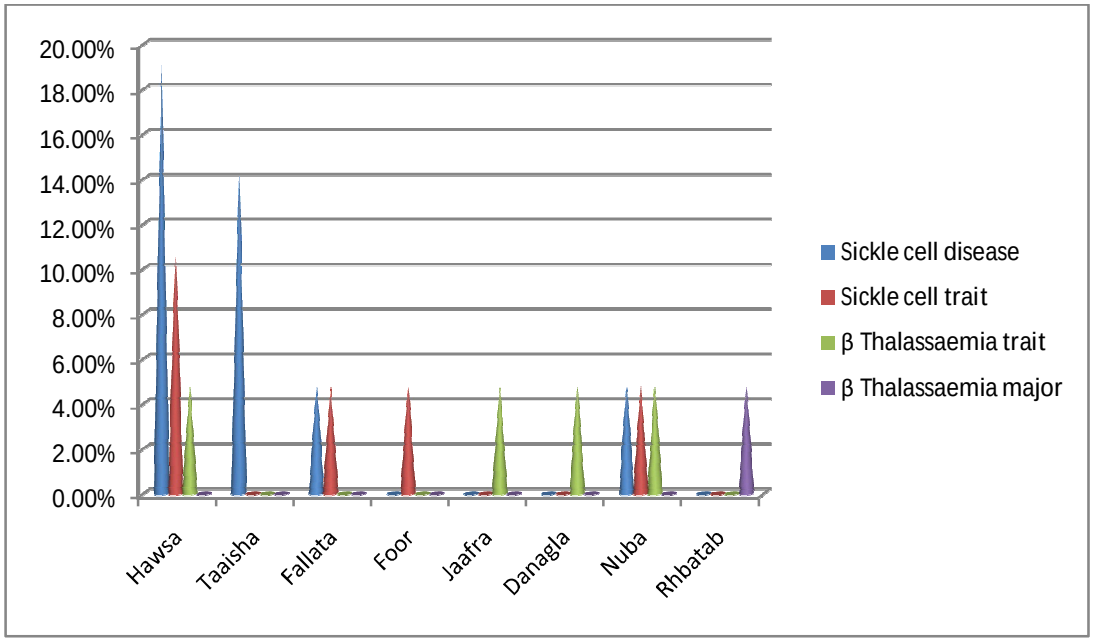
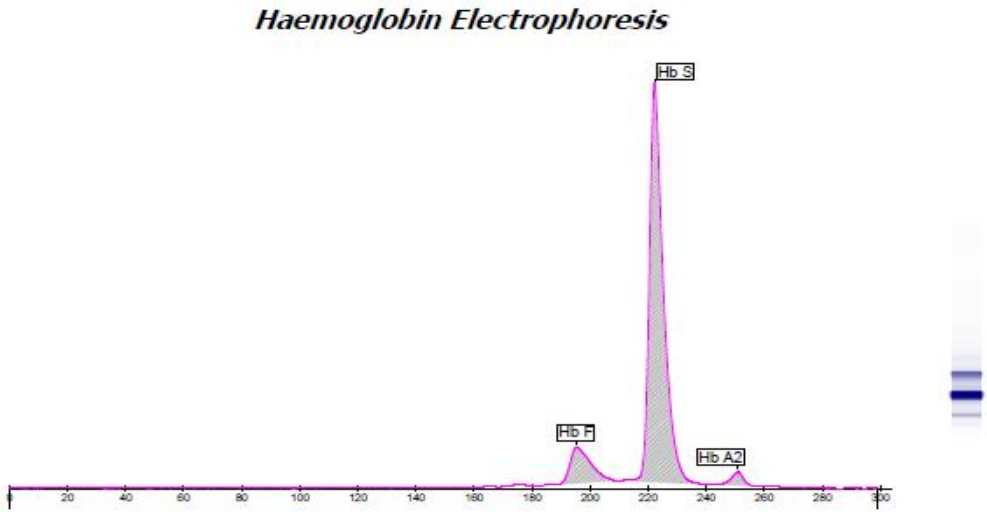
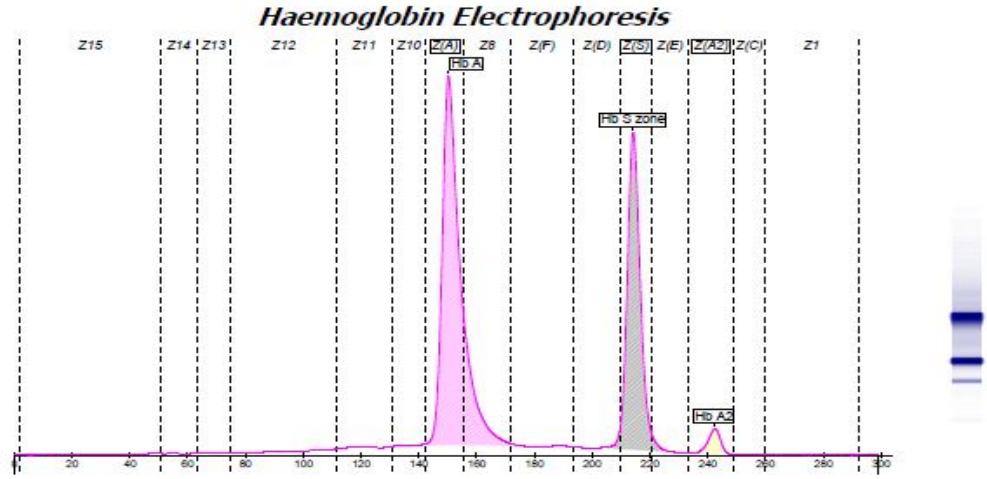


Figure 3-2: Electrophoretic results of tribes of the minor groups.



Name	%	Normal Values %
Hb F	10.5	
Hb S	87.1	
Hb A2	2.4	

Figure 3-3: Capillary electrophoretic pattern of a 46 years old male belonging to Bani Amer tribe showing a typical pattern of sickle cell disease (SS).



Name	%	Normal Values %
Hb A	60.7	
Hb S zone	36.6	
Hb A2	2.7	

Figure 3-4: Capillary electrophoretic pattern of a 31 years old female belonging to Hawsa tribe showing a typical pattern of sickle cell trait (AS).

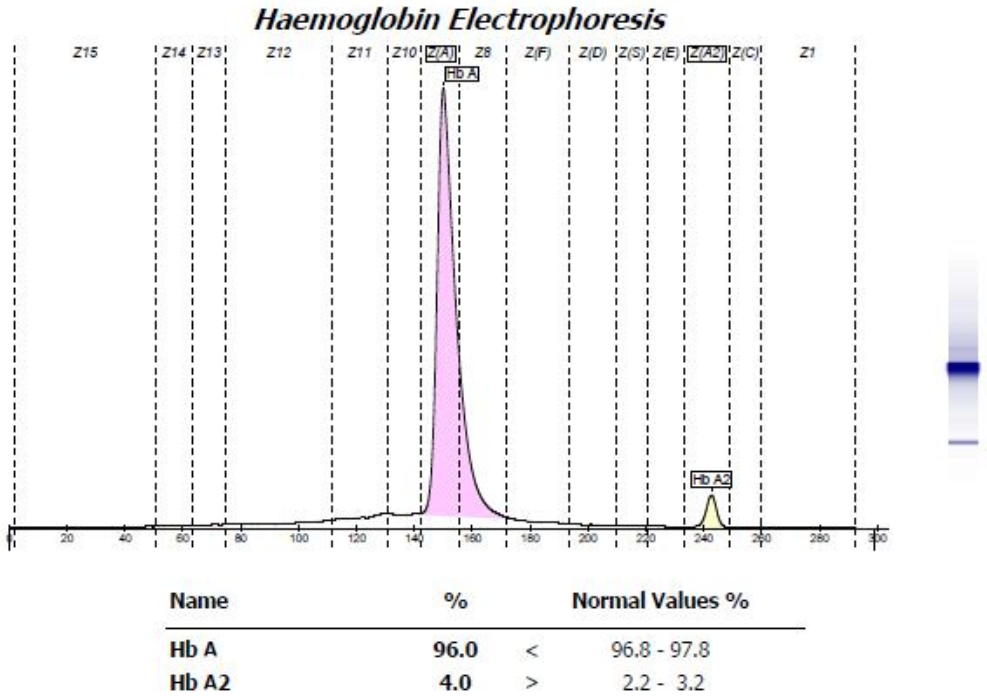


Figure 3-5: Capillary electrophoretic pattern of a 44 years old female belonging to Hadandawa tribe showing a typical pattern of Beta thalassaemia trait.

Table 3-3: Quantity of different Haemoglobins among patients subjected to Capillary Electrophoresis:

Patients with:	**Hb-A** (mean ± 1SD)	**Hb-A2** (mean ± 1SD)	**Hb-F** (mean ± 1SD)	**Hb-D** (mean ± 1SD)	**Hb-E** (%)	**Hb-S** (mean ± 1SD)
Normal Hb pattern	97.2 ± 1	2.6 ± 0.46	0.27 ± 0.7	0	0	0
Sickle cell trait	65±13.7	3.1 ± 0.52	2.9 ± 9.4	0	0	30±17.1
Sickle cell disease	0	3.1± 0.53	10 ± 5.6	0	0	87.1±5.2
Hb-D trait	93.4 ± 4	2.6 ± 9.21	0.45± 0.06	3.4±4.1	0	0
Hb-E trait (%)	95.8	2.7	0.9	0	0.6	0
β-thalassaemia trait	95.9±1.6	3.9 ± 1	0.26±0.76	0	0	0

The ARMS-PCR strategy was used to identify the tow common Afro-American (-29 and -88) Beta thalassaemia mutations in all patients (29) that showed positive results in the Capillary electrophoresis. Primer concentrations for each mutation were adjusted in the reaction in such a way that there was no possibility of obtaining false-positive results. -29 (A⟶G) and -88 (C⟶T) mutations in the promoter region were simultaneously amplified in a single tube with the common reverse primer F, thereby giving 310- and 369-bp bands, respectively. An 861-bp internal control band was also amplified in all the reactions, indicating successful PCR. No band other than the internal control was amplified in the negative control reaction, which was carried out with a dried blood spot normal for β-thalassaemia.

All the 29 samples produced a 369-bp band indicating the presence of -88 (C⟶T) β-thalassaemia mutations (Figure 3-6).

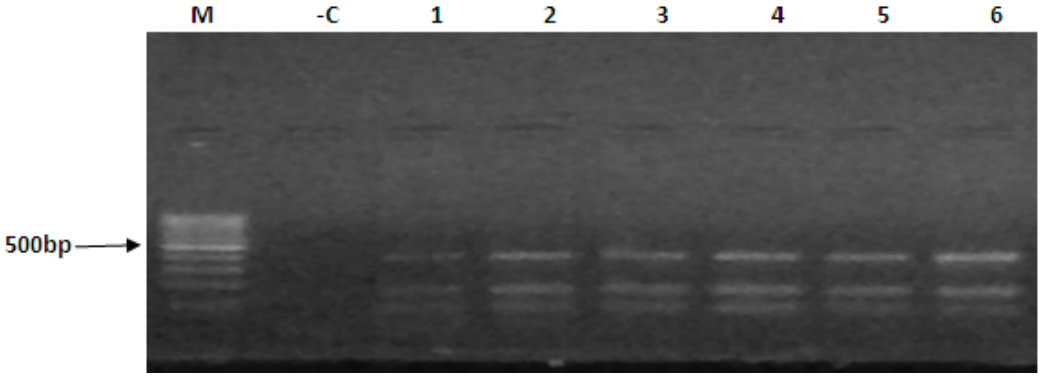


Figure 3-6: Ethidium bromide-stained agarose gel for multiplex ARMS-PCR for -88 (C⟶T) β-thalassaemia mutation (369 bp). A negative control can be seen next to the marker.

Table 3-4: Allele frequency for β-thalassaemia mutations among the study population:

Mutation	Allele frequency
-88 (C⟶T)	(29)100%
-29 (A⟶G)	(00) 0.0%

Table 3-5: Genotype distribution for β-thalassaemia mutations among the study population:

Genotype	Distribution
-88/-88	100%
-88/-29	0.0%
-29/-29	0.0%

Table 3-6: Tribal distribution of the -88 (C→T) β-thalassaemia mutations among the study population:

Tribe	Number (male: female)	Percent (%)
Hadandawa	11 (5:6)	37.9
Bani Amer	10 (8:2)	34.5
Abderahmanab	3 (1:2)	10.3
Armada	1 (1:0)	3.4
Hawsa	1 (1:0)	3.4
Nuba	1 (1:0)	3.4
Rubatab	1 (1:0)	3.4
Danagla	1 (1:0)	3.4
Total	**29 (19:10)**	**100**

The results obtained from sequencing of beta thalassaemia samples clearly confirmed that all samples possessed the -88 (C⟶T) mutation (see figure 3-7).

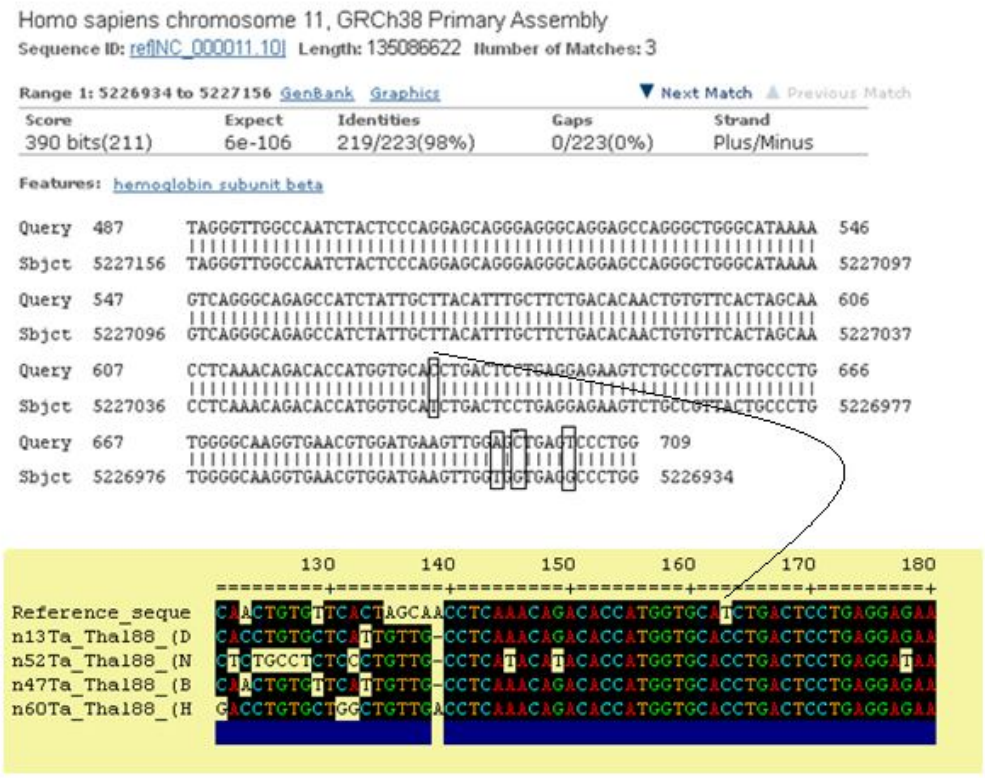


Figure 3-7: The -88 (C⟶T) mutations in four beta thalassaemia samples as compared to reference genomic sequence.

The phylogenetic tree (figure 4-8) indicated that the different tribal groups acquired the mutation independently and it seemed to have appeared in Bani Amer and Nuba long before Hadandawa and Danagla.

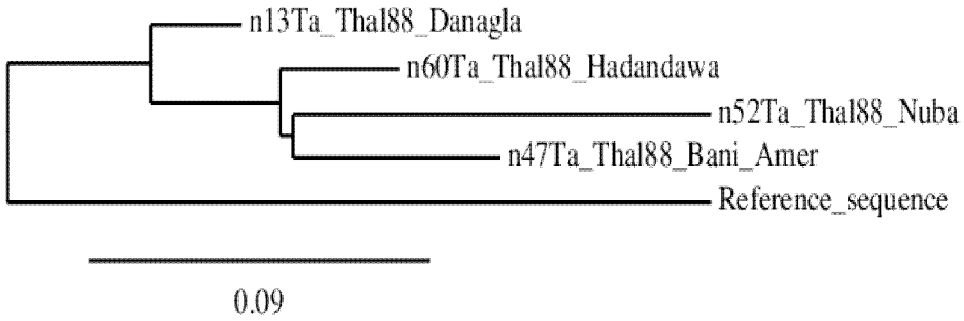


Figure 3-8: Phylogenetic tree of the beta thalassaemia mutation in four tribal groups.

Hb S was detected in 26 (12.91%) of whom 15 (55.6%) were heterozygous AS and 12 (44,4%) were homozygous SS. RFLP for these samples indicated that all were Benin haplotype (Figures 3-9 to 3-12).

Table 3-7: Tribal distribution of sickle cell anemia among the study population:

Tribe	SS:AS	Percent (%)
Hadandawa	2:1	11
Bani Amer	1:4	22
Taaisha	3:1	15
Fallata	1:1	7
Hawsa	4:4	30
Nuba	1:1	7
Attman	0:1	4
Rashaida	0:1	4
Total	**26**	**100**

On the other hand, sequencing tests were also carried out for selected samples with sickle cell hemoglobin. The sequences obtained for Hind III PCR products yielded no significant similarities when compared to the reference genomic strands on the BLAST, For Hinf I PCR products, comparison to the reference sequence has indicated the presence of the cutting site (GANTC) that is recognized by Hinf I enzyme (Figure 3-9). Three recognition sites are seen in that figure while the fourth one has been diminished due to a single base pair substitution in the query strand (A⟶ G).

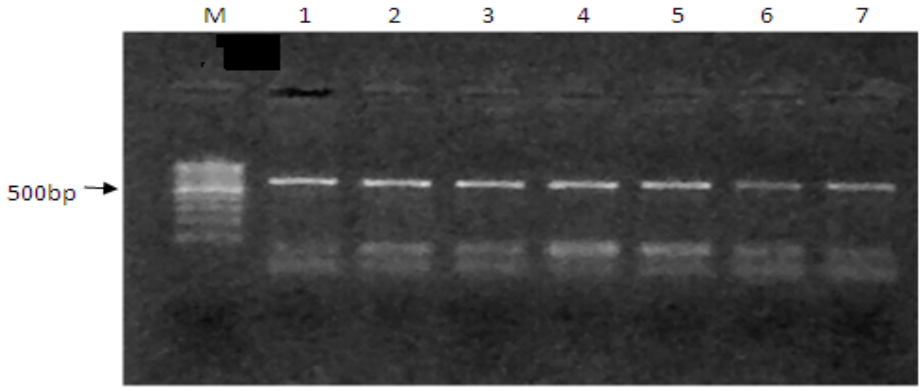


Figure 3-9: PCR/FRLP amplification product of Hinf I primers before digestion. Notice the 638 bp band in all samples.

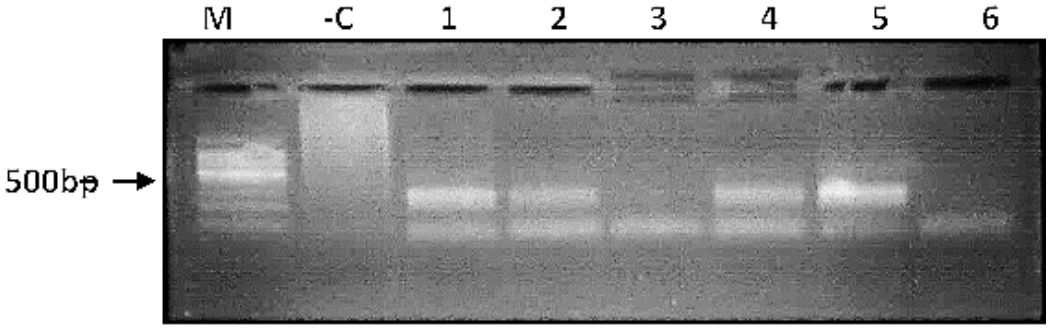


Figure 3-10: PCR amplification product of Hinf I enzyme digestion.

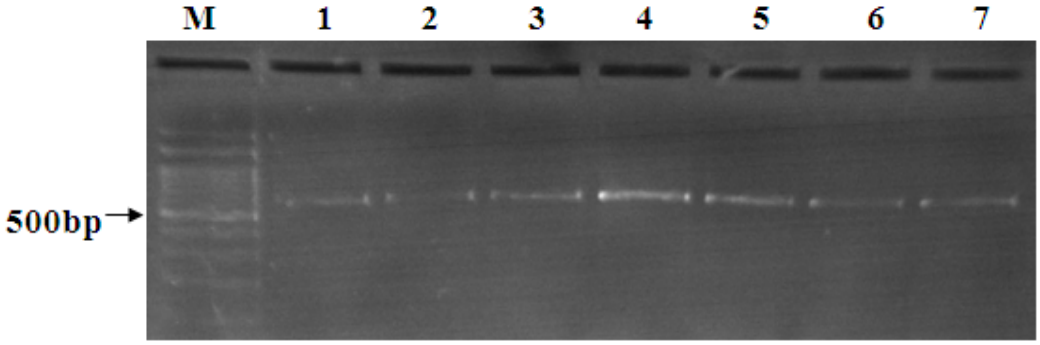


Figure 3-11: PCR/RFLP amplification product of XmnI primers before digestion. Notice the 638 bp band in all samples.

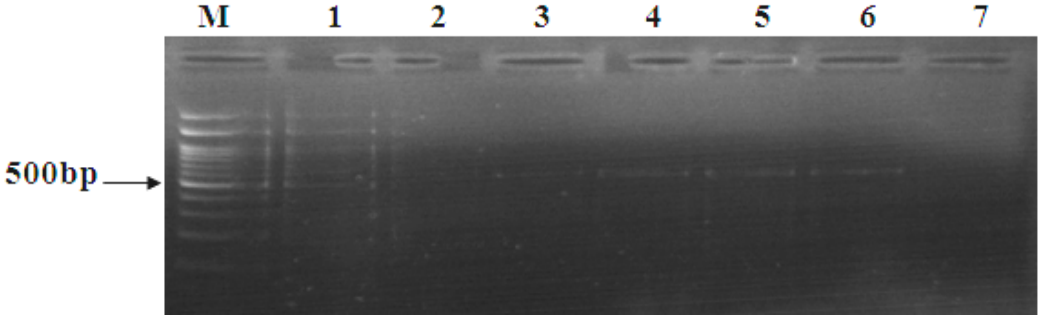


Figure 3-12: Figure 4-10: PCR amplification product of XmnI enzyme digestion.

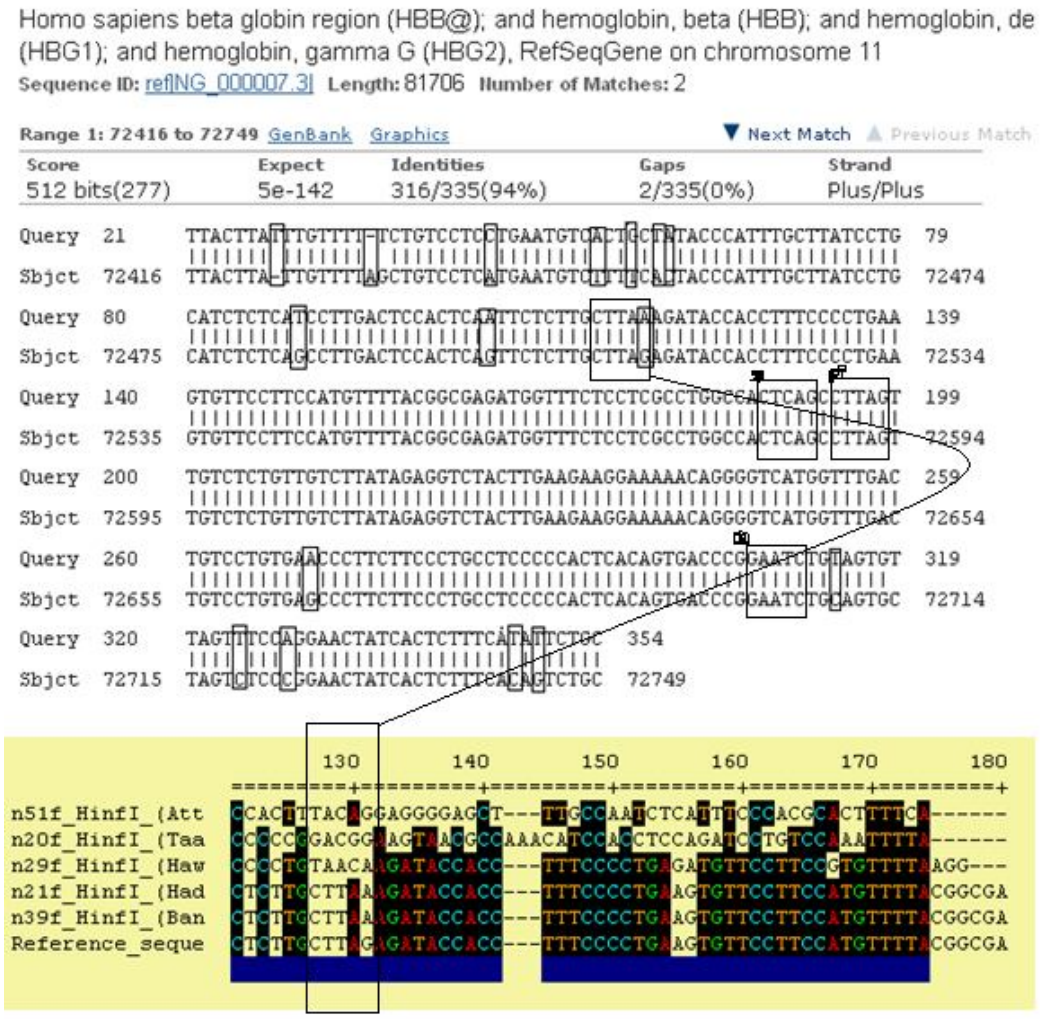


Figure 3-13: Sequencing alignment for Hinf I PCR product. The upper sample is from a patient belonging to Bani Amer Tribe. 15 base pair substitution can be seen as compared to the reference sequence.

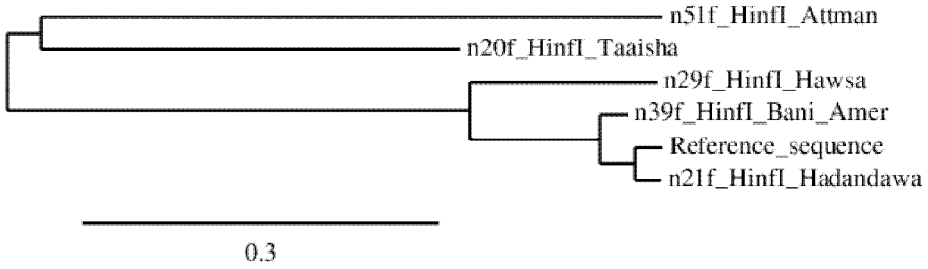


Figure 3-14: Phylogenetic tree of sickle cell (Hinf I) PCR product in five tribal groups as compared to the reference sequence.

4- Discussion

Little is known till now about the occurrence and manifestations of haemoglobinopathies in the population of Eastern Sudan, namely Red Sea State. That is because haemoglobinopathies are rare conditions and their diagnosis is consistently more difficult especially in heterozygote subjects. So, paucity of previous studies on the prevalence of haemoglobinopathies further complicates the matter. Nutritional anaemia is also still a worldwide problem especially among children and women in the child bearing age. Dela cruz *et al*, 2012, studying the nutritional causes of anaemia in Mexican children under 5 years reported that the overall prevalence of anaemia was 20.6 % of which 14 % were mild cases and 6.38 % were moderate. And that anaemia was associated with iron deficiency in 24.17 % of the case, whereas iron deficiency coexisted with either folate or vitamin B_{12} deficiency in 9 %. He also reported that only 2 % of cases of anaemia were associated with either folate or B_{12} deficiencies. These findings are agreed with the results obtained in our study. Ronald H. Lands 2009, concluded that the prevalence of anaemia increases with advancing age; under age 75 years anaemia is more common in females, but over age 75 years it is more common in males and that anaemia increases dramatically after age 85 in both sexes. Rohrig, 2012, has studied the anaemia and iron deficiency in elderly; Prevalence, diagnostics and new therapeutic options in Germany reported that the prevalence of anaemia can reach 40 %. Again these finding are agreed with our findings. Rojroomwasinkul, 2013, in his research entitled nutritional status and dietary intakes of 0.5-12 years old Thai children, reported that the prevalence of anaemia in rural areas was twice as those in urban areas. This is beyond the scope of our present study. Rohner, 2013, studying the infant and young child feeding practices in urban Philippines and their associations with stunting, anaemia, and deficiencies of iron and vitamin A, stated that 42 percent were anaemic and 28 percent were iron deficient These findings differ significantly from data obtained by the current study may be due to our smaller sample size.

Globally, it has been shown from previous studies that the incidence of Beta thalassaemia gene in Africa has been variable. In many developing countries the haemoglobinopathies (thalassaemias and sickle-cell disorder) are so common that they provide a convenient model for

working out a genetic approach to control of chronic childhood diseases. At present, about 250 million people (4.5% of the world population) carry a potentially pathological haemoglobinopathy gene. Each year about 300 000 infants are born with major haemoglobinopathies. Owing to increasing global migration, however the haemoglobinopathies have appeared in many areas where they were not endemic (Angastiniotis *et al*, 1995). Bain, 2006, reported that the prevalence of Hb S in Sudan is less than 1-17%, Hb C is 0% and Beta thalassaemia is 1-10% (overall 4%). These figures are completely agreed with what we have found in our current study where Hb C is not detected in our patients and the 29 patients with beta thalassaemia trait represents only 4.8% of the patient population (Table 3-6, page 65). The absence of Hb C in the Sudanese population was later diminished by Elderdery et al, 2014 who has found that two schoolgirls aged 11 and 16 were admitted to the Military Hospital in Omdurman on the 2nd of April 2009. They had Hb (haemoglobin) SC and their parents were carriers of abnormal Hb; the mother was (Hb AS) and father (Hb AC). These girls were the first reported cases of Hb SC in Sudan. Elderdery *et al*, 2011, studying the tribal distribution of Haemoglobinopathies in a Sudanese patient population, reported that the S gene was the most common variant found (6.1%) and was most prevalent in the Western tribes in Sudan (12.%). This was agreed with our current findings and further potentiates our reliance on patients from Western Sudan descent for comparison. Another study, SCT Among relatives of sickle cell patients in western Sudan, carried out by Munsoor and Afaf, 2011, reported that 54% of target samples were heterozygous carriers (Hb AS) while 42% were normal (Hb AA) and 4% were diagnosed as sickle cell disease. The significantly higher proportion than ours (Table 3-7 page 66) of sickle cell trait in this study and in the study entitled "Haemoglobinopathies in the Sudan" carried out by Elderdery *et al*, 2008 and indicated that Hb S [β6(A3) Glu→Val] was the most common abnormal Hb, is due to that their study included a "known" population of patients with the disease. Elderdery *et al*, 2012, studying the Molecular analysis of the beta-globin gene cluster haplotypes in a Sudanese population with sickle cell anaemia, has reported that four of the five typical βS-globin haplotypes were identified. The most frequent was the Cameroon (35.0%), followed by the Benin (29.4%), the Senegal (18.2%) and the Bantu (2.8%). The Indian-Arab

haplotype was not observed. Three atypical haplotypes were identified in 17 patients, occurring at a combined frequency of 14.6%. One of these, found at the high frequency of 11.8%, possibly represented a new Sudan haplotype. The same author further concluded that βS Haplotyes were demonstrated successfully from dried blood samples and that a new haplotype is apparent in Sudan, in addition to the four African haplotypes. In our study, all patients with sickle cell anemia were of Benin haplotype and that may possibly indicate that the mutant gene was introduced by male immigrants of African descent rather than being autochthonous. Genetic polymorphism related to the multi-diverse make up of the Sudanese population is clearly indicated when our findings are compared with those of Podhorodecka *et al*, 2011, who reported that the S mutation in the HBB gene was not detected among the Shagia and Manasir individuals investigated and that lack of HB-S mutation in the HBB gene in the previously unstudied Shagia group confirms that the frequency of sickle cell gene in Sudan is tending to decrease in a northerly direction. Kolita, 2010, in his review about guidelines for the diagnosis of the haemoglobinopathies in Nigeria, mentioned that the prevalence of homozygous SCD (Hb SS) and (Hb S+C) disease is 3.1% and 1.1% respectively thus giving a combined prevalence of 4.1% while the trait (HB AS) has a prevalence of 23.7%. Again the prevalence of sickle cell trait was significantly higher than our findings and that is because the S gene is much more prevalent in West African populations (Barbara, 2006). Elgari *et al*, 2014, studying the Hematological Characteristics in Sudanese adults with sickle cell disease in Khartoum state, have found that the frequency distribution of sickle cell phenotypes of study group on basis of hemoglobin electrophoresis were SS 29(36.2%), AS 49(61.2%), SC 1(1.3%), SF 1(1.3%). The mean hemoglobin concentration, mean packed cell volume, mean cell volume (MCV), mean cell hemoglobin (MCH) and mean cell hemoglobin concentration (MCHC) were 9.2 ± 3.2 gm/dl, 30.2 ± 9.2%, 76.6 ± 7.9/fl, 23.2 ± 3.0/pg, 30.1 ± 2.3gm/dl. Significant lower hemoglobin (p value <0.05), The mean TWBC, RBC and platelet counts were 11.1 x 109/L ± 8.6, 3.9 x 1012/L ± 1.3, 450 x 109/L ± 15.8, significant higher TWBC and lower RBC counts of p value <0.05 for both. The findings indicate for moderate to severe anemia. Their findings agreed with ours (Table 3-2 page 60) except for the presence of Hb SC which is absent in our patient

population. Globally, there are more carriers of thalassaemia than of SCD, but the high frequency of the sickle-cell gene in certain areas leads to a high birth rate of homozygotes. As a result, SCD accounts for about 70% of haemoglobin disorders worldwide (Angastiniotis *et al*, 1995). Since the frequencies of beta thalassaemia trait and sickle cell disease displayed by this study are derived from patient population, we can hardly say that the frequencies can be slightly lower considering the whole community, but the problem is still there. The bands of Hb D (0.6 and 0.2%) and Hb E (0.3%) detected in three Beja patients by Capillary Electrophoresis were not consistent with the establishment of the diagnosis of the correspondent hemoglobin variants (see table 1-9, page 39) and molecular analysis was not carried out for these samples.

The results obtained from sequencing of mutant genes for both sickle cell anemia (Figures 3-13 and 3-14 page 68) and beta thalassaemia (figure 3-7 page 65 and figure 3-8 page 66) suggest that those genes were loaded to the target population through the migration of many tribes of African descent into the area early in the beginning of the last century rather than malaria endemicity for that, although Beja tribes reside by the "Red Sea", the environmental conditions are not suitable for the vector of malaria transmission.

Conclusion:

1- The data obtained from this study clearly indicates that nutritional anaemia is still a major problem in red sea state with the main cause of nutritional anaemia being iron deficiency anaemia which is relatively prevalent among children.

2- Haemoglobinopathies (haemoglobin variants and thalassaemia, brought about by the migration of many tribes of African descent into the Sudan during the early years of the eighteenth century) is not uncommon amongst the population of Beja tribes and other minor ethnic groups inhabiting the study area.

Recommendations:

1- It is strongly recommended to establish a governmental centre for screening, diagnosis and prognosis of Haemoglobinopathy in Sudan.

2- The implementation of population based educational programmes to raise the community awareness about the haemoglobinopathies is a very necessary issue.

3- Neonatal and premarriage screening for Haemoglobinopathy should be regarded as a great concern.

4- All patients with low values of haemoglobin, MCV and MCH in the target group should be screened for Haemoglobinopathy regardless their iron profile results.

5- Community based studies that search and screen for unknown mutations in the HBB gene are necessary to explain the large number single base pair substitutions encountered during this study.

6- Further studies that aim at screening of Haemoglobinopathy in different Sudanese populations are needed to complete our local data base table.

References

Abbas MY (2014). Haematological parameters in Sudanese children with sickle cell disease. American Journal of Research Communication, 2(2): PP. 20-32.

Abouh AME and Abdalla MHA (2014). D-dimer level in Sudasese children with sickle cell anemia. International Journal of Current Research. 6 (05): PP.6599-6601.

Ajjack EA, Awooda HA and Abdalla SE (2014). Haemoglobin Patterns in Patients with Sickle Cell Haemoglobinopathies. International Journal of Hematological Disorders, 1(1), 8-11.

Alizadeh S, Bavarsal MS, Dorgalaleh A, Khatib ZK, Dargahi H, Nassiri N, Hamid F, Rahim F, Jaseb K and Saki N (2014). Frequency of beta-thalassaemia or beta-haemoglobinopathy carriers simultaneously affected with alpha thalassaemia in Iran. Clin. Lab. 60(6): 941-9.

Angastiniotis M, Modell B, Englezos P and Bouleyjenkov V (1995). Prevention and control of haemoglobinopathies. Bulletin of the World organization, 73: 375-386.

Bakri MH, Ismail EA, Elsedfy GO, Amr MA, and Ibrahim A (2015). Behavioral impact of sickle cell disease in young children with repeated hospitalization. Saudi Journal of Anesthesia, 8(4): 504-509.

Ballinger A and Patchett S (2007). Clinical Medicine. Fourth edition. London, Saunders ELSEVIER. P. 200.

Bain BJ, Bates I, Laffan MA, and Lewis SM (2012). Dacie and Lewis Practical Haematology. 11th edition. U.K: Elsevier Churchill Livingstone,PP. 301-330.

Bain JB (2006). Haemoglobinopathy Diagnosis. Second edition, Oxford UK: Blackwell publishing. PP. 70, 223.

Balne NG and Rao C.S.V. R (2013). Role of XmnI restriction site polymorphism and JAK2 gene mutation in β-Thalassaemia. Int. Res. J. Biological Sci. 2(1): 41-45.

Bashir N, Barkawi M, Sharif L, Mamoni A and Gharaibeh N (1992). Prevalence of Haemoglobinopathies in North Jordan. Trop. Geog. Med. 44(1-2): 122-125.

Brant A (2010). Haemoglobinopathy and thalassaemia detection: Traditional methods and a Novel method – Capillary Electrophoresis Technology. Seperations keeping you in track with the latest in EP. 1(11): 1.

Cantan D (2014). Thalassaemias and Haemoglobinopathies in Turkey. Hemoglobin. 17: 1-3.

Dassanayake RS, Mahadevan K and Gunawardene Y (3013). Development of single-strand conformational polymorphism to screen for mutations in hotspot regions of Beta-globin gene of Beta-thalassaemia patients of Sri Lanka. SoutheaSt aSian J trop Med public health. 44 (1). PP. 115-125.

Day RA (1996). How to write and publish a scientific paper. Fourth edition. United Kingdom, Cambridge university press.

Dirar MS(1984). The History of Habab and Hamasyin Tribes in Sudan and Eretria. First edition. Khartoum: The Sudan Publishing House. PP. 16-38.

Dereeper A., Guignon V., Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J.F., Guindon S., Lefort V., Lescot M., Claverie J.M., Gascuel O. *Phylogeny.fr: robust phylogenetic analysis for the non-specialist.* Nucleic Acids Res. 2008 Jul 1;36(Web Server issue):W465-9. Epub 2008 Apr 19. (PubMed) : joint first authors.

Dereeper A., Audic S., Claverie J.M., Blanc G. *BLAST-EXPLORER helps you building datasets for phylogenetic analysis.* BMC Evol Biol. 2010 Jan 12;10:8. (PubMed).

Elderdery AY, Mohamed BA, Karsani ME, Ahmed MH, Knight G and Cooper AJ (2008). Hemoglobinopathies in the Sudan. Hemoglobin. 32:323-6.

Elderdery, Abozer Y, Mohamed BA, Cooper AJ, Knight G and Mills J (2011). Tribal distribution of Haemoglobinopathies in a Sudanese patient population. J. Med. Lab. Diag. 2 (4): PP. 31-37.

Elderdery AY, Mills J, Mohamed BA, Cooper AJ, Mohammed AO, Eltieb N and Old J (2012). Molecular analysis of the beta-globin gene cluster haplotypes in a Sudanese population with sickle cell anaemia. Int J Lab Hematol; 34:262-6.

Elderdery AY, Abozer Y, Mohamed BA, AL- Mijalli SH and Mills J (2014). First recorded case of haemoglobin SC in Sudan. Hematology and Leukemia, ISSN 2052-434X: 10.7243/2052-434X-2-2.

Elgari MM, Ahmed H A, Younis M S, Waggiallah H A (2014). Hematological Characteristics in Sudanese adult with sickle cell disease in Khartoum state. J Am Sci; 10(8):8-11.

Firkin F, Chesterman C, Penington D and Rush B (1989). de Gruchy's Clinical Haematology in Medical practice. Fifth edition. United Kingdom, Blackwell Science. PP. 137-170.

Gaafar TM, ELBeshlawy AM, Aziz MI and Abdelrazik HN (2006). Rapid screening of β-Globin gene mutations by Real-Time PCR in Egyptian thalassemic children. Afr J Health Sci. 13: 70-77.

Gabig-Cimińska M, Kloska A and Malinowska M (2013). Human Genetic Disorders. Part 2 – Diagnosis and Treatment. Kapital Ludzki, Unia Europejska. 8-16.

Galanello R and Offit R (2010). Beta thalassaemia. Orphanet J. Rare Dis. 5: 11.

Gongora DG (2012). Nutritional causes of anaemia in Mexican children of under 5 years: results from the 2006 National Health and Nutrition survey. Salud Paplica Mexico: 54(2): pp. 108-115.

Greer JP, Foerster J, and Lukens JN (2003). Wintrobe's Clinical Haematology. 11th edition. London: Lippinkott Williams and Willkins publishers. P 1019.

Hassan YF (1989). An introduction to the History of Islamic Kingdoms in Eastern Sudan. Third edition. Sudan: Khartoum University Press. PP. 3-10.

Hewitt R, Krause A, Goldman A, Campbell G, and Jenkins T (1996). β-Globin Haplotype Analysis Suggests That a Major Source of Malagasy Ancestry Is Derived from Bantu-Speaking Negroids. Am. J. Hum. Genet. 58:1303-1308.

Hoffbrand AV, Catovsky D and Edward GD (2005). Postgraduate haematology. Fifth edition, Volume 1. United Kingdom, Blackwell publishing. PP. 85-103.

Ibrahim SA and Mustafa D (1967). Sickle-cell haemoglobin O disease in a Sudanese family. Br Med J. 3:715-7.

Kaddah N, Rizk S, Kaddah AM, Salama K and Lotfi H (2009). Study of possible genetic factors determining the clinical picture of thalassaemia intermedia. J. Med. Sci. 9 (3): 151-155.

Karkun A and Bhagat S (2011). Diagnosis of sickle cell anemia by Dde 1 enzyme based RFLP. The Bioscan. 6(4): 575-578.

Kraemer K and Zimmermann MB (2007). Nutritional Anaemia. Switzerland, Sight and Life press. PP. 4-17.

Kolita TP (2010). Guidelines for the diagnosis of the haemoglobinopathies in Nigeria. Ann. Ibadan Postgrad. Med. 8 (1): 25.

Lewis SM, Bain BJ and Bates I (2006). Practical Hematology. 10th edition. United Kingdom: Church hill Livingstone. PP. 272-303.

Marder VG and Conley CL (1959). Electrophoresis of hemoglobin on Agar Gel, Frequency of Hemoglobin D in a Negro population. Bulletin of the Johns Hopkins Hospital (105). PP. 77-88.

Modell B, Khan M, and Dorlison M (2000). Survival in Beta thalassaemia major in the UK: data from the UK thalassaemia register. Lancet. 355 (9220): 2051-2052.

Mohammed AO, Attalla B, Bashir FM, Ahmed FE, El Hassan AM, Ibnauf G, Jiang W, Cavalli-Sforza LL, Karrar ZA and Ibrahim ME

(2006). Relationship of the sickle cell gene to the ethnic and geographic groups populating the Sudan. Commun. Genet. 9(2): 113-120.

Munsoor MM and Afaf A (2011). SCT Among relatives of Sickle cell patients in Western Sudan. Can. J. of Med. 2 (2): 20-26.

Nagar R, Sinha S and Raman R (2014). Haemoglobinopathies in eastern Indian states: a demographic evaluation. J. Commun. Genet. 1-8.

Nerweiy FF, AL-Allawins, Jubrael J and Dawood RS, (2010). The Application of the Amplification Refractory Mutation System (ARMS) for characterization of β-Thalassaemia mutations in Duhok. Duhok Med J;4(2):8-20.

Ohaj MA (1986). From the History of Beja. First edition. Sudan: Khartoum University Press. P. 8-21.

Old JM. DNA Diagnosis of Hemoglobin Mutations. In: Nagel RL (2003). Hemoglobin Disorders: Molecular Methods and Protocols. Humana Press Inc., Totowa, NJ. PP. 101-116.

Old J, Harteveld CL, Traeger-Synodions J, Petruo M, Angastiniotis M and Ganalleno R (2012). Prevention of thalassaemias and other Hemoglobin disorders. Second edition; volume 2: Laboratory protocols. Thalassaemia international federation press, Cyprus. PP. 3-7.

Olivieri N (1998). Thalassaemia: Clinical management. Clinical Haematology 11: 147-162.

Omati CE (2005). Beta thalassaemia trait in Nigerian patients with sickle cell anaemia. J. Biomed. Sc. 4(1): P. 40.

Parker JN and Parker PM (2007). Beta Thalassaemia. ICON Group International, Inc. USA. PP 3-17.

Paul AA (1954). History of the Beja tribes of the Sudan. London: Cambridge. P. 64.

Podhorodecka AK, Knap OM, Parczewski M, Kuleta AB and Ciechanowicz A (2011). Sickle cell anaemia-associated Beta-globin

mutation in Shagia and Manasir tribes from Sudan. Pol. J. of Environ. 20 (6): PP 1225- 1530.

Provan D, Singer CRJ, Baglin T and Lilleyman J (2004). Oxford Hand Book of Clinical Haematology. Sesond edition, U. K. Oxford University press. PP. 77-94.

Ranzo G and Raffaella O (2010) . Beta thalassaemia. Orphan. J. Rare Dis. 5: P. 8-13.

Rohrig G(2012). Anaemia and iron deficiency in the elderly. Prevalence, diagnostic and new therapeutic options. 2 Gerontol Geriatr. 45(3). PP. 191-96.

Rohner F, Woodruff BA, Aaron GJ, Yakes EA, Lebanon MA, Rayco-Solon P and Saniel OP (2013). Infant and young child feeding practices in urban Philippines and their association with stunting, anaemia, and deficiencies of iron and vitamin A. Food and nutrition Bull, 34(2): PP. 73-81.

Rojroongwasinkul N, Kijboonchoo K, Wimonpeerapattana W, Purttiponthanee S, Yamborisut U, Boonpraderm A, Kunapan P, Thasanasuwan W and Khouw I (2013). The nutritional status and dietary intakes of 0.5-12 years old Thai children. British Journal of nutrition, 110(3): PP. 36-44.

Roudknar MH, Najmabadi H, Derakhshandeh P and Farhud DD (2003). Detection of Rare and Unknown Mutations in β- tathalassaemia Traits in Iran. Iranian J Pub Health, 32(1), pp.11-14.

Sabahelzain MM and Hamamy H (2014). The ethnic distribution of sickle cell disease in Sudan. Pan African medical Journal, 18:13 doi:10.11604/pamj.2014.18.13.3280.

Schneider RG. Identification of Hemoglobin by Electrophoresis. In: Schmdt RM, Haisman THJ, Lehmann H (1974). The Detection of Haemoglobinopathies. Cleveland, OH: CRP press. PP. 9-28.

Steinberg MH (1998). Pathophysiology of sickle cell disease. Clinical Haematology. 11: 163-180.

Serjeant GR (1992). Sickle cell disease. New York, Oxford University Press. PP. 88-116.

Steinberg M.H. (2001). Disorders of hemoglobin. New York, Cambridge University Press. PP. 95-146.

Ware RE (2013) Is Sickle Cell Anemia a Neglected Tropical Disease? PLoS Negl Trop Dis 7(5): e2120. doi:10.1371/journal.pntd.0002120.

Weatherall DJ (2004). The role of molecular genetics in an evolving global health problem. Am. J. Hum. Genet. 74(3): 385-392.

Weatherall DJ and Clegg JB (2001). Inherited Hemoglobin Disorders: an increasing Global Health Problem. Bulletin of the World Health Organization. 79: 704-11.

Weatherall DJ and Clegg JB (2001). The thalassaemia syndromes. Oxford, Blackwell Science.

Williams TN, Mwangi TW, Wambua1 S, Neal D. ND, Kortok M, Snow RW, and Marsh K (2005). Sickle Cell Trait and the Risk of Plasmodium falciparum Malaria and Other Childhood Diseases. J Infect Dis. 192(1): 178–186.

Appendix

Appendix I

Questionnaire

No.:……

Name:…… Age:……

Tribe:…… Sex:……

Residence:…… Occupation:……

Marital status:……

History of previous hemolytic tendency: Yes ☐ No ☐

If yes, specify:……

History of previously diagnosed hemoglobinopathy: Yes ☐ No ☐

If yes, specify:……

Hemoglobin profile:

Hemoglobin g/dl	
Packed cell volume L/L	
Mean cell volume Fl	
Mean cell hemoglobin Pg	
Mean cell hemoglobin concentration g/dl	

Hemoglobin Electrophoresis:

	Cellulose Acetate		Agar Gel	
	Band	**Quantity%**	**Band**	**Quantity%**
Hemoglobin A				
Hemoglobin A_2				
Hemoglobin F				
Hemoglobin S				
Hemoglobin E				
Hemoglobin C				
Others ( )				
Others ( )				
Others ( )				

Research Consent

أقر أنا ..... بأنني قد تبرعت لأن أكون أحـــد أفراد عينة هذه الدراسة الخاصة بدراسة أنماط خضـــاب الدم لدى قبائل البجـــا. و رضيت بأن تجمع مني عينة دم (2.5 سى سى) يتم استخدامها فقط لغرض البحث العلمي.

التوقيع : .....

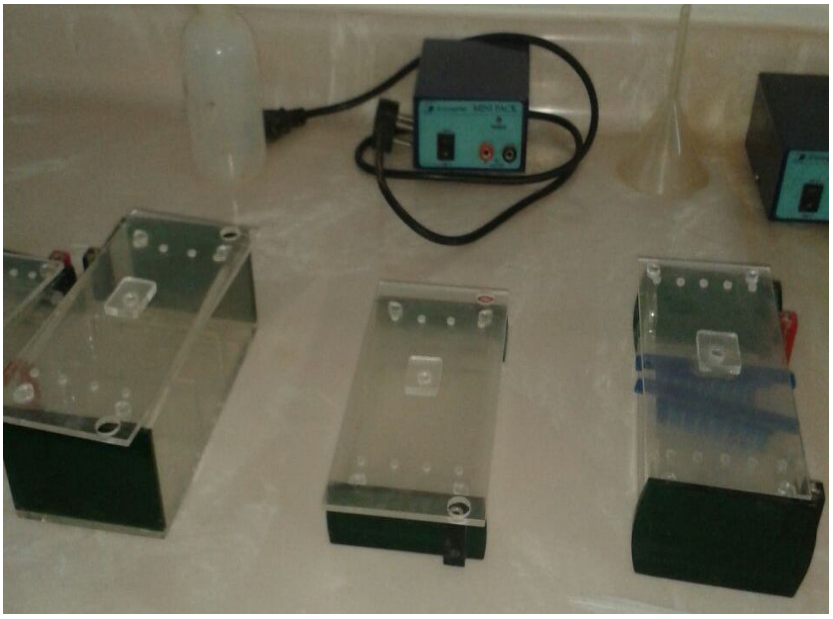
التأريخ: .....

Appendix II

Apparatus:

1- Cold centrifuge TGL 60M, Made in china.

2- Gel electrophoresis Aristogene MINIPACK Mfd 090227, SI No. P 50575, Germany.



3- Labcare SC-R Recirculating class II Microbiological Safety Cabinet. To BS 5726:1992 and BS EN 12469:2000.

4- Microwave, Palson Sapphire Microondas microwave COD: 30534, MOD: WP 700 J-520 Germany.

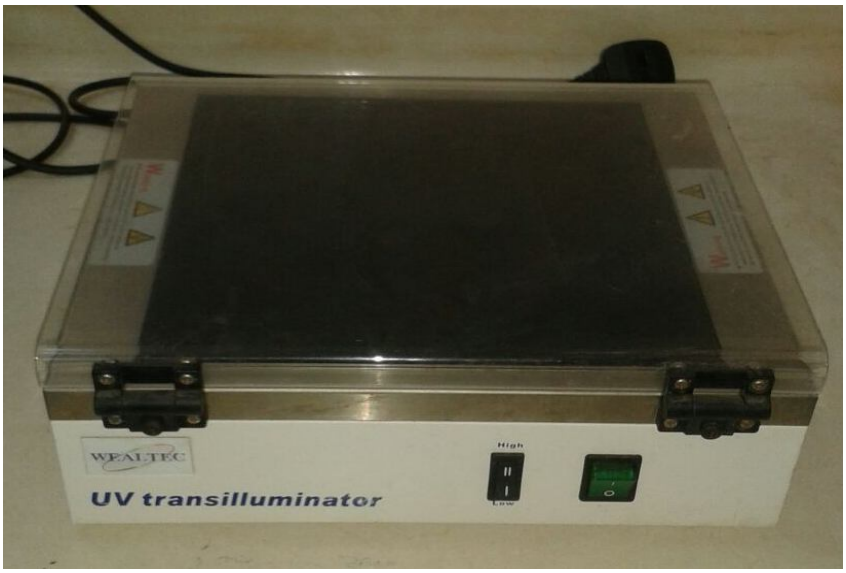
5- Oven, BTC Biotech company for medical and laboratory equipments, Made in Egypt.

6- Pipettes, accupipet, U 39576, U 39370, U 38522, U 39811.

7- Sensitive Balance, Sartorius BL 310, AG GOTTENGEN Germany 13509567.

8- Sysmex KX 21N. Serial No. 87151, Japan.

9- UV transilluminator, WEALTEC Model MD-25, 312nm, S No. UBW 250522 made in Taiwan.



10- Vortex mixer, LMV 202, T 0155155312/11 Scott Science UK.

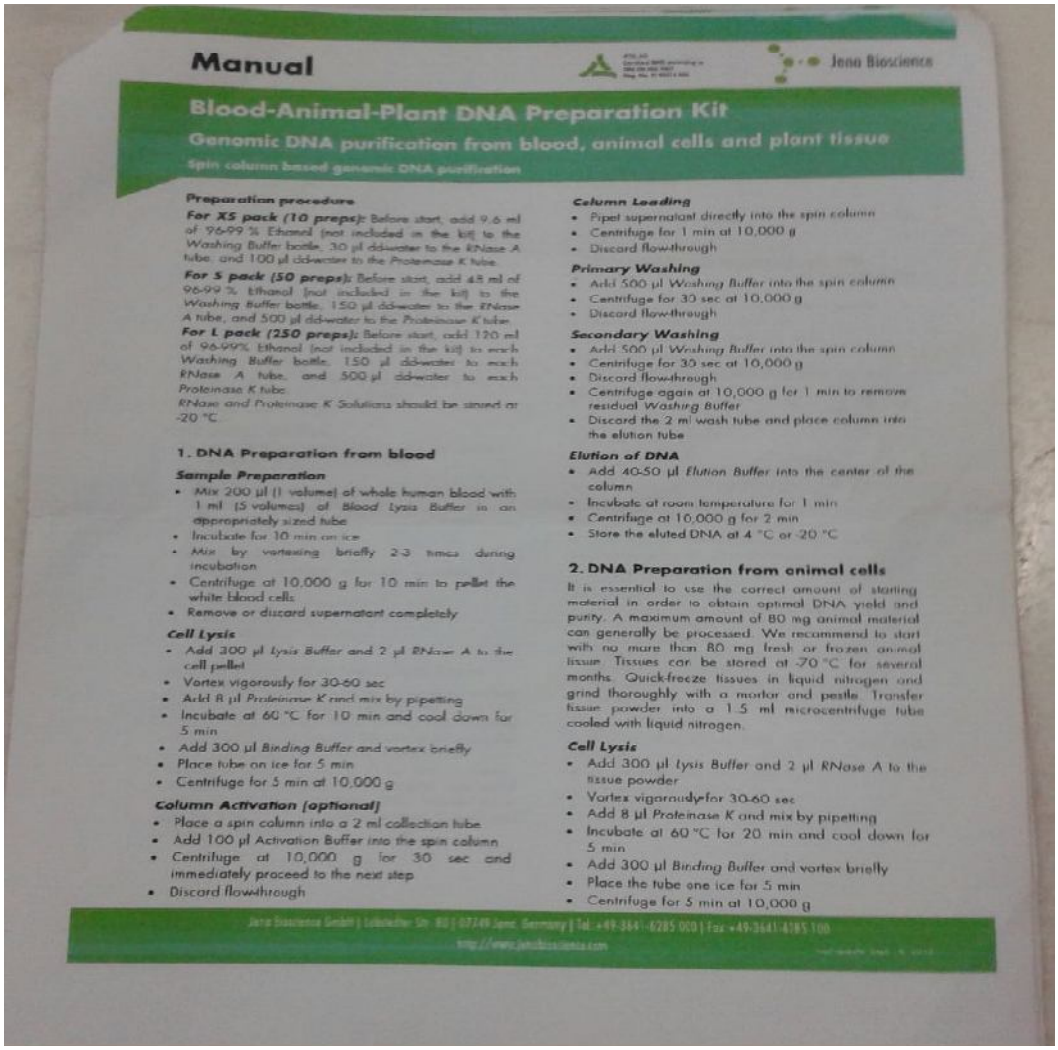
11- Thermo cycler, Convergent technologies GmbH & Co. KG. SN: CT-TCL-205, Germany.

Reagents:

Agarose powder (nuclease free), BCL—001M, Lot # 90003.

Buffer stock, 50X TAE BCC-013, Lot # 131011.

DNA extraction kit:



Manual

Jena Bioscience

Blood-Animal-Plant DNA Preparation Kit

Genomic DNA purification from blood, animal cells and plant tissue

Spin column based genomic DNA purification

Preparation procedure

For XS pack (10 preps): Before start, add 9.6 ml of 96-99 % Ethanol (not included in the kit) to the *Washing Buffer* bottle, 30 µl ddwater to the *RNase A* tube, and 100 µl ddwater to the *Proteinase K* tube.

For S pack (50 preps): Before start, add 48 ml of 96-99 % Ethanol (not included in the kit) to the *Washing Buffer* bottle, 150 µl ddwater to the *RNase A* tube, and 500 µl ddwater to the *Proteinase K* tube.

For L pack (250 preps): Before start, add 120 ml of 96-99% Ethanol (not included in the kit) to each *Washing Buffer* bottle, 150 µl ddwater to each *RNase A* tube, and 500 µl ddwater to each *Proteinase K* tube.

RNase and Proteinase K Solutions should be stored at -20 °C.

1. DNA Preparation from blood

Sample Preparation

- Mix 200 µl (1 volume) of whole human blood with 1 ml (5 volumes) of *Blood Lysis Buffer* in an appropriately sized tube
- Incubate for 10 min on ice
- Mix by vortexing briefly 2-3 times during incubation
- Centrifuge at 10,000 g for 10 min to pellet the white blood cells
- Remove or discard supernatant completely

Cell Lysis

- Add 300 µl *Lysis Buffer* and 2 µl *RNase A* to the cell pellet
- Vortex vigorously for 30-60 sec
- Add 8 µl *Proteinase K* and mix by pipetting
- Incubate at 60 °C for 10 min and cool down for 5 min
- Add 300 µl *Binding Buffer* and vortex briefly
- Place tube on ice for 5 min
- Centrifuge for 5 min at 10,000 g

Column Activation (optional)

- Place a spin column into a 2 ml collection tube
- Add 100 µl Activation Buffer into the spin column
- Centrifuge at 10,000 g for 30 sec and immediately proceed to the next step
- Discard flowthrough

Column Loading

- Pipet supernatant directly into the spin column
- Centrifuge for 1 min at 10,000 g
- Discard flowthrough

Primary Washing

- Add 500 µl *Washing Buffer* into the spin column
- Centrifuge for 30 sec at 10,000 g
- Discard flowthrough

Secondary Washing

- Add 500 µl *Washing Buffer* into the spin column
- Centrifuge for 30 sec at 10,000 g
- Discard flowthrough
- Centrifuge again at 10,000 g for 1 min to remove residual *Washing Buffer*
- Discard the 2 ml wash tube and place column into the elution tube

Elution of DNA

- Add 40-50 µl *Elution Buffer* into the center of the column
- Incubate at room temperature for 1 min
- Centrifuge at 10,000 g for 2 min
- Store the eluted DNA at 4 °C or -20 °C

2. DNA Preparation from animal cells

It is essential to use the correct amount of starting material in order to obtain optimal DNA yield and purity. A maximum amount of 80 mg animal material can generally be processed. We recommend to start with no more than 80 mg fresh or frozen animal tissue. Tissues can be stored at -70 °C for several months. Quick-freeze tissues in liquid nitrogen and grind thoroughly with a mortar and pestle. Transfer tissue powder into a 1.5 ml microcentrifuge tube cooled with liquid nitrogen.

Cell Lysis

- Add 300 µl *Lysis Buffer* and 2 µl *RNase A* to the tissue powder
- Vortex vigorously for 30-60 sec
- Add 8 µl *Proteinase K* and mix by pipetting
- Incubate at 60 °C for 20 min and cool down for 5 min
- Add 300 µl *Binding Buffer* and vortex briefly
- Place the tube one ice for 5 min
- Centrifuge for 5 min at 10,000 g

Jena Bioscience GmbH | Löbstedter Str. 80 | 07749 Jena, Germany | Tel.: +49-3641-6285 000 | Fax +49-3641-6285 100
http://www.jenabioscience.com

Appendix III
Communications

CALL FOR PERMISSION (3)

People
me Hello dear professor Swee Lay Thein, I'm a Sudanese student at Sudan University of Science and Technology. I'm performing a study entitled "Molecular analysis of haemoglobinopathy in Beja tribes". I w
Dec 19 at 12:40 PM
Thein, Swee
To
me
CC
barnaby.clark@nhs.net
Dec 20 at 1:56 AM

Hello,

With pleasure, you can use the figure with our permission.

Kind Regards,

Swee Lay Thein
Professor of Molecular Haematology / Consultant Haematologist
King's College London School of Medicine / King's College Hospital
James Black Centre
125 Coldharbour Lane
London SE5 9NU
UK
email: sl.thein@kcl.ac.uk
swee.thein@nhs.net
Tel: 44-(0)20-7848 5443/5445
44-(0)20-3299 31679/31689
Fax: 44-(0)20-7848 5444
Reply, Reply All or Forward | More
me
To
Thein, Swee
Today at 2:02 AM
I really thank you very much professor
With my best wishes.
Mohamed Omer Gibreel
Hematology department, Port Sudan Ahlia College, Sudan
Reply, Reply All or Forward | More
Click to reply all

Figures

Citation: Ware RE (2013) Is Sickle Cell Anemia a Neglected Tropical Disease? PLoS Negl Trop Dis 7(5): e2120. doi:10.1371/journal.pntd.0002120

Published: May 30, 2013

Funding: No funding was received to produce this work.

Competing interests: The author has declared that no competing interests exist.

To
myamya88@hotmail.com my.mohammed@sau.edu.sa
Today at 4:13 AM
Dear Dr Mohammed Yousif Abbas
I am performing a project entitled " Genotyping of haemoglobinopathy in Sudanese Beja tribes" at the Hematology department, College of graduate studies, Sudan University of science and technology.
I am looking forwards to kindly offer me a permission to use your "Figure 1. Map of Sudan. The red circles demonstrate the distribution of tribes with high prevalence of SCD." that is illustrated in your article entitled
Haematological parameters in Sudanese children with sickle cell disease.
With all my best regards

Yours sincerely, Mohammed Omer Abubaker Gibreel
Coordinator of MLCs, Port Sudan Ahlia College
E-mail: m_omer8164@yahoo.com
Phone number: 00249912908164.
Dr. Mohammed Yousif Abbas
To me
Today at 5:35 AM
Dear Dr. Mohammed Omer,

I hope all success for you, and I give you the permission to use the mentioned figure. I am also ready for any help you want. all the best

Yours sincerely

Date: Tue, 3 Mar 2015 12:13:38 +0000
From: m_omer8164@yahoo.com
To: myamya88@hotmail.com; my.mohammed@sau.edu.sa
Subject: CALL FOR PERMISSSION

Appendix IV

The HBB gene sequence

TGATGGTCTTAGTATAGCTTGCAGCCTTGTCCCTGCAGGGTATTATGGGTAATAGAAAGA
AAAGTCTGCGTTACACTCTAGTCACACTAAGTAACTACCA**TTGGAAAAGCAACCCCTGCC**
TTGAAGCCAGGATGATGGTATCTGCAGCAGTTGCCAACACAAGAGAAGGATCCATAGTTC
ATCATTTAAAAAAGAAAACAAAATAGAAAAAGGAAAACTATTTCTGAGCATAAGAAGTTG
TAGGGTAAGTCTTTAAGAAGGTGACAATTTCTGCCAATCAGGATTTCAAAGCTCTTGCTT
TGACAATTTTGGTCTTTCAGAATACTATAAATATAACCTATATTATAATTTCATAAAGTC
TGTGCATTTTCTTTGACCCAGGATATTTGCAAAAGACATATTCAAACTTCCGCAGAACAC
TTTATTTCACATATACATGCCTCTTATATCAGGGATGTGAAACAGGGTCTTGAAAACTGT
CTAAATCTAAAACAATGCTAATGCAGGTTTAAATTTAATAAAATAAAATCCAAAATCTAA
CAGCCAAGTCAAATCTGCATGTTTTAACATTTAAAATATTTTAAAGACGTCTTTTCCCAG
GATTCAACATGTGAAATCTTTTCTCAGGGATACACGTGTGCCTAGATCCTCATTGCTTTA
GTTTTTTACAGAGGAATGAATATAAAAAGAAAATACTTAAATTTTATCCCTCTTACCTCT
ATAATCATACATAGGCATAATTTTTTAACCTAGGCTCCAGATAGCCATAGAAGAACCAAA
CACTTTCTGCGTGTGTGAGAATAATCAGAGTGAGATTTTTTCACAAGTACCTGATGAGGG
TTGAGACAGGTAGAAAAAGTGAGAGATCTCTATTTATTTAGCAATAATAGAGAAAGCATT
TAAGAGAATAAAGCAATGGAAATAAGAAATTTGTAAATTTCCTTCTGATAACTAGAAATA
GAGGATCCAGTTTCTTTTGGTTAACCTAAATTTTATTTCATTTTATTGTTTTATTTTATT
TTATTTTATTTTATTTTGTGTAATCGTAGTTTCAGAGTGTTAGAGCTGAAAGGAAGAAGT
AGGAGAAACATGCAAAGTAAAAGTATAACACTTTCCTTACTAAACCGACATGGGTTTCCA
GGTAGGGGCAGGATTCAGGATGACTGACAGGGCCCTTAGGGAACACTGAGACCCTACGCT
GACCTCATAAATGCTTGCTACCTTTGCTGTTTTAATTACATCTTTTAATAGCAGGAAGCA
GAACTCTGCACTTCAAAAGTTTTTCCTCACCTGAGGAGTTAATTTAGTACAAGGGGAAAA
AGTACAGGGGGATGGGAGAAAGGCGATCACGTTGGGAAGCTATAGAGAAAGAAGAGTAAA
TTTTAGTAAAGGAGGTTTAAACAAACAAAATATAAAGAGAAATAGGAACTTGAATCAAGG
AAATGATTTTAAAACGCAGTATTCTTAGTGGACTAGAGGAAAAAAATAATCTGAGCCAAG
TAGAAGACCTTTTCCCCTCCTACCCCTACTTTCTAAGTCACAGAGGCTTTTTGTTCCCCC
AGACACTCTTGCAGATTAGTCCAGGCAGAAACAGTTAGATGTCCCCAGTTAACCTCCTAT
TTGACACCACTGATTACCCCATTGATAGTCACACTTTGGGTTGTAAGTGACTTTTTATTT
ATTTGTATTTTTGACTGCATTAAGAGGTCTCTAGTTTTTTATCTCTTGTTTCCCAAAACC
TAATAAGTAACTAATGCACAGAGCACATTGATTTGTATTTATTCTATTTTTAGACATAAT
TTATTAGCATGCATGAGCAAATTAAGAAAAACAACAACAAATGAATGCATATATATGTAT
ATGTATGTGTGTATATATACACACATATATATATATATTTTTTCTTTTCTTACCAG**AAGG**
TTTTAATCCAAATAAGGAGAAGATATGCTTAGAACCGAGGTAGAGTTTTCATCCATTCTG
TCCTGTAAGTATTTTGCATATTCTGGAGACGCAGGAAGAGATCCATCTACATATCCCAAA
GCTGAATTATGGTAGACAAAACTCTTCCACTTTTAGTGCATCAACTTCTTATTTGTGTAA
TAAGAAAATTGGGAAAACGATCTTCAATATGCTTACCAAGCTGTGATTCCAAATATTACG
TAAATACACTTGCAAAGGAGGATGTTTTTAGTAGCAATTTGTACTGATGGTATGGGGCCA
AGAGATATATCTTAGAGGGAGGGCTGAGGGTTTGAAGTCCAACTCCTAAGCCAGTGCCAG
AAGAGCCAAGGACAGGTACGGCTGTCATCACTTAGACCTCACCCTGTGGAGCCACACCCT
AGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGCTGGGCATAAAAG
TCAGGGCAGAGCCATCTATTGCT**TACATTTGCTTCTGACACAACTGTGTTCACTAGCAAC**
CTCAAACAGACACCATGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGT
GGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGTTGGTATCAAGGT
TACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAGACTCTT
GGGTTTCTGATAGGCACTGACTCTCTCTGCCT**ATTGGTCTATTTTCCCACCCTTAGGCTG**
CTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCT
GATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTT
AGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTG
CACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGACGCTTGA
TGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTCATAGGAAGGGGATAAGTAAC
AGGGTACAGTTTAGAATGGGAAACAGACGAATGATTGCATCAGTGTGGAAGTCTCAGGAT
CGTTTTAGTTTCTTTTATTTGCTGTTCATAACAATTGTTTTCTTTTGTTTAATTCTTGCT
TTCTTTTTTTTTCTTCTCCGCAATTTTTACTATTATACTTAATGCCTTAACATTGTGTAT
AACAAAAGGAAATATCTCTGAGATACATTAAGTAACTTAAAAAAAAACTTTACACAGTCT
GCCTAGTACATTACTATTTGGAATATATGTGTGCTTATTTGCATATTCATAATCTCCCTA
CTTTATTTTCTTTTATTTTTAATTGATACATAATCATTATACATATTTATGGGTTAAAGT
GTAATGTTTTAATATGTGTACACATATTGACCAAATCAGGGTAATTTTGCATTTGTAATT
TTAAAAAATGCTTTCTTCTTTTAATATACTTTTTTGTTTATCTTATTTCTAATACTTTCC

CTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCATGCCTCTTTGCACCATTCT
AAAGAATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATCTCTGCATATAAATAT
TTCTGCATATAAATTGTAACTGATGTAAGAGGTTTCATATTGCTAATAGCAGCTACAATC
CAGCTACCATTCTGCTTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAG
CTAGGCCCTTTTGCTAATCATGTTCATACCTCTTATCTTCCTCCCACAG**CTCCTGGGCAA**
CGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCACCCCACCAGTGCAGGC
TGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCACTAAGC
TCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTA
AACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTA
TTTTCATTGCAATGATGTATTTAAATTATTTCTGAATATTTTACTAAAAAGGGAATGTGG
GAGGTCAGTGCATTTAAAACATAAAGAAATGAAGAGCTAGTTCAAACCTTGGGAAAATAC
ACTATATCTTAAACTCCATGAAAGAAGGTGAGGCTGCAAACAGCTAATGCACATTGGCAA
CAGCCCCTGATGCATATGCCTTATTCATCCCTCAGAAAAGGATTCAAGTAGAGGCTTGAT
TTGGAGGTTAAAGTTTTGCTATGCTGTATTTTACATTACTTATTGTTTTAGCTGTCCTCA
TGAATGTCTTTTCACTACCCATTTGCTTATCCTGCATCTCTCAGCCTTGACTCCACTCAG
TTCTCTTGCTTAGAGATACCACCTTTCCCCTGAAGTGTTCCTTCCATGTTTTACGGCGAG
ATGGTTTCTCCTCGCCTGGCCACTCAGCCTTAGTTGTCTCTGTTGTCTTATAGAGGTCTA
CTTGAAGAAGGAAAAACAGGGGTCATGGTTTGACTGTCCTGTGAGCCCTTCTTCCCTGCC
TCCCCCACTCACAGTGACCCGGAATCTGCAGTGCTAGTCTCCCGGAACTATCACTCTTTC
ACAGTCTGCTTTGGAAGGACTGGGCTTAGTATGAAAAGTTAGGACTGAGAAGAATTTGAA
AGGCGGCTTTTTGTAGCTTGATATTCACTACTGTCTTATTACCCTGTCATAGGCCCACCC
CAAATGGAAGTCCCATTCTTCCTCAGGATGTTTAAGATTAGCATTCAGGAAGAGATCAGA
GGTCTGCTGGCTCCCTTATCATGTCCCTTATGGTGCTTCTGGCTCTGCAGTTATTAGCAT
AGTGTTACCATCAACCACCTTAACTTCATTTTTCTTATTCAATACCTAGGTAGGTAGATG
CTAGATTCTGGAAATAAAATATGAGTCTCAAGTGGTCCTTGTCCTCTCTCCCAGTCAAAT
TCTGAATCTAGTTGGCAAGATTCTGAAATCAAGGCATATAATCAGTAATAAG